Linking Ontological Classes and Archaeological Forms

Vincenzo Lombardo(✉) , Rossana Damiano , Tugce Karatas ,
and Claudio Mattutino

Università di Torino, Torino, Italy
{vincenzo.lombardo,rossana.damiano,
tugce.karatas,claudio.mattutino}@unito.it

Abstract. Archaeological studies are a trans-disciplinary endeavor, where a number of different scientists collaborate to get a reasonable account of material artefacts, through the various phases of recovery, analysis, and, recently, also exhibition. A large amount of digital data support the whole process, and there is a high value of keeping the coherence of the information and knowledge contributed by each discipline. The paper introduces a modular computational ontology, which is in use in a comprehensive archaeological project, Beyond Archaeology. The ontology provides the information structure to all the phases of the project, from the excavation phase, to the archaeometric analyses, up to the design and the implementation of the exhibition. The computational ontology is compliant with CIDOC-CRM reference model and introduces a number of novel properties and classes to link the description of the archaeological world with the forms traditionally used by the archaeologists to record the excavation and data about findings on the field and in the lab. The forms are implemented through a CMS structured site, for the creation of a data base, that is also filled with multimedia items that are to be employed in interpretation and exhibition, respectively.

Keywords: Archaeology · CRMarchaeo model · CMS

1 Introduction

Archaeological projects are more and more digital, in many accounts, as it happens in many areas of cultural heritage: data collection, curation, and visualization (see, e.g. [6,11], among others), analysis (e.g., GIS [2]), exhibition (starting from the virtual archeological reconstructions of the 1990s [1,9] and addressing general public outreach and participation [10]).

The scientific community of the archaeologists has been realizing the importance of the digital data curation, alongside with physical artefacts. Projects such as the Digital Archaeological Record[1], the catalogue section of the Central Institute of Cataloguing and Documentation of the Italian Ministry of Cultural

---

[1] http://www.tdar.org/.

Heritage[2], and the Archaeology Data Service[3] are archiving and making available a number of archeological data for quantitative testing and processing. These data can be reused by people other than the original creators, in ways that they had not even envisioned before (see, e.g., [12]).

There also are projects that have been carried out with the goal of maintaining the data as long as possible. The Çatalhöyük Database and the Çatalhöyük Image Collection Database[4] make available the documentation of the Çatalhöyük excavation site. These are custom platforms, indeed searchable data management systems, updated during every excavation season, which have been made available through the Çatalhöyük Living Archive[5], an experimental web application that provides access to the data from two decades of excavation and analysis at a Neolithic settlement in Turkey. They also provide an API to query the database.

However, although languages and tools seem to be available and effective, there exist, in general, many limits concerning sharing and standardization of data [3]. A very recent survey made within the AriadnePlus project[6] reports that researchers are not very aware of the issues of data sharing and Linked Data. They also find useful to raise the competence in the alignment of terminologies through the usage of international thesauri (e.g., Getty AAT[7]) and to promote the usage of domain computational ontologies (e.g., CIDOC-CRM collaboration family[8]).

In this scenario, the Semantic Web approach has been invoked to support the sharing of data, particularly in the trans-disciplinary endeavours [5], as in the case of archaeology. Though we agree that a thorough development of the need for data sharing goes with the growth of awareness that is achieved through pervasive data modeling, training, and knowledge (see AriadnePlus report above), we believe that a boost in this direction can come up by the successful implementation of truly trans-disciplinary projects, where research questions arise through the collaboration and peer-to-peer cross-fertilization of several disciplines [8]. Archaeology is an ideal testbed, especially in its multiple relations with archaeometry and laboratory science, philosophy and social sciences, activities on the field (including the negotiation with contractors and the public) and in the university rooms, where "ordering and reconstructing the past" co-exist with "articulating activist political positions in the present" [13].

This paper describes a Semantic Web approach to the conduction of an ongoing EU project named Beyond Archaeology (BeArchaeo)[9], which consists in an

---

[2] http://www.iccd.beniculturali.it.

[3] http://archaeologydataservice.ac.uk/.

[4] http://www.catalhoyuk.com/research/database (last visited on 15 May 2020).

[5] http://catalhoyuk.stanford.edu.

[6] D2.1 Initial Report on Community Needs https://ariadne-infrastructure.eu/wp-content/uploads/2019/11/ARIADNEplus_D2.1_Initial-Report-on-Community-Needs-1.pdf, dated 31 October 2019.

[7] https://www.getty.edu/research/tools/vocabularies/aat/.

[8] http://www.cidoc-crm.org/collaborations.

