Preprint version.

Qualitative Evaluation of Cultural Heritage Information Modelling Techniques

CHARLOTTE HUG

Centre de Recherche en Informatique,
Université Paris 1 Panthéon-Sorbonne, Paris, France
Laboratoire d'Informatique de Grenoble,
Grenoble University, Grenoble, France

AND

CESAR GONZALEZ-PEREZ

Institute of Heritage Sciences (Incipit),
Spanish National Research Council (CSIC),
Santiago de Compostela, Spain

_____

In this paper we evaluate three modelling techniques coming from information system engineering to represent cultural heritage domain concepts. Evaluations of the modelling techniques were performed by carrying out experiments with cultural heritage specialists. The first experiment consisted in evaluating UML class diagrams and a conceptual graph. The second experiment evaluated a value cluster modelling approach. These qualitative evaluations were conducted using focus groups composed of cultural heritage specialists. We evaluated the perceived ease of use, the perceived usefulness and the intention to use of the three techniques (UML class diagrams, conceptual graphs and the value cluster approach). Thanks to these evaluations, we will be able to propose a combined approach using value cluster modelling and UML class diagramming techniques to describe cultural heritage information.

_____

Authors' addresses: Charlotte Hug, Université Paris 1 Panthéon-Sorbonne, Centre de Recherche en Informatique, 90, rue de Tolbiac 75013 Paris France, Charlotte.Hug@univ-paris1.fr; César Gonzalez-Pérez, Institute of Heritage Sciences (Incipit), Spanish National Research Council (CSIC), San Roque, 2, 15704 Santiago de Compostela, Spain, cesar.gonzalez-perez@incipit.csic.es.

1. INTRODUCTION

In the context of computerising information, automating activities, and, in brief, easing users' jobs through the use of new technologies at the Institute of Heritage Sciences (Incipit) at the Spanish National Research Council (CSIC), we were faced with the problem of representing the structure of a future information system aimed to replace the now obsolete SIA+ (González-Pérez [1997]).

The new information system, currently in requirements analysis phase, is expected to manage all the research and technical information generated and used by Incipit in the fields of archaeology, anthropology, architecture and related areas. Specifically, we tried to tackle the problem of how researchers in the above mentioned cultural heritage fields could model their domain concepts in an accurate and easy way. Indeed, involving cultural heritage specialists at this stage of the project should ensure the relevance and the adequacy of the defined domain concepts to the users view.

Requirements analysis of such a domain information system is usually carried out by experienced engineers, usually trained in modelling languages such as UML (ISO [2005], OMG [2010]) and usually a software development method. However, in the context of the future information system of Incipit, no resources and skills were available for a trained engineer to perform this task. Moreover, the diversity and the complexity of the cultural heritage domain impose the participation of different domain specialists. It would be very costly for an engineer to understand all the concepts and have a global vision of such a large and complex domain. The communications between engineers and domain specialists would be multiplied and therefore complicated. It was proposed that cultural heritage specialists (both academics and technical staff) were put in charge of creating the conceptual models of the information that they would need to enter, use, and retrieve from/to the future information system.

Information systems engineers, as well as information systems researchers, are trained to model and use models everyday in their job, which give them a subjective vision of the relative difficulty (or ease) of modelling. As we decided that the modelling task would be carried out by cultural heritage specialists with no (or little) previous training in information technologies, usability of the modelling approach was a crucial concern. So, one of the first things we needed to do was to look for an easy and pertinent way of modelling:

—Easiness: Cultural heritage specialists are not information systems engineers, and are not willing or able to spend hours learning a modelling language. The modelling language that they were to employ had then to be extremely easy to understand and use.

—Pertinence: The modelling language to be employed must be able to serve the specific needs of the cultural heritage fields (archaeology, anthropology, architecture, etc.); these needs are strongly grounded in the social sciences and humanities and, arguably, pose constraints and specificities that differ from "classic" information systems engineering, which tends to be oriented towards "hard science" subject fields; cf. *Principle of Scientific Simplicity* in (Gonzalez-Pérez [2002]), p. 185.

In this situation, we decided to roll out two experiments in order to collect information about conceptual modelling carried out by cultural heritage specialists. The first experiment aimed to qualitatively evaluate the perception of, and intention to use (Moody [2003]) the class diagramming and the conceptual graph (Hug [2009]) techniques when used to build cultural heritage information models. The second experiment aimed to qualitatively evaluate the perception and intention to use of what we call the value cluster ("valster" for short) modelling approach. These techniques were chosen according to their relevancy with the context of cultural heritage modelling. These

experiments allowed us to select one of the evaluated techniques to model cultural heritage information in the framework of the Incipit information system design.

As the main objective of the project was to represent the structure of the future information system at Incipit, we only considered models that allowed representing domain concepts, not processes. ER modelling (Chen [1976]), could have been considered but its expressiveness is lower than UML class diagrams. Moreover, we did not try, in these experiments, to relate the evaluated conceptual modelling techniques to archaeological theories such as the attribute-artefact-type (Clarke [1978]) and Gardin's theories (Gardin [1958]).

This paper presents the qualitative evaluations and their results. Section 2 presents related work, including a summary of the state of the art on modelling by non computer specialists and qualitative evaluations. Section 3 presents the first experiment, related to the evaluation of class diagramming and conceptual graph techniques, and Section 4 introduces the experiment related to the evaluation of the value cluster ("valster") modelling approach. Finally, Section 5 concludes this paper.

