The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage
VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

# Ontological Modelling for Archaeological Data

Andrea D'Andrea[1] Giulia Marchese[2] and Tommaso Zoppi[2]

[1]CISA- Università degli Studi di Napoli L'Orientale
[2]PIN Scarl PRATO

**Abstract**

*The need to investigate an archaeological context causes the inevitable destruction of the upper strata in favour of the lower ones. On-site research activity therefore translates into a documentation activity (forms/photographs/surveys), which must be as neutral as possible and not necessarily biased by the scientific interests of the archaeologist. This documentation activity has recently been subjected to an in-depth analysis and evaluation. A wide spectrum of standards and reference regulations is being defined by different national and international Organisations. The various attempts made at normalizing the production of documents of excavations have lead to the elaboration of a large number of forms. The introduction of computer science in the management of archaeological records has actually complicated the picture because it has added different formats, software and operating systems, chosen by each individual researcher. Considering how dangerous it is to convert older data to newer digital formats, as is any translation from one language to another, we started a project aiming at defining an ontology able to guarantee interoperability between different archives without modifying, altering or sacrificing the archives created by each archaeologist. The primary objective of our research has been to analyse the Italian documentation produced during the stratigraphical excavation, which represents the most consistent corpus of data available also in digital format. We chose the CIDOC-CRM because it is event-oriented. In terms of content, the archaeological documentation activity may be easily schematized: it documents a past event occurring during an archaeological era and, at the same time, it documents the action of the modern-day scholar. Any excavation activity and its pertinent methodology may be easily described following this conceptual formalism.*

Categories and Subject Descriptors (according to ACM CCS): I.2.4 Knowledge Representation Formalisms and Method [Artificial Intelligence]: Representation (procedural and rule-based)

## 1. Introduction

Excavation is an activity aimed at investigating an archaeological context, both horizontally at the open-area level, and vertically at the depth level. It causes the inevitable destruction of the upper strata (even when dealing with monumental structures) in favour of the lower ones.

Archaeological excavation is distinguished by two types of activities in the field and a successive one in the laboratory. The first two are characterized by the actions of excavation and by the documentation of all procedures carried out in the field, while the third one is performed in the laboratory, often when the excavation activity has already been completed. The documentation activity is carried out during and after the digging action, often starting with the filling of forms on field and terminating in the laboratory after the analysis of all the documents produced (photos, drawings, reports, forms and findings).

From a methodological point of view all these activities can be altered by the experience of each individual researcher, by local archaeological tradition or by the type of archaeological context investigated (tombs, monumental structures, surfaces, etc.). Thus it is possible to change the way to carry out the excavation, adjusting the methodology and research strategy to the specific archaeological area.

The archaeological survey in the field is a process involving the destruction of the data required by the archaeologist in order to reconstruct a past event. In order to avoid losing important information it is necessary to accompany all the ground investigations with other types of documentation activities. So, while an archaeologist digs, other archaeologists or technicians are involved in the "translation" of the bulk of materials coming from the excavation into written, photographic and graphic documentation.

On-site research activity therefore translates into the documentation (forms/photographs/surveys) of all the activities carried out in order to allow the archaeologists to provide all the information indispensable for analysis and

study at a later time. Irrelevant information is often stored as well, as archaeologists are not always really conscious – during excavation – of the content of their investigations.

This descriptive process, called archeaography [Mob81] permits other scholars to re-use the documentation gathered during fieldwork for in-depth studies and perhaps for new hypotheses.

The documentation activity should therefore be as neutral as possible and not exclusively related to the scientific interests of the archaeologist. But as I. Hodder [Hod99] highlighted "*…whether an object in the ground has any chance of becoming an 'archaeological object' depends on the perspectives and methods of the recovery process*". All data collected on-site are necessary for the archaeologists, but only certain raw data become archaeological information through the filter of experience or scientific sensitivity of the researchers.

For these reasons the documentation activity - both for the excavation and for the documentation preceding the on-site investigation - has been in recent years the object of in-depth analysis and evaluation [RB96)].

Documentation activity of all the actions carried out in the field has a long history in the archaeological tradition and perhaps arose with the birth of modern archaeology. From ancient reports, composed of "artistic" drawings and photos, to the modern technologies used to collect digital data directly on field, the history of documentation has known many different strategies and goals. It is possible however to state that it is precisely the documentation activity to have given a scientific status to modern archaeological investigations. From the first paper reports/forms filled by archaeologists, which can represent a sort of history of the sensitivity of the archaeologists on field and their methodological strategies, we've now moved on to a sophisticated way of recording the archaeological process. I. Hodder [Hod00] has suggested using a video recording system during the excavation to improve not only the techniques of documentation of all the actions, but also to develop a "reflective" stance.