[9] https://www.bearchaeo.com/ (last visited on 15 May 2020).

archaeological excavation, the consequent archaeometric analyses of the site and the excavated materials, the interpretation of the findings, and the dissemination of the results through physical and virtual exhibitions. The whole project depends on the creation, maintenance, and employment of digital data documentation, that ambitiously supports all the project phases, from the excavation to the exhibitions. The effort aims to overcome some of the limits that have been raised for IT applications in archaeology, which, on the one hand have been appointed to bring, notwithstanding a number of criticisms, some data-driven theory-neutrality to archaeological investigations (together with data recording and visualization), while, on the other, have been appraised as "unrealized 'great expectation"' [7]. In this paper, we introduce the core of BeArchaeo ontology that encodes the conceptual model of the data base. The ontology is publicly available at /purl.org/beArchaeo. In many cases, the ontology classes and properties specialize the entities of the well-known CRMarchaeo model[10] and represent a concrete realization of the application of the ontology from the initial phases of a project. In particular, the ontology captures the entities that are necessary to encode the knowledge that supports the archaeologists in filling the forms that document the excavation and the interpretation phases in an archaeological project. With respect to CRMarchaeo, we have addressed the descriptive issues that are recorded in the documentation rather than the processes that cause the existence of some encountered object.

In the next section, we introduce the BeArchaeo project. Then, we illustrate how we encoded the knowledge about the forms and how it is related to the knowledge of the archaeological entities and processes. Finally, we describe how we have implemented the forms in a CMS structure to allow the filling operation in the field. Some comment on the lesson learnt and conclusions end the paper.

2   The Bearchaeo Project

Project Beyond Archaeology (BeArchaeo) is a RISE European project that consists in the archaeological excavation, archaeometric analyses, interpretation of the findings, and eventually dissemination of the results about the Tobiotsuka Kofun (Soja city in Okayama Prefecture), together with other Kofun burial mounds and the related archaeological material in ancient Kibi and Izumo areas (present Okayama and Shimane Prefectures), in Japan[11]. Archaeologists and archaeometrists (e.g., chemists, physicists, ...) from both Europe and Japan work on a major period (the Kofun period) of the Japan history with a truly transdisciplinary vision of archaeology combined with archaeometry; the project activities and outcomes are accessible to the general public through engaging media communication along the project development and two final exhibitions in Italy and Japan, to be held at the end of the project.

---

[10] http://www.cidoc-crm.org/crmarchaeo/, (last visited on 15 May 2020).

[11] BeArchaeo website https://www.bearchaeo.com/ (last visited on 15 May 2020) and RISE programme https://ec.europa.eu/research/mariecurieactions/news/research-and-innovation-staff-exchange-rise-bridging-ri-sectors-europe-and-worldwide_en.

A preliminary achievement of this research has been the design and implementation of a semantic database for the encoding and storing of the digital data concerning the documentation of the archaeological excavation and the account of the metadata for the several disciplines[12]. We have developed a domain ontology centred around the major classes that appear in the archaeological projects, according to the major forms that are in use, namely the forms for the stratigraphic units and the archaeological findings, respectively. We have analyzed the major documentation sources about the forms in use by the European and the Japanese archaeological teams and we have encoded the related knowledge into the ontology, while keeping the alignment with the CRMarchaeo model. The documentation sources are mostly published by the national organizations of cultural heritage (see, e.g., the documentation records of the Italian Central Institute for the Catalogue and the Documentation[13]). This documentation is rarely related to some shared knowledge source, although in some cases there has been some post-alignment of relevant data bases. For example, the NIOBE database (concerning the Colosseum, the Roman age National Museum, and the Rome Archaeological Area) has been recently mapped onto CRMarchaeo[14] in the context of the ArcheoSITAR project[15].

As far as we know, BeArchaeo is the first archaeological project that assumes a Semantic Web approach from the start. In fact, the multi-disciplinary, multi-cultural, and multi-lingual characters of Be-Archaeo raise a high demand of interoperability of knowledge and data. The alignment with CIDOC-CRM is pursued at the disciplinary level, by aligning the archaeologic and the archaeometric descriptions through the CRMarchaeo and CRMsci models, wherever possible. These issues are particularly relevant in the mapping of the forms to be filled by the researchers onto the ontology classes and properties; so, we designed both a practical workflow and the form interfaces for collecting the data as the excavation goes on, to be continued in the analysis labs, and eventually with the design of the exhibition.

In the next section, we describe the modeling of the BeArchaeo ontology, by highlighting both the encoding of the forms and the alignment with CIDOC-CRM and then we see how it is interfaced on the documentation website.

3   The Modeling of the BeArchaeo Ontology

There have been two major guiding principles in the development of the BeArchaeo ontology. The first is that it should capture concepts and properties in archaeology and archaeometry and how these are connected. The long term aim is a truly trans-disciplinary approach between the archaeologic and

---

[12] https://bearchaeo.unito.it/omeka-s (last visited on 15 May 2020).

[13] http://www.catalogo.beniculturali.it/sigecSSUFE/, in Italian (last visited on 15 May 2020).

[14] http://www.archeositarproject.it/wp-content/uploads/2017/06/Mapping-NIOBE-towards-CIDOC-CRM-final_12.10.2016.pdf, (last visited on 15 May 2020).