2. RELATED WORK

2.1. Modelling by cultural heritage specialists

Cultural heritage is often considered to be part of the humanities and, as such, is a low-tech field as compared to "hard" scientific areas such as medicine or construction. However, cultural heritage is quite high-tech *within* the humanities; the development of techniques in the fields of geographic information systems, 3D documentation (photogrammetry, laser scanning, etc.) or virtual reconstructions is tightly connected to the needs stated by, and funding obtained from, cultural heritage users and organisations. Still, cultural heritage belongs in the humanities, and this manifests through the additional challenges that are rarely present in "hard" scientific areas; the need to explicitly manage subjectivity in the information being collected and processes is a good example, as is the fact that the informational categories used by cultural heritage specialists are sometimes of a radial nature (Lakoff [1990]) rather than criteria-based. This means that information modelling is especially difficult in cultural heritage, and the modelling languages, tools and guidelines that are often valid in other, more mainstream fields with a harder scientific orientation, are only partially applicable here.

Being part of the humanities, the theoretical and methodological underpinnings of cultural heritage are far removed from those of information technologies, much further than those of experimental scientific areas or, of course, engineering fields. This is visible, for example, through the fact that the hypothetico-deductive method is rarely used within cultural heritage studies, or by acknowledging that the humanities in general, and cultural heritage in particular, are approached from theoretical positions that are often of a strongly narrative, discursive and interpretive nature, whereas engineering and experimental fields are approach from positions that tend to rely more on analysis and prediction. As a consequence, information modelling efforts in cultural heritage that are carried out by end users rather than specialised experts are less likely to enjoy the necessary formal support; and scarcity of resources and lack of conceptual infrastructures mean that these efforts tend to be fragmented and highly local in scope. At Incipit we have decided to attempt to minimise these barriers and explore the possibility of having end users construct their own information models, taking this task away from the engineer or consultant. We are aware that, for this to be workable, the necessary resources must be made available; these resources include an adequate modelling theory,

a modelling language that is capable of rich expressivity, and a toolset that enables users to employ that language with ease.

This paper focuses on modelling theory and, mainly, on modelling languages, by empirically assessing how end users perceive and employ three different modelling languages underpinned by two major alternative theoretical approaches. On the one hand, class diagrams and conceptual graphs are supported by a modelling theory based on the type/instance paradigm; on the other hand, value cluster modelling is supported by a theory of typeless modelling.

2.2. Qualitative evaluations

Our aim was to collect the maximum of ideas and comments about the different modelling techniques. We wanted to be able to select one of the evaluated techniques according to the feedback of the cultural heritage specialists at Incipit. We chose to carry out qualitative evaluations based on the Method Evaluation Model (MEM) (Moody [2003]) and the focus group technique that we present below.

2.2.1. The Method Evaluation Model

In order to evaluate the different modelling techniques to represent cultural heritage information, we used part of the Method Evaluation Model (Moody [2003]). This model takes into account the Technology Acceptance Model (TAM) (Davis [1989]) that is based on two characteristics:

—The perceived usefulness, defined as "the degree to which a person believes that using a particular system would enhance his or her job performance",
—The perceived ease of use, defined as "the degree to which a person believes that using a particular system would be free of effort."

TAM is also pertinent for evaluating methods. Moody (Moody [2003]) integrated TAM with Methodological Pragmatism (Rescher [1977]) which states that an objective validation of a method should be measured by its efficiency and effectiveness with which it achieves its objectives. The Method Evaluation Model includes then (see Figure 1):

—The actual efficiency, which is the effort required to apply a method,
—The actual effectiveness, which represents the degree to which a method achieves its objectives,
—The perceived ease of use,
—The perceived usefulness,
—The intention to use, which is the extent to which a person intends to use a method,
—The actual usage, which is the extent to which a method is used in practice.
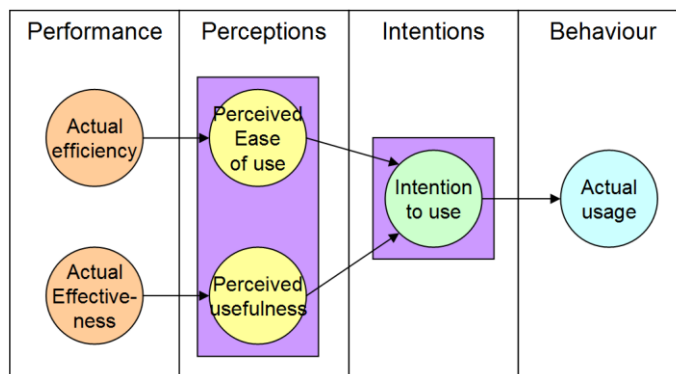
Fig. 1. The Method Evaluation Model.

The performance and the behaviour of a method or a technique can be evaluated through quantitative evaluations. However, to run quantitative evaluations, a large number of subjects are needed in order to establish valid results in terms of statistics. We did not have the sufficient number of subjects at our disposal at Incipit, so we focused our evaluations on the qualitative side only, specifically, on the evaluation of the perceived ease of use, the perceived usefulness and the intention to use, as framed in Figure 1.