**2. Standards and "Best Practices"**

Nowadays, in order to contain the excavator's natural tendency to be a narrator of stories (often already present in the mind of the archaeologist), rather than a neutral excavation "technician", a wide spectrum of standards and reference regulations are being defined by different national and international organisms [ads.ahds.ac.uk/arena/links/standards.html].

The main goal of these standards is to normalize the production of excavation documents through the elaboration of a large quantity of forms. These have been designed to record the overall activities of an archaeological excavation (be they structures or objects, finds, artefacts) thus guaranteeing a further analysis of the documentation. All the forms are really accurate and designed to document in a homogeneous way the information gathered on field, without being influenced by the sensitivity and experience of the archaeologists and their scientific targets.

It is possible to distinguish various types of standards which only in certain cases may also cover the problems regarding the management of an archaeological excavation:

- **Catalogue** Standard: it determines the rules for cataloguing various objects;
- **Standards** regarding the terms used to describe an object (for example vocabularies and thesauri);
- **Metadata** Standard: it gives structure to information making it interchangeable independently from the databases used.
- **Interchange** Standard: it allows communication between computers.

The simplest and most immediate type of standard is the "thesaurus". It characterizes archaeology almost as its foundation and disciplinary premise, though we can't really speak of a general reference (and generic) vocabulary, as it is connected to specific sectors (prehistoric, classic, oriental, medieval archaeology, etc.) which are divided into sub-disciplines that further specialize the different thesauri (Etruscology, Greek archaeology, the Iron Age, etc.). At this level the primary goal of a standard is to guarantee "homogeneity" in the description of any object, a sort of single description protocol with values defined by pre-fixed lists (vocabularies, thesauri, dictionaries) chosen by each archaeological team according to specific needs.

Many standards devised especially for the management of museum collections exist [http://www.fish-forum.info]. These are often implemented to guarantee a rapid inventory of the objects found during an archaeological investigation in the field. Thanks to widespread computerization, these standards allow museums to be efficiently automated and kept up-to-date for any future developments in the sector.

Notwithstanding the release of a number of standards, often widely distributed at the national level, the archaeologist's work in the field has not changed substantially. On the contrary, the university teams, often released from the obligation to produce and deliver "standard" documentation, make use of a variety of different methods and systems for the recording of information.

For this reason, the initial goal of the documentation activity, that is to guarantee a sort of scientific interoperability of data, failed. Nowadays there are too many different standards and information systems that compete against each other, all aimed at recording the archaeological excavation. All the attempts made to "create" an interoperability about the structures and the content collided with many, problems not easy to solve: it is impossible to figure out a sort of super-form able to synthesize all the various forms. In the same way it is hard to foresee a single dictionary/thesaurus/vocabulary on which all the archaeologists agree, as it is to oblige archaeologists to use a specific system or format or platform. The design of a unique methodology effective for all the researchers should also be foreseen.

The recent introduction of computer science has complicated the picture because it added different formats, software and operating systems, chosen by each individual researcher.

Although the passage from Relational Databases to "mark-up language" made the data normalization process more flexible, thus more "aware" of the quality of the information, rather than the quantity, it hasn't always been possible to understand how the researcher's ability has lead

him/her to the classification/interpretation, thus to the formalization of the data. This difficulty emerges even more clearly if one thinks of the need, common to most archaeological investigations, to re-use previously compiled documentation, often formalized according to systems and excavation methodologies which differ from those based on the principles of archaeological stratigraphy. Each new archaeological hypothesis will be built upon the documentation previously acquired during on-site investigation. It is easy to figure out how each historical and graphical (3D or virtual) reconstruction is influenced by the quality of the recorded documentation.

For these reasons, in order to avoid the implementation of a new system incapable of communicating with other systems and archaeological archives, we decided to experiment the use of the knowledge-based system in order to understand and formalize the conceptual model underlying the design of the archaeological documentation.

The absence of interoperability, caused by the diffusion of different platforms, programs and formats, is determined mainly by a profound semantic diversity, only partially solved by the definition and circulation of thesauri and thematic dictionaries. Perhaps what still impedes a real and true interoperability between on-line resources, the portals and thus the data is the absence of a clear formal representation of knowledge-based models that are at the basis of the normalization and computerization of information.

### 2.1. Italian ICCD Standard

The reference model for cataloguing the Italian archaeological heritage was published in 1984 by the Istituto Centrale per il Catalogo e la Documentazione (www.iccd.beniculturali.it), an institute of the Ministry of Cultural Heritage. It refers to a limited number of forms: Stratigraphic Trench (SAS), Stratigraphic Unit (US), Archaeological Find (RA), "Wall Covering" Stratigraphic Unit (USR), Paleo-Anthropological Remains and Archaeological Monuments (MA).