[15] http://www.archeositarproject.it, (last visited on 15 May 2020).

the archaeometric disciplines mediated by the formal ontology. There has been an improvement from the 1990s, when the natural sciences were to deliver data to be interpreted by the archaeological theories, to nowadays, when the data emerging from scientific analysis are viewed as more objective and a stable foundation for interdisciplinary analysis [13]. The second is that the ontology must align with the forms currently used by the archaeologists. In particular, the ontology implementation must provide a detailed account of the archaeological knowledge that can lead to the publishing of the record forms that are typically filled by the archaeologists on the field and in the lab, and are the object of a continuous review and interpretation. This, in turn, requires that the interface should recall the traditional forms filled by the archaeologists, in order to match their standard working practices and consequently achieve their full collaboration in the filling of the database. These forms, though sharing a number of features (see below), are usually provided by the individual national institutions, and can be more or less rich. Also, in this project, these forms are, for the first time, augmented with fields that encode the archaeometric analyses, so to achieve a transparent management of the interpretations.

Finally, as an add-on requirement, the ontology must capture features that can support the work of the designers and professionals that will work on the main exhibition, which will be both physical and virtual, at the end of the project. This requirement has been currently limited to the storage of the media (3D models and images) that are associated with the items, together with the algorithms and the procedures used to achieve them.

The modeling has worked in parallel between the encoding already realized by CRMarchaeo and the forms provided by the Italian Ministry of Culture and translated into English for one excavation mission in Pompei. In the following, we introduce the two knowledge sources, the RDF-formalized CRMarchaeo and the forms of the Italian Ministry of Culture, respectively, and then we address the modeling of the BeArchaeo ontology.

3.1   CRMarchaeo

CRMarchaeo, an extension of CIDOC–CRM, was developed to support the activities concerning any archaeological project (actually, mainly the excavation process). Also, it has been growing from standards and models already in use by national and international cultural heritage institutions as well as from the metadata contained in the archaeological documentation. The Conceptual Reference Model is a formal ontology for the integration and interchange of cultural heritage information, which displays an heterogeneous nature. It provides the semantic definitions that underlie the construction of coherent resources, with a super-institutional perspective, to enable semantic interoperability.

The CIDOC-CRM family of models (Fig. 1, top, right) extends the general documentation model through specialised thematic models for the needs of projects and organisations. In particular CRMdig is a model for

provenance metadata, CRMgeo is a model spatio-temporal entities. As super-models of CRMarchaeo, and of particular interest for the BeArchaeo project, are

– CRMinf, a formal ontology about argumentation and inference making, provides a formal description of the semantic relationships between premises, reasoning activities, and conclusions. For example, class *CRMinf/I2 Belief* encodes the fact that some associated proposition set is held to have a particular belief value about some subject on behalf of some actor (e.g., "Italian team believes that Archaeological finding AF29 is of 6th Century AD").
– CRMsci, a formal ontology about scientific observation, measurements and processed data in descriptive and empirical sciences, provides a formal description of the causal relationships in scientific investigations. For example, class *CRMsci/S3 Measurement by Sampling* encodes activities of taking a sample and measuring or analyzing it, in which the sample is typically not identified and preserved (e.g., "S3 Metabarcoding of microbial taxonomic diversity in sample SU202A has observed presence of Rhodosporidiobolus.").

CRMarchaeo, as well as BeArchaeo, take inspiration from Harris' model [4], which takes into account the stratified arrangement of an archaeological excavation. The excavation model includes the description of the dichotomy between the (natural or human) phenomena that produced the stratification (centred around the class *CRMarchaeo/A1 Excavation Process Unit*) and the units that are the outcome of the generation/modification process (centred around the class *CRMarchaeo/A8 Stratigraphic Unit*). Stratigraphic units contain some remains, classified as physical objects (centred around the class *CIDOC CRM/E18 Physical Object* of the core ontology). Stratifications and their contents are analyzed and interpreted to determine the relative chronological order of the strata, together with the classification and functionality of the objects therein, till the high-level reconstruction of the beliefs and behaviors of some group of people in the past in that place.

Figure 1 illustrates the major relationships between BeArchaeo ontology and CRMarchaeo, as well as the reference to the two thesauri (BeArchaeo Archaeological Finding Thesaurus – AFT, for a taxonomy of Japanese history materials, built within the project, and Getty Art and Architecture Thesaurus – AAT). While the class *CRMarchaeo/A8 Stratigraphic Unit* has been imported as it is, the class *bearchaeo/ArchaeologicalFinding* specializes the generic *CIDOC CRM/E18 Physical Object*, in order to connect it to the corresponding catalogue record (see below). We also see that the stratigraphic relation, existing between stratigraphic units, is specialized into several subproperties, as reported by the catalogue record forms below, namely the following spatial relations:

– isEqualTo, for two stratigraphic units that are claimed to be the same;
– isBoundTo, for a stratigraphic unit that is a limit for another one;
– Abuts/isAbuttedTo, for a stratigraphic unit that edges another one;
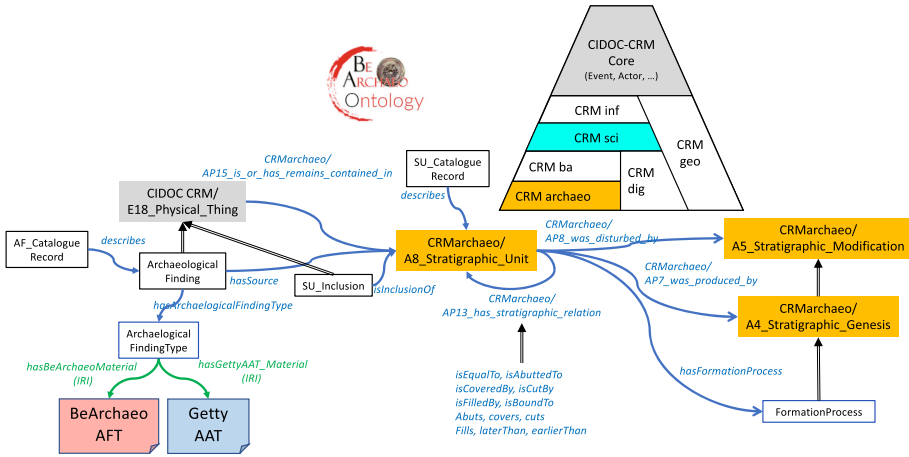– Cuts/isCutBy, for a stratigraphic unit that introduces a discontinuity into another one;

Fig. 1. Major relationships between BeArchaeo and CIDOC-CRM. Colors are employed to distinguish provenances. (Color figure online)

– Covers/isCoveredBy, for a stratigraphic unit that covers (stands over) another one;
– Fills/isFilledBy, for a stratigraphic unit that has filled a cut (see above);

Also, there are two temporal relations, laterThan and earlierThan, resulting from the interpretation of the stratigraphy.

3.2 The Archaeological Forms

When on the field as well as when in the lab, archaeologists fill forms that are prepared by the national authorities to track all the entities that have been recognized and to update the interpretations of the findings. So, it is important, for the practical application of the encoded ontological knowledge, that knowledge and forms are connected. The solution devised in BeArchaeo has been to also encode the form fields as properties of the ontology, developed as modules that included the archaeologic knowledge, the archaeometric knowledge, and the catalogue record knowledge (see below). The forms we have encoded, with some adjustments after long and productive discussions with the archaeological team of the project, have been the ones distributed by the Italian Ministry of Culture, and in particular, the forms of the Stratigraphic Unit and the Archaeological Finding.

In Fig. 2 we see an excerpt of the Stratigraphic Unit form, in an English translation (for the sake of understanding). The upper left part is the registry part of the Stratigraphic Unit, reporting identifiers and locations; then going down, after the informal definition and position, we find, among others,

– the distinguishing criteria (three-valued multiple choice), which were employed by the archaeologist to identify such a stratum in the soil;

Fig. 2. The form provided by the Italian Ministry of Culture, in the English translation provided by an excavation mission carried out by the universities of Missouri and Mount Alison in Pompei (Courtesy of Ivan Varriale).

– the formation process (connected to the genesis or modification of the unit), which can be valued with a number of common processes, together with the possibility of free insertion ("other", in the form);
– the type of soil matrix, which can also be a combination of values;
– the inclusions, i.e. the generic types of physical objects found in the unit (a list is provided, together with the freedom of some further insertion), with their (three-valued) frequency of occurrences;
– the (five-valued) compaction attribute;
– the color of the unit.

The form for the Archaeological Finding record that we have taken into account is an extract from the very articulate documentation reported in the web platform of the Italian Ministry of Culture, named SigecWeb[16]. The reduction was due to the fact that the original format is designed to track the existence of the finding in passing through various institutions during its entire life cycle (restoration centres, museums, churches, ...), which is possibly very long and departs from the goals of BeArchaeo. The major fields of the record concern:

– the source, i.e. the stratigraphic unit that contains it (actually a subproperty of CRMarchaeo property *AP15 is or contains remains of*);

---

[16] http://www.sigecweb.beniculturali.it.

– the type of the finding (in terms of materials and functions), which we implemented as a reference to the two thesauri mentioned above;
– an indication of the guessed (or confirmed) chronology, together with a motivation for it.

Now we see how both the CRMarchaeo model and the archaeological excavation forms have contributed to the BeArchaeo ontology.

3.3 The BeArchaeo Ontology

The BeArchaeo ontology is geared to the description of the objects rather than the forming processes, and merges the general classes and properties provided by CRMarchaeo with the fields of the archaeological catalogue records. We did not employ specific design patterns because we were not aware of patterns for connecting knowledge and forms, and because our solution was straightforward (see below). The result is an application ontology that connects three types of knowledge: the archaeologic knowledge (lower left part of Fig. 3), the archaeometric knowledge (lower right part of Fig. 3), and the catalogue record knowledge (upper part of Fig. 3).