2.2.2. The focus group experiment
To carry out the qualitative evaluations, we decided to organize the experiments using the focus group technique (Nachmais, et al. [2008]). A focus group is "a group of people who have been brought together to discuss a particular subject in order to solve a problem or suggest ideas." (Cambridge [2010]) This technique allows bringing together several subjects at the same time in order to carry out the same experiment. It allows time saving, it facilitates the emergence of new ideas through the debate and it "allows researchers to capture subjective comments and evaluate them" (Edmunds [1991]). The focus group is by definition a technique to carry out qualitative evaluations and ideally met our needs to collect cultural heritage specialists 'opinions about the modelling techniques.

3. EVALUATION OF THE CLASS DIAGRAMMING AND CONCEPTUAL GRAPH TECHNIQUES
In this first experiment, we wanted to evaluate the perceived ease of use, the perceived usefulness and the intention to use of two different modelling approaches: UML class diagramming (OMG [2010]) and the conceptual graph (Hug [2009], Hug et al. [2010]).

We first describe the two modelling techniques; we then present the followed protocol and the results of the evaluation.

3.1. The modelling techniques

3.1.1. UML class diagramming
UML class diagrams (OMG [2010]) are well-known and widely used in the software industry and academic world. They allow representing static information using classes, attributes and associations between classes. UML class diagrams have proven to be an efficient means to represent information in a formal way and easy to implement in object-oriented code.

We built a class diagram to represent archaeological information from different sources (González Pérez [1999], [2002]). Figure 2 presents an extract of the UML class diagram describing the primary entities that are potentially useful for archaeological projects within Incipit.

For example, an Archaeological Site (*Site* class) has a type, a chronology and a cultural ascription, by inheritance of the attributes of the *Archaeological Entity* class. An Archaeological Site can be located in any other *Primary Entity*, as a *Geographic Entity* or in another Archaeological Entity, by inheritance of the *IsLocatedOn* association defined on the *Primary Entity* class.
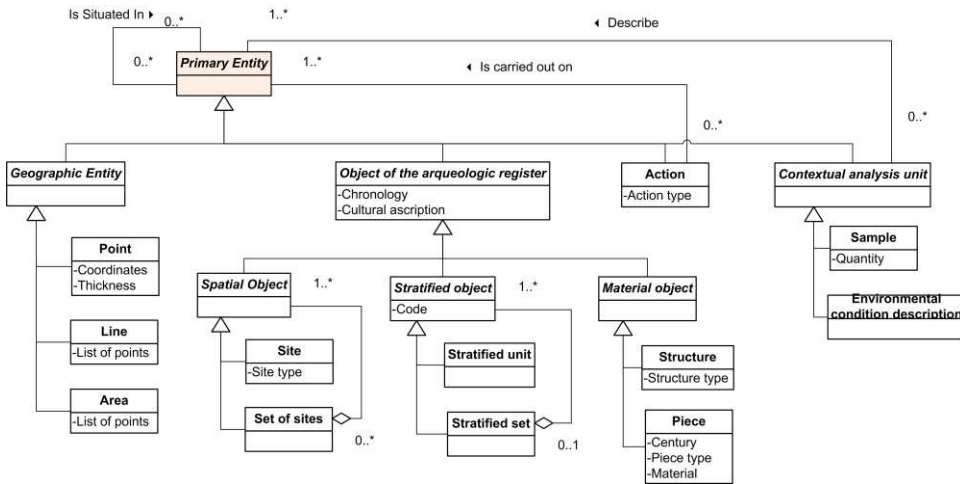
Fig. 2. Extract of the UML class diagram describing the primary entities

3.1.2. Conceptual graph

The conceptual graph technique proposed in (Hug [2009], Hug et al. [2010]) was designed for information systems engineering processes. It allows people with no training or specific knowledge in process metamodelling for information systems engineering to build process models by selecting concepts in a graph. Each concept in the conceptual graph corresponds to a class. The relations defined between the concepts in the graph help the selection of the concepts according to what the user wants to do with the model: completing it, specialising it or adding properties. The selection of a concept generates the creation of a UML class diagram that is not shown to the users. The conceptual graph allows building UML class diagrams by hiding the difficulties of modelling from the users since it only shows the concepts and abstracted relations. (Hug et al. [2010b]) conducted an evaluation of the conceptual graph in the context of information systems engineering processes. Results showed that the conceptual graph technique was useful and helpful to build process metamodels. In this experiment, we used the conceptual graph to build archaeological information models.

We designed a conceptual graph for archaeological information using the class diagram previously created as a source (Figure 2); an extract of this conceptual graph can be seen in Figure 3. The classes and attributes in the class diagram are represented as concepts in the graph; the associations in the class diagram were abstracted by three types of relations in the graph:

—"Is specialized by" to represent specialization between classes,

—"Is completed by" to represent associations between classes,

—"Is characterized by" to represent attributes of classes.