As one may read in the foreword of the Normative that ICCD released in March 2004 (*Sistema Informativo Generale del Catalogo* - Normativa 3.00) [Nor04]: "*…the intense job of systematization of the entire catalographic process in its methodological and operative aspects… has entailed an accurate revision of the most frequently used catalographic forms and of the regulations regarding their compilation… To these requirements, of a technical, practical and operative nature, linked to the refinement and natural evolution of a normative occurring in its emerging aspects over a ten-year period of time, one must add those made necessary by the current and delicate phase of data diffusion which must be reconciled with the principles of privacy, safeguard and intellectual property; such demands have rendered necessary the insertion of appropriate fields in which the institutions responsible for the catalographic procedure are necessarily involved regarding the "sensitivity" of the information and the consequent differentiated access according to the user profile*".

The activity of the ICCD however is not limited to the definition of forms, but also of the format standards in order to guarantee an optimal management of all the available resources. Through the years, along with a consistent investment of resources for the planning and implementation of forms to be compiled according to formalized ways of describing objects (vocabularies), a reflection regarding format standards has also developed. In reference model 3.00 [Nor04] there is a section reserved for multimedia applications standards, which underlines the advantages of using widely available commercial products and which are characterized by the use of interchangeable formats (*.tiff, *.jpeg, *.dwg, etc.) that allow for some interoperability between different sources.

Notwithstanding the coordinating action of the ICCD in promoting and favouring the adoption of format standards (dictionaries and forms) and of digital support (file types), the activity of documentation in universities has moved along different lines due to widespread computerization of the archaeological excavation. Proprietary programs have been implemented for specific research projects (GIS solutions, Database, Multimedia tools, etc.). The absence of a common framework determines an important issue not easy to solve: the migration and the integration of these proprietary data and "standards" in a unique system. This situation becomes more complex (and difficult to avoid and overcome) when different teams, using a variety of recording systems and standards, work in the same archaeological areas. Unfortunately in this case, as in other similar conditions, the integration is carried out at the level of historical synthesis and not of raw data. So it is impossible re-use this *oriented-data* in order to propose new historical reconstructions and hypotheses.

### 2.2. Syslat

Alongside standard systems and national regulations for the formalization of archaeological documentation, a number of informatics systems are available, often implemented by a single excavation or research team. Their objective is to simplify the process of recording, filing and researching of the information stored in the document archive (forms, images, drawings, etc.).

During the excavation of Cumae, the most ancient western Greek colony, the University of Naples l'Orientale adopted the Syslat registration system since 1994.

Originally conceived as a tool to record excavation data of the proto-historic site of Lattes (Montpellier – France), since 1984, when it was first experimented, Syslat (Système Lattes) underwent several releases that progressively assimilated the suggestions deriving from use. Outside France, the system has been used in the investigation carried out by the University of Naples l'Orientale in the ancient Greek site of Cumae. The experimentation created a portable system provided with a wide set of personalization functions, which can be configured according to the needs of any excavation.

Syslat is not only a translation in digital format of the stratigraphical excavation model formalized by Harris. Even if it is based on the fundamental principles of stratigraphy, it is a tool for a "guided" organization of field data (from recording of actions to graphical and photographic documentation and sample collection) and for an integrated management of all data, including statistic-quantitative analysis of materials.

Syslat represents a great container in which appropriate scripts allow access to recorded information: from

stratigraphical units to "fact" and "set" records, from photos to graphic archives, from quantification records of materials to the typology of individual ceramic finds and so on. The global archive is structured in five different modules (terrain, objects, samples, documentation, utility) from which additional hierarchical sub-levels may be accessed.

The system includes an ample iconographic dictionary to classify ceramics and several vocabularies guide the operator in compiling forms. Syslat also has a module to personalize the database and to add new definitions to glossaries and dictionaries.

Some years ago we realized the porting of the tool into an interoperable framework, converting all data gathered during the excavation of Cumae in XML, and we implemented a customized management system [DN02] using an open-source program.

## 3. A Preliminary Ontology for Archaeological Data

### 3.1. Introduction

By comparing the two standards (ICCD and Syslat) we immediately found out that, apart from a few items and a kind of in-depth fields for the description of the archaeological context, these two systems were different for the conceptualisation and organisation of data: for instance, the numeration system assigned to the stratigraphical units. While the former is similar to the traditional Harris system, with a numeration not linked to the interpretation of the excavation, the latter vice-versa relies on the sub-division of the investigated area into "zones" corresponding to the ancient organization of the site (blocks, quarters, streets, functional areas).