Figure 3 provides an overview of an instantiated knowledge, where the rectangles in grey or black are the individuals, and the white rectangles are the classes; object properties are depicted as blue lines, while datatype properties are depicted as green lines; the three elements in Courier font, highlighted in yellow, are the strings that are actually written in the final form interface. Going left to right: the stratigraphic unit "SU 202" (content of the title field of the catalogue record for this unit) is the source of the archaeological finding "AF 59" (content of the title field of the catalogue record for this finding); the type of the finding is "Sue (ceramics style)", as selected from the Getty AAT thesaurus and "sekki", as selected from the BeArchaeo thesaurus; the finding body[17] has undergone some chemical test, which has produced a composition descriptor (the individual is actually a table reporting the presence of substances), which in turn is input to a Data Evaluation process (related to CRMsci ontology), which assigns some dimension, namely an attribute for the body composition ("Calcareous").

In the Figs. 4 and 5 there are the major relations of the stratigraphic unit class and the archaeological finding class, respectively. Going clockwise, a stratigraphic unit has inclusions (i.e., entities that are contained in stratum), which are of some type, that can be generic or specific, and has a frequency of occurrence in the unit, qualitatively valued as rare, medium, or frequent. Inclusions have types that are taken from partially overlapping vocabularies, based on the practical experience of the archaeologists (these may change and should be aligned with the types included in the thesauri for the archaeological findings). Some informal properties, noted as free text, are the state of preservation of the unit

---

[17] Usually, for chemical tests, an archaeological finding is considered as composed a body, a coating, and an embellishment.
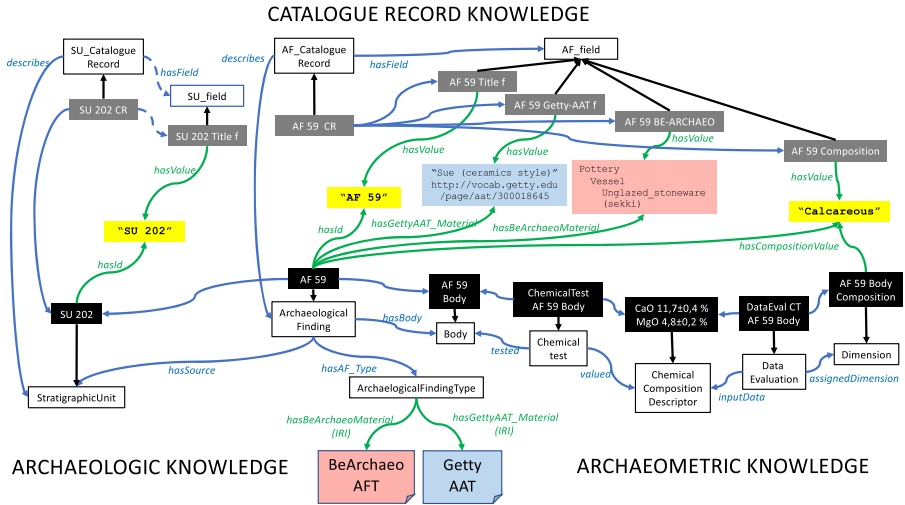
Fig. 3. Modeling of the archaeological finding, exemplifying archaeologic and archaeometric knowledge, respectively, and the corresponding fields in the archaeologic finding record.

and the measurements taken during the excavation, with a particular concern for Elevation. The distinguishing criterion determines how this unit has been identified: the terms that concern this attribute are three (Color, Composition and Compaction) and there are other three properties that possibly specify the actual values for such attributes (namely 6-valued soil/matrix term for composition, 5-valued term for compaction, and a free string for color). Color, in the relationship with archaeometrists (specifically, the soil scientists) has been augmented with the encoding provided by the well-known Munsell color system, in use in pedological studies[18]. Finally, the formation process concerns a specialization of the processes that are responsible for the creation and modification of the unit, with a frequent term vocabulary, which can be further augmented with free text insertion. The properties in the center of the figure specialize the stratigraphic relations, in spatial and temporal terms (see above).

An archaeological finding (Fig. 5) can be part of another archaeological finding (frequent is the case of fragments to be composed afterwards) and is sourced by some stratigraphic unit as well as museum collection or other places. This variety of sources concerns the goals of the BeArchaeo project, because of the employment of the ontology into the design of the final exhibition. The archaeological finding has a type, referring to terms in the widely acknowledged Getty