Starting from the *Primary Entity* concept, users can complete it (by the *Context Entity* or the *Intervention* concepts) or specialize it (*Archaeological Entity* or *Geographic Entity* concepts).
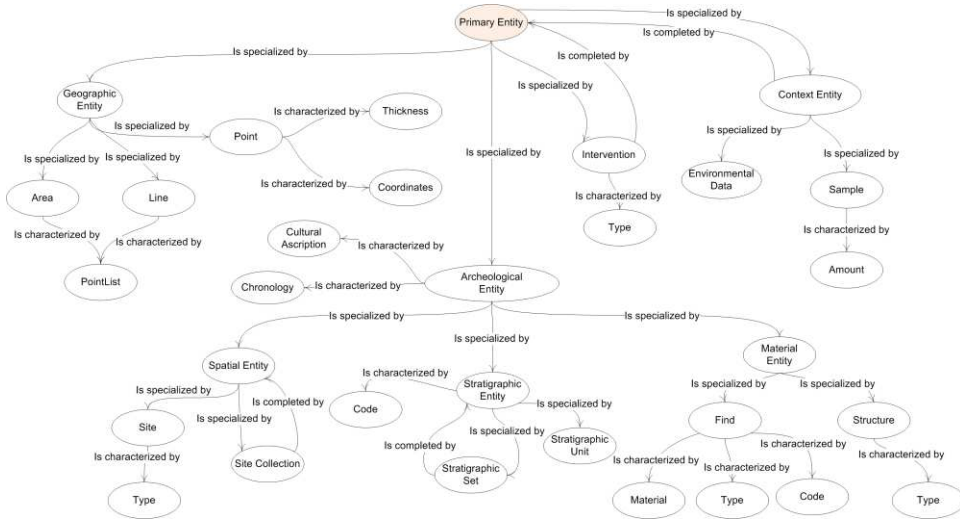
Fig. 3. Extract of the conceptual graph describing the primary entities

3.2. The protocol

3.2.1. The objectives

The objectives of this experiment were to evaluate the perceived ease of use, the perceived usefulness and the intention to use of the class diagramming and the conceptual graph modelling techniques, in order to define which of them was more adequate to model archaeological information. As mentioned earlier, the evaluation of the performance and behaviour of a method entail quantitative measurements and therefore they cannot be conducted in our context, as the number of individuals is too low.

3.2.2. The subjects

As we needed specialists in cultural heritage, we contacted the staff of Incipit, which is composed of research assistants, PhD students, post-doctoral researchers, staff scientists and one professor. The focus group comprised fifteen subjects who were split into two different groups. One group used the UML class diagramming technique (eight subjects) and the other group worked with the conceptual graph technique (seven subjects). The groups were homogeneous in terms of gender, age and positions held, as described in Table I.

In the UML class diagramming group, five subjects out of eight had already modelled archaeological information: one subject uses natural language, two use databases, and three subjects use specific archaeological languages. In the same group, three subjects out of eight had never modelled archaeological information, that is to say, they were only users of existing models or databases.

In the conceptual graph group, six subjects out of seven had already modelled archaeological information: five subjects use natural language, two use data bases, and three subjects use specific archaeological languages. Only one subject out of seven in this group had never created models before the experiment.

Table I. Profile of the subjects

Technique	UML class diagramming	Conceptual graph
Gender	3 women, 5 men	3 women, 4 men
Average age	35	34
Position	5 research assistants 2 PhD students 1 professor	4 research assistants 1 PhD student 2 post-doctoral researchers
Previous experience	Natural language: 1 Databases: 2 Specific languages: 3 Never modelled: 3	Natural language: 5 Databases: 2 Specific languages: 3 Never modelled: 1

3.2.3. Process execution

The two groups were convened at different times. We used PowerPoint slides as a guide to explain what the subjects had to do at each moment of the experiment.

The subjects first had to fill out a questionnaire about their experience in modelling, the answers to which are described in the previous section. Then we presented them with the modelling technique they would use; the specific technique (class diagramming or conceptual graph) was briefly described and exemplified. The subjects were then given a case study describing a fictional archaeological project. The objective of the exercise was to design individually, a UML class diagram or a conceptual graph that represented the information described in the case study. To do that, the subjects had to select the needed classes/concepts from the class diagram/conceptual graph, using a printed version of the corresponding model and a highlighter pen. They could also add new classes/concepts and relations using a regular pen. If parts of the model were useless or wrong, they could cross them out. Figure 4 shows the solution produced by one of the subject of the conceptual graph group: the selected concepts are highlighted in blue, three concepts were added ("Area de actividad", "Numero de UE", "Contexto"), two concepts ("Tipo de actuación" and "Actuación") were transferred to another area of the diagram and the concept "Grosor" was renamed "Tamaño". A translation of Figure 4 into English is available in the appendix.
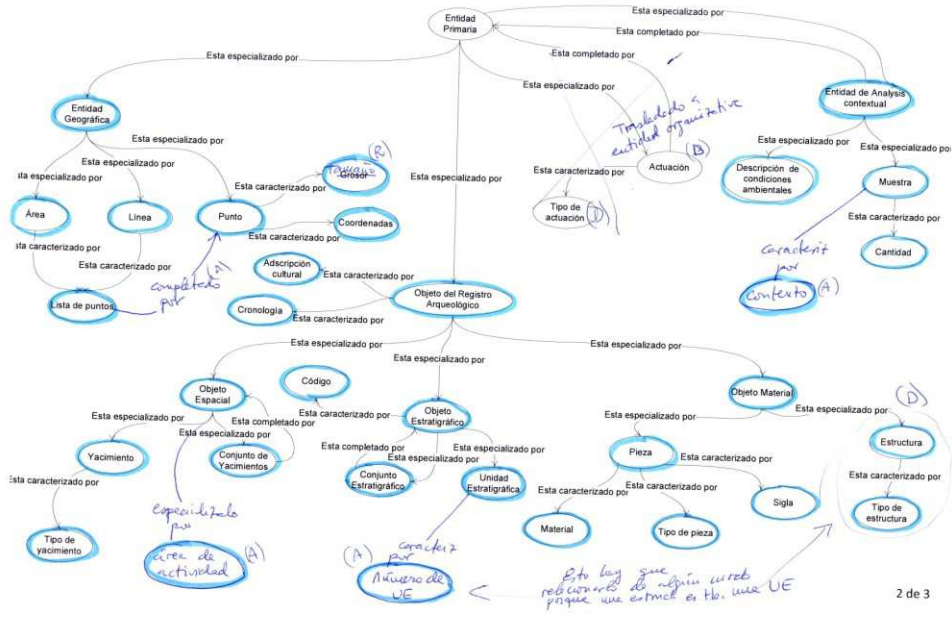
Fig. 4. Extract of the solution proposed by one subject