Moreover these two systems used different vocabularies (one in Italian and one in French) and many different items corresponding to the sensitivity and the experiences of each institution. For example while the ICCD system offers a greater wealth of descriptions for the identification of stratigraphic relations, in particular of wall structures. Syslat is based only on anteriority, posteriority and contemporaneity, the three fundamental characteristics of the relations between US.

A further element of differentiation is given by the aggregation of US into groups. The ICCD is based on the grouping of archaeological levels according to functional units, which should correspond to sets of actions and activities referable to a single function (habitation, house, private/public building, sacred area, etc.). Syslat introduces instead an intermediate element, the fact, whose function is to schematize the construction process of the matrix based on the simple observation that some actions, although split amongst more than one US, represent in any case a single action: for example, trench and fill, wall and foundation, etc.

Considering how varied the methodology used to investigate an archaeological area is, we immediately abandoned the idea of integrating the metadata of the two systems into a single database. The differences were in the conceptual models used to dig and to collect data, which followed a different strategy, the former technical, the latter interpretative, and not in the software or in the design of database or the management system.

Furthermore considering how dangerous it is to "convert" older data to newer digital formats, as is any translation from one language to another, we started a project aimed at defining standards that guarantee interoperability between different archives without modifying, "altering" or sacrificing the archives created by each archaeologist. We wanted at the same time to maintain not only the data and the documentation collected, but also the methodology utilized in order to better understand how the data was created and then processed.

Consequently our project does not consist of recognizing and underlining the fields in common with the different structures present in the forms (for example: place, location, period, chronology, phase, stratigraphical relations, etc.) for it might appear as a simple definition of metadata. Instead it consists of "extracting" from each definition a representation of the conceptual model the archaeologist has referred to during his/her fieldwork and which translated into documentation. The objective is not limited to attempting the integration of multi-temporal and stratified databases, but rather to the necessity of "comprehending", thus of "representing" in a "transparent" manner, the processes carried out by the archaeologist in his/her own knowledge-based domain (stratigraphical excavation, open-area excavation, in-depth trench excavation, artificial strata excavation).

To guarantee interoperability between different repositories (formalized using different standards) the first step should be to have a common standard. It is impossible however to force archaeologists to use the same protocol or standard, thus sacrificing their point of view.

To avoid constructing a new standard for the entire community of archaeologists (a super-standard), we decided to work on a new approach using an ontology to understand the work of the archaeologist on field. Other proposals, with our same scope, failed because they were based on the idea of a common distributed infrastructure to guarantee the interoperability among archives [http://www.progettodice.it].

The ontology was deemed particularly useful to formally "describe" the archaeologist's activity; the ontology is synthetically defined as formal cognitive models in a certain domain. In this perspective it is possible to implement a system based on the common ontology in order to integrate different repositories created using Database or other systems into one management system.

We decided to formalize these methodologies using the CIDOC-CRM in order to highlight the semantic connections between these two recording and documenting systems.

### 3.2 CIDOC-CRM

The CIDOC-CRM (http://cidoc.ics.forth.gr) [CDG*05] is an ontology created in order to offer "definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation". "The CIDOC CRM is intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework that any cultural heritage information can be mapped to. It is intended to be a common language for domain experts and implementers used to formulate requirements for information systems and to serve as a guide for good

practice of conceptual modelling. In this way, it can provide the 'semantic glue' needed to mediate between different sources of cultural heritage information, such as those published by museums, libraries and archives". After about 10 years the CIDOC-CRM has become an ISO recommendation, while a definitive elaboration is expected from the technical committee ISO/CD 21127.
Even though the CIDOC-CRM is composed of over 80 classes and 110 properties, it may reduce to a "light" meta-scheme (diagram) that has at its core a correlated event, through bi-directional property, to other classes: Object, Subject, Place and Time. Such a simplified scheme (diagram) works with a rapid integration of data with different formal structures. So CIDOC-CRM is a sort of *Top-level ontology* describing very general concepts like space, time, matter, object, event, action, etc., which are independent of a particular problem or domain [Gua98]; different kinds of ontology according to their level of generality.

Although this knowledge-based model has been implemented especially for the documentation of museum collections, CIDOC-CRM may also be adopted to describe the documentation gathered in the course of field investigations. We chose the CIDOC-CRM because it is *event-oriented.*

In terms of content, the archaeological documentation activity may easily be schematized: it documents a past event occurring during an archaeological era (archaeological strata) and, at the same time, it documents the action of the modern-day scholar. Any excavation activity and its pertinent methodology may be easily described following this conceptual formalism.