---

[18] Munsell color system is based on the three-dimensional model, where each color is defined by a triple of hue (the color of the color), value (how light or dark is the color), and chroma (or saturation/brilliance of the color), set up as a numerical scale with visually uniform steps https://munsell.com/about-munsell-color/how-color-notation-works/.
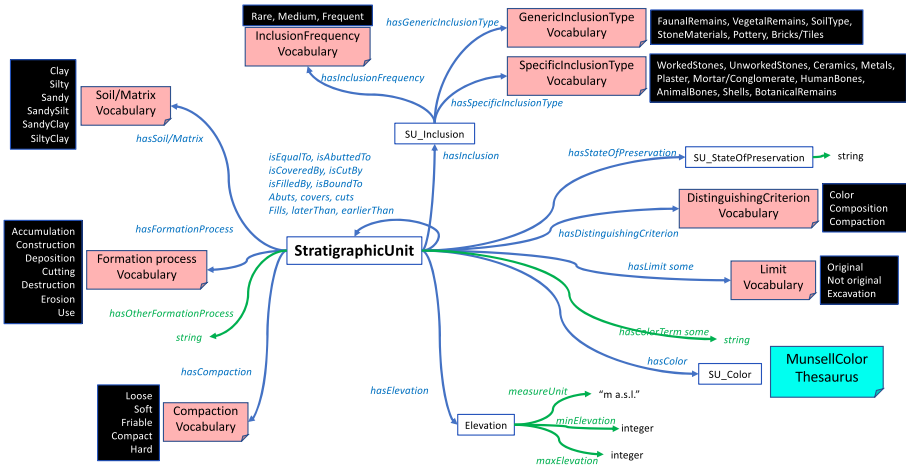
Fig. 4. Modeling of the stratigraphic knowledge (including references to thesauri and vocabularies (with list of terms)).

AAT thesaurus and the BeArchaeo AF thesaurus, the latter encoding knowledge from an authoritative Japanese reference [14]. Finally, an archaeological finding is marked with its chronology, currently limited to a free text insertion, together with its motivation, but with the idea of linking to a time ontology in the LOD panorama.

From a technical point of view, the model has been described as a number of subontologies that address different sections of the forms. In particular, there are five subontologies for the stratigraphic unit catalogue record: "registry" (which contains identifiers and formal issues concerning the location, the trench, the section, ...), "description" (concerning inclusions and soil attributes), "stratigraphy" (concerning the relations with other stratigraphic units), "dating" (where dating elements and chronology are represented), "sampling" (concerning some data about the excavation process). Then, there is one ontology for the archaeological finding record. And, finally, there is the encoding of the archaeological knowledge. The several subontologies for the sections are connected to the main ontology for the record through the properties *hasField* and *hasSection*, while the ontologies for the records are connected to the archaeological knowledge are connected through the property *arco/describes*, as introduced by project ArCo[19] for the relationship between an entity that describes another entity in the field of cultural heritage. The ontology is expressed in OWL/RDF formats and published at the permanent address /purl.org/beArchaeo[20].

---

[19] http://wit.istc.cnr.it/arco/.
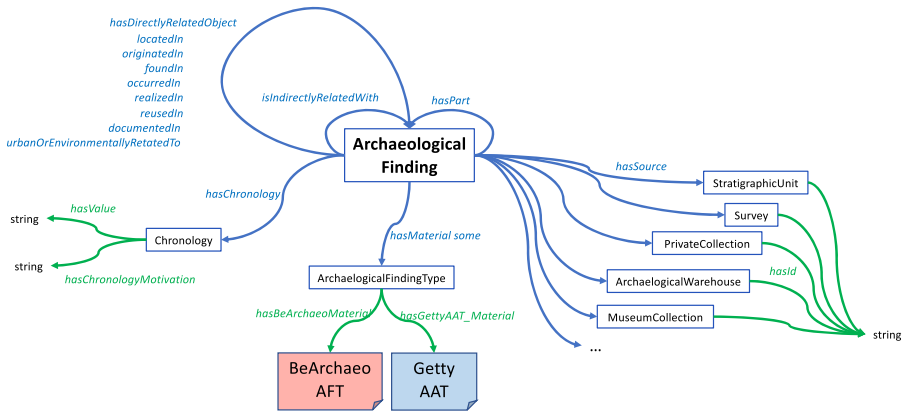[20] File "BeArchaeo_merge_all.owl" merges all the other sub-ontologies.

Fig. 5. Modeling of the archaeological finding. Exemplifying archaeologic and archaeometric knowledge, respectively, and the corresponding fields in the archaeologic finding record.

4 Ontology BeArchaeo in Use: CMS Approach to Form Filling and Lesson Learnt

In this section, we describe how the ontology has been employed for the excavation campaign carried out by the BeArchaeo team in August 2019 and reported on the project website[21]. In order to make the knowledge available to the archaeologists on the field, we built a website, based on an installation of a Content Management System (CMS), for achieving an immediate deployment of the forms. The CMS Omeka-S[22] is particularly suited for the import of semantic properties defined in a RDF file, the definition of customized vocabularies, and the construction of templates for the instantiation of filling forms. Also, from the inserted items, one can easily build a website for immediate check of the data base content, sharing of the data, and the development of specific functions, based on the native API.