Once the exercise was over, the subjects had to fill out a questionnaire to evaluate the perceived ease of use, the perceived usefulness and the intention to use of the modelling technique they had worked with. The possible answers were "strongly disagree", "disagree", "neither agree nor disagree", "agree" and "strongly agree", according to the Likert scale (Likert [1932]). The questions are presented in the next sections along with the results.

In order to stimulate the emergence of ideas, a discussion between the subjects was organized. During ten minutes, the subjects were encouraged to discuss their opinions about the modelling technique aloud, including its strong and weak points, what was easy or hard to do. Finally, the subjects had to fill out a summary questionnaire, which questions are presented in section 3.3.4.

We planned the experiment to take two hours, including the presentation of the objectives to the subjects, the description of the exercise and the filling out of the different questionnaires. The whole experiment has been recorded and the conversations have been transcribed in order to extract the relevant comments.

3.3 The result

3.3.1. Perceived ease of use

In order to measure the perceived ease of use, we used the questionnaire proposed by (Moody [2003]) and adapted it to archaeology models. Table II presents the questions asked to the subjects, independently of their group. We next present the results of the perceived ease of use for each group.

Table II. Items about the perceived ease of use

I found the procedure for applying the technique complex and difficult to follow
Overall, I found the technique difficult to use
I found the technique easy to learn
I found it difficult to apply the technique to model archaeology projects
I found the rules of the technique clear and easy to understand
I am not confident that I am now competent to apply this technique in practice
How much time would you need to learn and use this technique?

- UML class diagramming group

The perceived ease of use of the UML class diagramming was quite negative. First, it was difficult for the subjects to follow the process of the method. The exercise was difficult and complex, and there was not enough time to do it, according to the subjects. The case study provided was too large and all-encompassing to be modelled in the allotted time. "It's a very long and complex exercise to do in the allotted time" (Subject number MB3). Another problem came from the lack of knowledge: the subjects were more knowledgeable in some part of the case study than others. They spent more time where they were more at ease, and added more details to those areas in the model. "There were parts in which I felt more comfortable and have responded with more knowledge and others which clearly cost me more." (Subject number MB3).

As a result, the subjects found difficult applying the technique to model archaeology information. A training period would have been useful before beginning the experiment to better use the technique. A learning phase would be necessary to properly use the technique and assimilate it. "To apply it [the method] properly, it requires a level of ability, accurate knowledge and a precise understanding of the method" (Subject number MB4). This phase would last from a few hours (four subjects out of eight), one day (one subject), to a few weeks (one subject).

- Conceptual graph group

The perceived ease of use of the conceptual graph was negative to moderate. Only two subjects out of seven found that following the process of the method was difficult, the others gave a neutral answer. The relations of the conceptual graph were difficult to understand and the tree representation of the graph was confusing. "I agree that the tree structure is useful for its visualization, but it also distorts a little or makes difficult the way we consider the relationships"(Subject number CG4). As in the class diagramming group, the subjects of this group thought that they did not have enough time to complete the exercise. A learning phase would also be necessary: a few hours (two subjects), a few days (two subjects) to a few weeks (two subjects) (one subject did not answer). "The concepts are not very complex but require time to be assimilated" (Subject number CG6).

3.3.2. Perceived usefulness

The second part of the questionnaire concerned the perceived usefulness of the two modelling techniques. Table III presents the questions concerning the perceived usefulness adapted from (Moody [2003]).

Table III. Items about the perceived usefulness

I believe that this technique would reduce the effort required to build model on archaeological information
Archaeological models represented using this technique would be more difficult for users to understand
This technique would make it easier for users to verify whether archaeological models

are correct
Overall, I found the technique to be useful
Using this technique would make it more difficult to maintain archaeological models
Overall, I think this technique provides an effective solution to the problem of representing archaeological models
Using this technique would make it easier to communicate archaeological models to archaeologist

- UML class diagramming group

Although the perceived ease of use was rather negative (see previous section), the perceived usefulness was quite positive. Seven subjects out of eight thought the technique would reduce the effort required to build models of archaeological information, but the produced models would be more difficult for users to understand (four subjects). They also thought that the technique was useful (seven subjects) because it allowed identifying objects and relations between them. "The method has many strengths as it allows the identification of objects and relationships between objects" (Subject number MB4).All the subjects thought that the technique provided an effective solution to the problem of representing models for archaeology. Four subjects thought that the produced models would help archaeologists to communicate among each other.

- Conceptual graph group

The perceived usefulness of the conceptual graph was moderate to positive. Three subjects thought that the technique would reduce the effort required to build archaeological information models (three subjects selected the median answer, one did not agree). The produced models would be more difficult for users to understand (three subjects, four median answers).

Overall, five subjects thought the technique was useful to analyse and self-criticize the information collection system. "[The method] helps you to think about the concepts, to consider the concepts and relations that we use intuitively, and also to study the systematisation of the models that we use" (Subject number CG4).