### 3.3 Mapping of Archaeological ICCD and SYSLAT Forms into CIDOC-CRM

The main peculiarity of our project was based on *a highly interdisciplinary approach* playing a fundamental role in analyzing the structures of the different recording system.

Even if the two systems were based on the same vocabulary, there is no guarantee that they can agree on certain bits of information unless they commit to the same conceptualization. As each recording system has its own conceptualization, a condition that is necessary in order to make an agreement possible is that the intended models of the original conceptualizations overlap.

Each archaeological method or strategy can simply be described as a task-ontology which, as object of its activity, produces a specific documentation. *A task-ontology* describes generic task or activity by specializing the terms introduced in the top-level ontology. In our case it was possible to illustrate the two different systems by considering the selected strategies as a specialization of the concept of the Stratigraphical Excavation.

So it is impossible to verify the documentation without knowledge of the method adopted by the archaeologist. Similarly it is impossible to guarantee the integration of different data acquired in the same area by different archaeological teams.

The primary objective of our research has been to analyse the documentation gathered during the stratigraphical excavation, which currently represents the most consistent corpus of data available also in digital format according to national regulations. An experimental project regarding the mapping of excavation forms according to the CIDOC-CRM standard has been initiated and its primary role is "*to enable information exchange and integration between heterogeneous sources of cultural heritage information*".

We've used CIDOC-CRM in order to supply the data with the semantic definition required in order to transform the archaeological documentation into a coherent global framework.

Our project is based essentially upon recognizing the correspondence of meaning and content between different classes of CIDOC-CRM and the fields defined in the forms chosen as samples for this experiment. In synthesis our objective is to propose an extension of the CIDOC-CRM classes and to pinpoint as many sub-classes and/or relationships for the structure of certain excavation forms (which are standard on ministerial forms and on those forms created by a university team).

As defined by CIDOC-CRM, extension signifies that CRM classes subsume all classes of the extension, and all properties of the extension are either subsumed by CRM properties, or are part of a path for which a CRM property is a shortcut. Thus the model may be enriched and "customized" without losing "compatibility" with CIDOC-CRM.

The model we chose does not limit itself to a simple description of the archaeological "categories", but rather to a detailed one, since it inherits all the sub-classes and relationships of CIDOC allowing a more sophisticated formal analysis of the recorded data.

This research – linked to a larger project named AMA carried out in the framework of the European Network EPOCH (www.epoch-net.org) – relies on a previous research having similar goals, among others the one concerning excavation forms from the archaeological site of Cumae [CDF02] [DN02].

The Cumae forms, converted in XML, have been edited in RDF and visualized with the Protégé, an extensible, platform-independent environment for creating and editing ontologies and knowledge bases [protege.stanford.edu].

The graph shows a plane structure without relations and with circular-type references. According to this method of structuring data, all information is placed on the same level, thus no hierarchy is visible.

In order to extract the relationships between the data, we have extended the ICCD and Syslat forms beginning from an initial analysis which has lead to the "conversion" of certain types of information into classes and of others into properties.

Analyzing the ICCD forms, it appears that several fields, such the Cumae forms, are placed according to implicit relations not included in the form structure. The most evident "anomalies" for instance concern chronology/dating and author. Dating may refer to the archaeological event, to the excavation or to the documentation. In a similar way, the author may be the person who produced the event in the past, the excavator or the form compiler. In fact the form label simplifies its understanding, but a simple conversion shows, as in the previous case of Cumae, how difficult it is to extract a semantic structure from the fields. A difficulty clearly emerges – showing a definite impact on the quality of

documentation – regarding the field "stratigraphical reliability": does it depend on the nature of the excavation (for example an investigation performed with a mechanic device provides less data than a brush-made one) or on the archaeologist's perplexity about the interpretation of the stratigraphy (basing on experience/competence)? Neither appropriate metadata system is sufficient. Why did the archaeologist chose one system (the mechanic device) instead of another (the brush)? Possibly he/she was interested in reaching quickly the lower layers and considered as irrelevant the sacrifice of a layer in the overall economy of the excavation management and interpretation. Often such information is located in the "interpretation" field, a sort of "black hole" where the archaeologist puts the most important data for the laboratory reconstruction of the excavation. In a similar way, the attached documentation (photos, drawings) is compressed by the methodology chosen.

Another aspect emerging from the semantic reading of the US form derives from the so-called stratigraphical relations, that is the spatial and temporal relations among different Stratigraphical Units (natural/anthropic actions discovered during the excavations and individually numbered). In this case they are not class attributes, but spatial relations among US forms which sometimes document events that happened before/after the one recorded in the form.