We exploited the possibility of the fast prototyping of a user interface for the back-end of the system, accessible by the people on the field, and a quick front-end, where supervisors and stakeholders could explore the development of the archive and the related findings. Also, we have started uploading a number of rich media materials (currently photos and 3D models acquired from photogrammetry and scanning), that are being used for interpretation as well as will be the base for the final exhibition. Figure 6 reports two images, from the back end and the front end, respectively, of the production website[23]. The archaeologists have used the back end on the field, introducing data through tablets (stratigraphic units and archaeological findings), and afterwards in the labs through laptop

---

[21] https://www.bearchaeo.com.
[22] https://omeka.org/s/ (last visited on 15 May 2020).
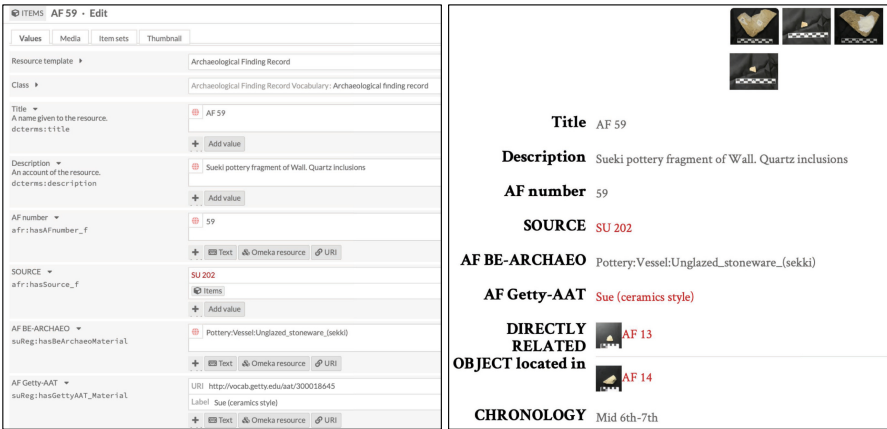[23] https://bearchaeo.unito.it/omeka-s.

Fig. 6. Screenshot from the BeArchaeo resources website, concerning the Archaeological finding no. 59, with the related fields and media. On the left, the back end; on the right, the front end. Elements in red are links to other elements of the documentation (e.g., Stratigraphic Unit 202) or to some external knowledge source (e.g., Getty AAT thesaurus).

and desktop computers (archaeological findings). They found the tool useful, especially for the digital support (usually, archaeological teams without an IT support notate item data on paper and then transfer data on Excel files); the organization provided by the CMS to split into Authors, Reviewers, and Editors, the roles of the archaeologists with respect to the platform has eased the check and revision of the inserted data.

Differences in interpretations have been annotated through the possibility offered by Omeka-S to insert more than one value for a property; this also happens with functional properties, and should be solved through the reasoning operated by the semantic server. Also, fields with free text have allowed to include motivations for the choices made and thus enable the communication within the team. Indeed, this feature of a centralized database has been particularly appreciated by the team, and editors have made a specific pressure on the rest of the team for an extensive use the platform. The database-based approach has been useful to start the collaboration with the archaeometric team in the interpretation process. Indeed, some archaeometric analyses, such as the geophysical ones and the biological ones, have started before the archaeological excavation process. However, the encoding of the archaeometric knowledge, which requires a mapping with the CRMsci model, is under development.

The splitting of the five sections for the stratigraphic units, as structured in the original forms, has been seen as a complication of the filling work: so, the archaeological team have required for a unique form to be navigated through scrolling for the access to a large number of fields. This will require some

programming effort to build an external website (with a specific connection with the database) and will be released in the next version of the platform. Also, the interface to the thesauri (both Getty AAT and BeArchaeo AFT) requires some modification. It is currently implemented as a keyword search/completion, in one case (Getty AAT), and a drop-down menu, in the other (BeArchaeo AFT). In particular, the terms are compiled from a number of features, namely Material, Object function, Sub-material, Object generic class, subclass, shape, subshape. For example, the following path can be built for a specific finding:

```
-- Metal
---- Objects Facet
------ Bronze (d ki)
-------- Weapon (buki)
---------- Sword (katana)
------------ Long sword (tachi)
-------------- Sword with a bulbous pommel (kabutsuchi no tachi)
```

The drop-down menu navigates the hierarchical classes by proposing allowed combinations of feature values; also, sometimes not all the features are overtly expressed. Archaeologists have found some difficulties in the navigation and prefer an interface arrangement that allows for the direct setting of the specific feature values. In this case, the system should then propose some terms from the thesaurus that are consistent with the setting provided. Again, a programmatic solution to be devised in the future.