3.3.3. Intention to use

The intention to use was evaluated through two questions presented in Table IV, adapted from (Moody [2003]).

Table IV. Items about the intention to use

I would definitely not use this technique to create archaeological models
After this exercise, I will use this technique to create archaeological models

- UML class diagramming group

Whereas the perceived ease of use was negative (see previous section), the intention to use was moderate. Four subjects reported that they would use the technique to create models for archaeology projects (four subjects selected the neutral answer). Some of them could not answer more precisely because of lack of knowledge about the technique.

- Conceptual graph group

The intention to use was moderate. Five subjects did not know if they would use the technique to create models for archaeology.

3.3.4. Synthesis

Table V presents the synthesis questions that were asked to the subjects in the last questionnaire, after the discussion.

Table V. Items about the synthesis

Overall, what did you think of this exercise?
What was the easiest thing to do?
What was the hardest thing to do?
Which are the strong points of the technique?
Which are the weak points of the technique?
Which are the improvements that could be done?

- UML class diagramming group

The exercise on modelling classes, attributes and associations was new for the majority of the subjects. Some of them were used to employ terms such as "model", "class" and "relationship", but not to organize and structure them as it was presented in the class diagram. The easiest thing to do during the exercise was to manipulate the well known concepts. The hardest thing to do was to use the language they did not know, in particular understanding and using the associations between the classes (five subjects out of eight).

The strong point of UML class diagramming is its capacity to represent, organize, and synthesize the information used during archaeological projects (five subjects out of eight). Three subjects underlined that once the technique is mastered, it could be very useful.

- Conceptual graph group

The subjects gave quite a lot of attention to the three pre-defined relations between concepts in the conceptual graph. As they were very abstract, it was not convenient for the subjects to use them. Although the relations were defined in order to help the users of the conceptual graph, they were not specific enough to represent the needed information.

The "tree problem" appeared clearly: the tree layout of the graph made the visualization and the understanding of the relations between the concepts a difficult task. The subjects then added new relations that were already defined higher in the tree.

3.4 Synthesis of the evaluation

To conclude, none of the two evaluated techniques emerge as clearly superior in terms of perceived ease of use, perceived usefulness and intention to use. Nevertheless, this evaluation allowed us to set out some interesting bases for cultural heritage information modelling.

Most of the subjects understood the interest and the need to model cultural heritage information, and they found the exercise interesting. Modelling was positively perceived. This acceptance was very important as cultural heritage specialists will support the modelling effort and participate in a constructive and efficient way to our project. As any information systems engineering project, user's acceptance is one of the basis for the success of the project.

In order to facilitate the models visualisation, we will have to introduce different levels of granularity. Each model has to represent all the necessary information in an intelligible and user friendly way. Any cultural heritage specialist has to be offered a language that allows him/her to model all the required information; the modelling language cannot restrict the amount of detail to represent.

The associations as defined in UML class diagrams allow a semantic richness essential in order to define relations between archaeological concepts. On the contrary, the conceptual graph intends to hide the difficulty and the details to the users; it does not seem the right solution in the cultural heritage context or at least at Incipit as the subjects looked for precision, particularly in the definition of the relations between the concepts.

It is worth noticing that cultural heritage specialists, or archaeologists in our context, are very cautious and moderate people who do not cast clear-cut opinions about a subject they are not at ease with. In the questionnaire we proposed the answer "neither agree nor disagree", which we found to be an error, as a lot of subjects selected this answer as a precaution. This kind of neutral option is often questioned as it allows the subject to give a non significant reply.

Finally, the model we used during the exercise raised questions about the organization of the concepts as a tree. The adequate structure of the models, the concepts and the relations has to be found.

4. EVALUATION OF THE VALUE CLUSTER MODELLING APPROACH

4.1. The value cluster modelling approach

The practical totality of the conceptual modelling approaches that are available to information systems engineers have (at least) one property in common: they assume that information is to be *categorised*. That is, they focus on the *categories* of entities of interest, rather than the entities themselves. Conventional object-oriented modelling uses the "class" construct to achieve this; entity-relationship approaches use the "entity" construct; relational database approaches describe information in terms of "tables"; etc. Although some modelling languages are perfectly capable of representing individual entities (an individual object, for example), these are almost always defined and described in terms of the category they belong to.

For the sake of clarity, we must say here that by *entity* we refer to the "things" that instance-level model constructs directly refer to. In an object-oriented model, entities are those things represented by objects; in an entity-relationship model, entities are those things represented by entities (rather than entity types, as intended by (Chen [1976])). We use the term "entity" with this meaning throughout this article.

The approach to modelling where the focus is placed on *categories* of entities, which we have called *metatypical* elsewhere (Gonzalez-Pérez et al. [2007]), is not the only possible one, but is certainly useful to organize and systematize large amounts of information, since categories often make excellent surrogates for the corresponding entities. However, metatypical modelling, by relying on the assignment of a category to any entity that is to be considered by the information system, assumes that a stable, well-known categorisation scheme exists. The humanities in general and cultural heritage in particular, are fields where this assumption is often challenged.

The existence of a pre-defined class hierarchy, domain type system or database table structure means that any information entity that is to be considered must conform to said hierarchy, system or structure. In a research setting, where knowledge is often being discovered or constructed on the fly, freezing a category scheme too soon is anything but advisable. Sophisticated information systems can implement models that allow a high degree of flexibility, allowing researchers to alter the categories greatly, incorporating new ones or modifying existing ones to accommodate new entities as necessary. Still, every entity in the system must conform to the categorisation scheme.