This preliminary analysis produced a first manual mapping consisting in the recognition of the correspondence between the classes of the US form and the entities of CIDOC-CRM. In this case, using Protégé we created a sub-class corresponding to the US form fields for every single entity.

The following tables show a synthetic exemplification of the draft predisposed for highlighting the correspondences between our forms and the CIDOC-CRM entities.

| **ENTITY-TAG** | **CONCEPT** | **CRM-ENTITY** |
|---|---|---|
| **SCHEDA_US** | US positive form | *E31_document* |
| **NOTES** | Each unit in the excavation has a corresponding US form. The tag refers to the database report or, alternately, to the paper sheet concerning the unit. | |
| **RELATION** | **WITH** | **NOTES** |
| P70_documents | AIE_13_us | |
| P4_has_time_span | AIE_06_anno | |
| P70_documents | E5_event | Event: excavation |

| **ENTITY** | **CONCEPT** | **CRM-ENTITY** |
|---|---|---|
| **DIRETTORE** | Director of excavation | *E82_actor_ appellation* |
| **NOTES** | Name of the director of the excavation. Tag used in forms: US, USM, USN. | |
| **RELATION** | **WITH** | **NOTES** |
| P11_participated_in | E5_event | E5_event (excavation) P108_has produced: scheda_us |
| P49_is_former_or_ current_keeper_of | scheda_us scheda_usm | |

| **PROPERTY** | **CONCEPT** | **CRM-PROPERTY** |
|---|---|---|
| **ANTERIORE_A** | Before to (US, USM, USN) | *P120_occurs_ before* |
| **NOTES** | Chronological relationship with positive and/or negative stratigraphical unit. It refers to the USM, US, USN chronologically and immediately before to the unit in exam. Tag used in US, USM, USN forms. | |

Using the Protegè editor and an RDF file present on the CIDOC-CRM website, we subsequently created sub-classes corresponding to the ICCD and SYSLAT standards for every CRM class singled out. It was thus possible to maintain the sub-classes and properties defined by CIDOC-CRM.

After checking the correspondences between our items and the entities of CIDOC-CRM, the mapping relied on the definition of possible new sub-classes and sub-properties. It often isn't sufficient to find the correspondences to highlight the semantic content without pointing out the paths or links that join each field/entity to one another. In Figure 1 we show how it was possible to link the compilation of the form to the author avoiding to link this task to the author of the excavation.

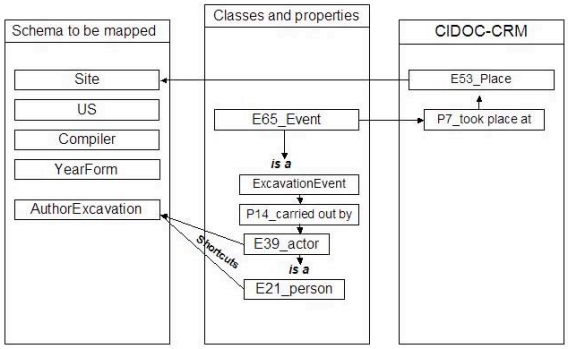


**Figure 1:** *Mapping mechanism.*

At the end of the process of analysis, we created the mapping procedure and generated the RDF file with the protege editor. Figures 2 and 3 show the hierarchy and the relationship among the classes CIDOC-CRM and those of the Italian standards.

The mapping thus determined is now available in the information management system of the excavation of Cumae, created as a re-adaptation of the old Syslat system in the open-source Exist environment [Fel06]. The data may now be researched and visualized in the proprietary format or exported in RDF in a format that complies with CIDOC-CRM.