Finally, two notes concerning the context of the BeArchaeo intercontinental project, because of the different excavation techniques that pertain the two traditional schools of archaeology and the linguistic issues for the interfaces. In the first case, we have that similar terms, such as trenches, sections, and rooms of the excavation call to slightly different definitions according to the two traditions; so, the interface must accomodate both methods and interpretations. At the current stage of development, the two teams are still looking for a common arrangement and the ontological vocabulary will be updated accordingly. In the second case, the interface in English was a limitation for the Japanese archaeologists in the form filling process, because the data insertion process could not be done in the native language (also because of the different interpretations above). On a development site[24], we are experiencing a number of innovations in preparation of the second excavation campaign (originally scheduled for August 2020, but now, due to the pandemic, postponed to 2021). In particular, we are addressing the encoding of the forms into Japanese: there are some Japanese resource templates for the Archaeological Finding and Stratigraphic Unit records, respectively, as well as a front end website in Japanese[25].

5   Conclusions

The paper has described an ontological approach to the encoding of the archaeological knowledge, in its relation with the archaeometric knowledge and the

---

[24] https://bearchaeo.di.unito.it/omeka-s.
[25] https://bearchaeo.di.unito.it/omeka-s/s/jtoppage/page/welcome.

practical forms for filling the archaeological/archaeometric information on the field and in the lab. In particular, we have presented how we have encoded an ontology of archaeological knowledge that is compliant to CRMarchaeo ontology, which is in use in a EU project concerning an excavation process in Japan. Also, we have seen how the ontology is the base for a CMS-based web platform for supporting the archaeologist's work in recording the excavation and interpretation activities. The encoding of the archaeological knowledge in an ontology that is compliant with CRMarchaeo, and so to CIDOC-CRM, and the implementation of a CMS-based solution for a concrete project can have a deep impact on fostering projects that adhere to the Semantic Web paradigm and address data sharing effectively. The publication of the ontology and the availability of a widespread CMS can be easily replicated in further projects. The ontology and the derived database will also be employed in the definition of the contents of the exhibition that will present the outcomes of the BeArchaeo project to large audiences in collaboration with museum institutions in Europe and Japan.

The BeArchaeo ontology has also been extended with the investigation processes and the related outcomes that concern the archaeometric part. However, we are working on how to build the interface forms to have them operational on the field and mostly in the labs afterwords. Also, we are going to connect the ontology with other resource for the cataloguing of cultural heritage assets (e.g. the Knowledge Graph for the Italian cultural heritage ArCo. We also want to improve the ontology interoperability, by replacing a number of customized vocabularies with domain ontologies (e.g., for chronology and formation processes).

6   Author Statement and Acknowledgements

References

1. Barcelo, J., Forte, M., Sanders, D.: Virtual Reality in Archaeology. ArcheoPress, Oxford (2000)
2. Conolly, J., Lake, M.: Geographical Information Systems in Archaeology. Cambridge University Press, Cambridge (2006)
3. Costopoulos, A.: Digital archeology is here (and has been for a while). Front. Digit. Humanit. **3** (2016). https://doi.org/10.3389/fdigh.2016.00004
4. Harris, E.: Principles of Archaeological Stratigraphy. Academic Press, London (1989)
5. Lampe, K.H., Riede, K., Doerr, M.: Research between natural and cultural history information: benefits and IT-requirements for transdisciplinarity. ACM J. Comput. Cult. Heritage **1**(1) (2008)
6. Lercari, N., Shiferaw, E., Forte, M., Kopper, R.: Immersive visualization and curation of archaeological heritage data: Çatalhöyük and the Dig@IT app. J. Archaeol. Method Theory **25**(2), 368–392 (2017). https://doi.org/10.1007/s10816-017-9340-4
7. Niccolucci, F., Hermon, S., Doerr, M.: The formal logical foundations of archaeological ontologies. In: Barcelo, J., Bogdanovic, I. (eds.) Mathematics and Archaeology, pp. 86–99. CRC Press, Boca Raton (2015)
8. Nicolescu, B.: Methodology of transdisciplinarity - levels of reality, logic of the included middle and complexity. Transdisciplinary Eng. Sci. **1**(1), 19–38 (2010)
9. Reilly, P.: Towards a virtual archaeology. In: Lockyear, K., Rahtz, S. (eds.) Computer Applications in Archaeology, pp. 133–139. Oxford: BAR 565 (1990)
10. Richardson, L.: A Digital Public Archaeology? Papers from the Institute of Archaeology. UCL, London (2013)
11. Roosevelt, C.H., Cobb, P., Moss, E., Olson, B.R., Ünlüsoy, S.: Excavation is digitization: advances in archaeological practice. J. Field Archaeol. **40**, 325–46 (2015). https://doi.org/10.1179/2042458215Y.0000000004
12. Silva, F., Linden, M.V.: Amplitude of travelling front as inferred from 14c predicts levels of genetic admixture among European early farmers. Sci. Rep. **7** (2017). https://doi.org/10.1038/s41598-017-12318-2
13. Stutz, L.N.: A future for archaeology: in defense of an intellectually engaged, collaborative and confident archaeology. Norw. Archaeol. Rev. **51**(1–2), 48–56 (2018). https://doi.org/10.1080/00293652.2018.1544168
14. Tadanao, Y.: Dictionary of Japanese Archaeological terms. Tokyo Bijutsu Publishing, Tokyo (2001)