Arguably, imposing any particular categorisation scheme on the observed reality injects a non-trivial dose of interpretive subjectivity into data collection process,

transforming "raw" data into tainted information that is, perhaps inadvertently, interpreted according to cultural, ideological and intentionality biases of whoever created the categories being applied. This effect was delightfully called "the Walcott's shoehorn" by Stephen Jay Gould in (Gould [1990]). If the shoehorning effect is deleterious in the "hard" sciences, it is even less desirable in the social sciences and humanities, where subjective interpretation plays such a considerable role in the research process. Our experience shows that researchers in cultural heritage, and in particular in the fields of archaeology and anthropology, have been reluctant to use pre-defined categorisation schemes to organise their data.

We are aware that different kinds of subjectivity may be taken into account here; Searle's ontic vs. epistemic subjectivity levels (Searle [1990], chapter 5) should at least be taken into account when trying to characterize what kinds of subjectivity is "allowed" to be injected into a model. A treatment of this topic, however, is beyond the scope of this paper.

Taking these ideas into account, the value cluster modelling approach builds upon the radical idea that no categorisation whatsoever is allowed during modelling. Entities of relevance may be represented in a model as value clusters (or "valsters", for short), i.e. collections of values where a value is a named and typed piece of data. For example, "Height (number) = 12" and "IsAvailable (Boolean) = false" are values. Since no categories can be established, only isotypical or prototypical modelling is possible (see Gonzalez-Perez et al. [2007]); in other words, entities can be represented in models, but no types, classes or categories of any sort.

The major advantage of describing cultural heritage information using valsters rather than a conventional, category-oriented approach, is that one gets rid of the implicit bias that a pre-existing categorisation scheme imposes on the data. There are some foreseeable disadvantages, such as the difficulty of resisting the natural impulse to categorise data as it arrives. Our objective was to verify whether advantages surpassed disadvantages in the usage of a valster modelling approach to describe archaeological information at the conceptual level. An in-depth description of the valster approach can be found in Gonzalez-Perez [2012].

4.2. The protocol

4.2.1. The objectives

The general objective of this evaluation consisted in evaluating if valster models were adequate to represent cultural heritage information. We especially wanted to evaluate the perceived ease of use of the valster modelling technique, its perceived usefulness and its intention to use (Moody [2003]). As the first evaluation, we focused on the qualitative evaluation of the valster modelling technique.

4.2.2. The subjects

We also used the focus group method to evaluate the valster modelling technique. The focus group was composed of ten persons from Incipit. Their profile is described in table VI. Because of lack of volunteers, some subjects were the same as the previous experiment.

Eight subjects out of ten needed representing information in a more formal way than in natural language. Two subjects did not need that at all.

Among the subjects, five never modelled before, two subjects had used natural language, two subjects had used databases and one had used specific tool for Geographical Information Systems (GIS).

The subjects had to work in pairs: each pair was composed by a beginner in modelling plus a more experienced subject.

Table VI. Profile of the subjects

Gender	4 women, 6 men
Average age	34,5
Position	3 research assistants 4 PhD students 1 post-doctoral researcher 1 staff scientist 1 professor
Previous experience	Natural language: 2 Databases: 2 Specific languages: 1 Never modelled: 5

4.2.3. Process execution

The experiment was conducted at two different times because of the difficulty to bring all the subjects together. We used a PowerPoint presentation to lead the experiment and explain the exercise.

The subjects first had to fill out a questionnaire about their experience in modelling which results are summarized in the previous section. We then presented the exercise. The subjects had to describe the replica of a pot from the Iron Age found in Cuntis, Pontevedra in Spain by Incipit archaeologists, see Figure 5. The subjects also had a short technical description of the pot in natural language stating the author of the replica, the cultural ascription of the original pot, the area of origin and a brief description of the pot.

Fig. 5. Replica of a Galician pot from the Iron Age

Using these resources, the subjects had to build a valster model, that is to say, model the pot without using any classifying attributes or assigning predefined types such as "artefact" or "pot" for example. In order to help them with describing the pot, we also gave them a list of properties that they might find useful. The subjects had one hour to carry out the exercise working in pairs. Figure 6 shows the valster model produced by one pair of subjects. A translation of Figure 6 into English is available in the appendix.
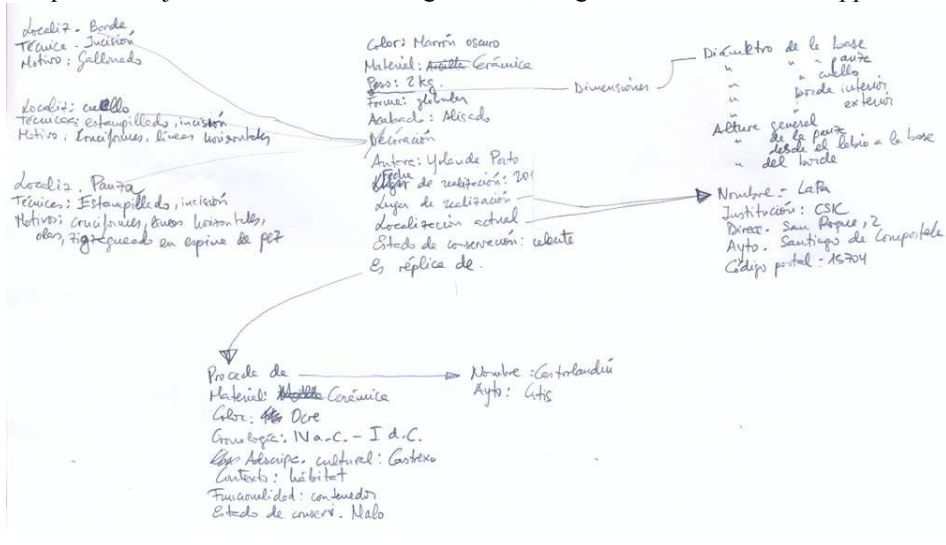