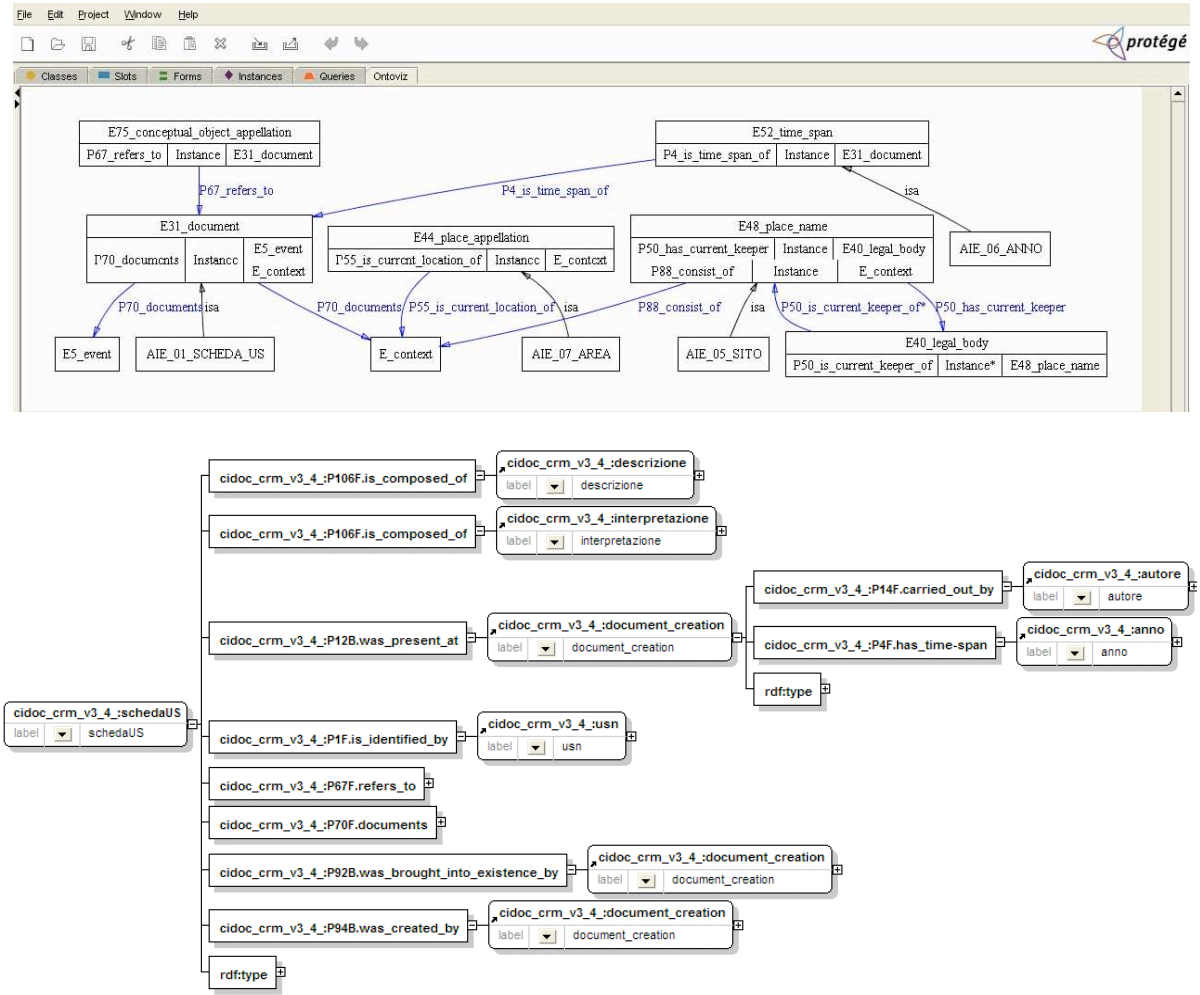


**Figure 2 and 3:** *The Graphs represent the hierarchy between the new sub-classes and CIDOC-CRM classes.*

## 4. Conclusion

The use of standards in the description of archaeological activities, both at the level of excavation and of the cataloguing of goods and handcrafted artefacts, represents a strong demand that cannot be limited, as in the case of the Italian ICCD, to ensuring a correct conservation and management of the archives, especially in the field of the management of archaeological goods. To this day a common descriptive system and a standard international vocabulary are lacking. There are difficulties turning the various attempts made into a single protocol.

The existence of national and/or local regulations and systems further complicates a picture already made complex by the habit of archaeologists to "create" new description criteria and new formalizations. As we have seen, the introduction of computer science hasn't simplified the situation. It has instead made it more articulate because of the unlimited possibilities to create and adapt local and temporary systems, which do not have the pretence of becoming object of new studies for new hypotheses.

The problem of standards, from whose solution a new, more scientific relation with the archaeological documentation - which represents the only testimony of the activity on the field - may result, does not appear to be an immediate necessity.

The benefit of using CIDOC-CRM is evident: on the one hand it forces archaeologists to conceptualize and formalize the excavation methodology and the recording of the data collected in the course of field work; on the other hand it allows the same researchers to maintain their systems guaranteeing the possibility for other archaeologists to check and re-use the data through CIDOC-CRM mapping.

Standard in our mind signifies primarily having the possibility of verifying the quality of the data in a discipline that, contrary to hard-sciences, does not allow the repetition of such a destructive experiment as the excavation. Furthermore it is possible to control the procedure and, no less important, to provide the framework for a better planning and implementation of the protocols chosen at the local level.

The disadvantages (minimal, if compared with the advantages) are to define in a conceptual and formalized way, according to a modality that is often uncongenial to the archaeologist, the methodological and theoretical itinerary that guides the researcher in his/her way in their process of cultural interpretation of the facts, starting from

the material and historically reconstructed culture.

From a methodological point of view, the work carried out so far seems to align itself with what has already been experimentally accomplished by the Centre for Archaeology of English Heritage [CGF*04] which, as declared in the foreword, is based upon the attempt to model a conceptual frame for all the archaeological data created by man. Rather than by an extension, the mapping in this case has been carried out by simplifying the CIDOC-CRM classes.

In the future our objective will be to formally "describe" other types of forms and/or document representation in order to test the potential of such an approach and to guarantee the interoperability between information pertaining to an archaeological context investigated over a long period of time.

In conclusion our scope was to understand the conceptual model used by archaeologists to record data excavation in order to propose a new form for the documentation according the semantic structure supplied by CIDOC-CRM. Only this way will it be possible to obtain real data integration and a re-use of data without altering and sacrificing the specific background and targets of each individual archaeologist.

The experimentation carried out during our project may represent an option for the creation of an ontology for the cultural heritage. Using CIDOC–CRM has definitely caused problems of adaptation and conversion of our requirements, specifically those regarding research, rather than those regarding museum conservation guaranteed principally by CIDOC-CRM. The risk would have been to formalize a new ontology for cultural heritage freed from the CIDOC-CRM. Its achievement would have certainly been simple without having to go through the "choices/decisions" container in the CIDOC-CRM not easily extendable to the archaeological area. However we would have created another new standard, further complicating a picture that today seems to be characterized by incommunicability.

Facilitating the exchange of information, on the contrary, allows an improvement of knowledge, the safeguard of local and world archaeological heritage.

**Acknowledgments**

Work on the global mapping system is presently carried on within the project AMA, by a partnership formed by CISA, CIMEC, IAA, Oxford ArchDigital, PIN, ROB, University of Kent, University of Oslo, UNIREL, and VARTEC.
AMA is developed in the framework of the EU funded Network of Excellence EPOCH (http://www.epoch-net.org/). EPOCH is funded by the European Commission under the Community's Sixth Framework Programme, contract no. 507382. However, this presentation reflects only the authors' views and neither EPOCH, nor the European Community are liable for any use that may be made of the information contained herein.

**References**

[CDF02] CRESCIOLI M., D'ANDREA A., NICCOLUCCI F.: XML Encoding of Archaeological Unstructured Data. G. BURENHAULT (Ed.), *Archaeological Informatics: Pushing the envelope*. In *Proc.* of the 29th CAA Conference, Gotland April 2001, BAR International Series 1016, Oxford, pp. 267-275.

[CGF*04] CRIPPS P., GREENHALGH A., FELLOWS A., MAY K., ROBINSON D.: *Ontological Modelling of the work of the Centre for Archaeology*, September http://cidoc.ics.forth/docs/Ontological_Modelling_Project_Report_%20Sep2004 .pdf)

[CDG*05] CROFTS N., DOERR M., GILL T., STEAD S., STIFF M.: *CIDOC CRM Special Interest Group*, Version 4.2, June 2005 (http://cidoc.ics.forth.gr/docs/cidoc_crm_version _4.2. pdf)

[DN02] D'ANDREA A., NICCOLUCCI F.: Database di scavo ed Internet. L'accesso remoto all'archivio Syslat. B. D'AGOSTINO, A. D'ANDREA (Eds.), *CUMA. Nuove Forme di Intervento per lo Studio del Sito Antico.* In *Atti della Giornata di Studio, Napoli 12 Febbraio 2001*, AION ArchStAnt, Quad. 14, Napoli 2002, pp. 167-175.

[Fel06] FELICETTI A.: MAD: Managing Archaeological Data. Paper presented at *7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2006), Cyprus.*

[Gua98] GUARINO N.: Formal Ontology and Information Systems Formal Ontology. In *Information Systems, Proceedings of FOIS'98*, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, pp. 3-15.

[Hod99] HODDER I.: *The Archaeological Process*. Blackwell, Oxford.

[Hod00] HODDER I. (Ed.): *Towards reflexive method in archaeology. The example at Çatalhöyük,* MacDonald Institute for Archaeological Research, Cambridge.

[Mob81] MOBERG C.A.: *Introduzione all'Archeologia*. Milano.

[Nor04] Normativa 300 (2004) http://www.iccd.beniculturali.it/standard/normative_300.html

[RB96] RENFREW C., BAHN, P.: *Archaeology: Theories, Methods, and Practice*. 2nd Edition, Thames and Hudson, London.

All the references have been visited on 10-07-2006