Fig. 6. Valster model produced by a pair of subjects

When the exercise was over, the subjects had to fill out a questionnaire to evaluate the perceived ease of use, the perceived usefulness and the intention to use of the valster modelling technique. This time, we proposed the following scale to answer the questions: "strongly disagree", "disagree", "agree" and "strongly agree", in order to void the neutral answer. Then, each pair of subjects quickly presented the valster model they produced and we presented them our version of the solution. We also organized a ten-minute discussion in order to let the subjects exchange their opinions and ideas about the exercise. Finally, the subjects had to fill out a summary questionnaire, which questions are presented in section 4.4.

4.3. The results

4.3.1. Perceived ease of use

Table VII presents the questions asked to the subjects about the perceived ease of use. They were extracted from (Moody [2003]) and (Hug [2010]) and adapted to cultural heritage information.

Table VII. Items about the perceived ease of use

| I found the procedure for applying the technique complex and difficult to follow |
| I found the technique easy to learn |
| I found it difficult to apply the technique to model archaeology projects |
| I found the rules of the technique clear and easy to understand |
| I am not confident that I am now competent to apply this technique in practice |

This technique need a learning phase
How much time would you need to learn and use this technique?
I would be willing to describe this technique to another person
I found the technique pleasant to use
I would have need advice during the use of the technique
Why?

Six subjects out of ten did not find the valster technique complex and difficult to follow; they also found it easy to learn. Those who found it difficult were the subjects that never modelled before, that it is to say, this exercise was their first modelling experience. Eight subjects out of ten did not find the valster technique difficult to use to represent cultural heritage information. Globally, the subjects were satisfied with the valster technique, taking into account that they did not know it before and for some of them it was their first modelling experience.

A learning phase was considered necessary but it would be quite short; a few minutes (one subject), a few hours (four subjects), a few days (two subjects) and a few weeks (three subjects of which two are beginners in modelling). Nine subjects felt able to describe the valster technique to someone else. These scores show that the valster technique was quite easy to learn, even though the subjects only used it during one hour.

The valster modelling technique was pleasant to use for half of the subjects (the others were beginners). However, nine subjects would have needed help during the exercise, above all about the subjectivity and the properties "because there were concepts and processes that were a bit confusing" (subject number 22).

4.3.2. Perceived usefulness
Table VIII presents the questions about the perceived usefulness.

Table VIII. Items about the perceived usefulness

I believe that this technique would reduce the effort required to build model on archaeological information
Why?
Archaeological models represented using this technique would be more difficult for users to understand
This technique would make it easier for users to verify whether archaeological models are correct
Why?
Overall, I found the technique to be useful
Using this technique would make it more difficult to maintain archaeological models
Overall, I think this technique provides an effective solution to the problem of representing archaeological models
Using this technique would make it easier to communicate archaeological models to archaeologist
I would recommend the use of the technique
I think this technique allows describing cultural heritage information in an objective way
I think this technique introduce subjectivity in the description of cultural heritage information

Nine subjects out of ten thought that the technique is useful. Six subjects thought that the technique would reduce the effort required to represent cultural heritage information in relation to their previous experiences because "the description is formalized until the end" (subject number 12) and "intuitive" (subject number 21) and "could help minimize the problems caused by an over-interpretation" (subject number 51). Conversely, one subject thought that "to introduce more categories in a discipline with categories already well 'settled' may be at risk of increasing the margin of subjectivity" (subject number 11). Six subjects thought that the technique would allow keeping cultural heritage information up to date.

Eight subjects thought that the cultural heritage information represented with this technique would not be difficult for archaeologists to understand and would make it easier for researchers to find out if the information is correct because it avoids *a priori* classification and it is clear and accurate. Eight subjects thought that the technique provides an effective solution to represent cultural heritage information and would enable researchers to communicate better.

Eight subjects would recommend the use of this technique because it allows describing objects with less subjectivity. It would also allow defining first the basic aspects and then the more complex aspects, although it does not eliminate completely the need for interpretation.

For eight subjects, the valster modelling technique allows describing cultural heritage information objectively; at least, the valster technique introduced less subjectivity and ambiguity than other techniques. However, nine subjects thought that the valster technique introduced subjectivity in the description of cultural heritage information because subjectivity is impossible to avoid when selecting the properties and creating relationships.

4.3.3. Intention to use

Table IX presents the items about the intention to use.

Table IX. Items about the intention to use

I would be willing to use this technique in my professional/investigation context
Why?

Six subjects out of ten would be willing to use the valster technique in order to "avoid *a priori* classification" (subject number 42) and because it allows introducing "more objectivity, less ambiguity" (subject number 21). On the contrary, one of the beginners would not use it because "it costs a lot to organize the concepts" (subject number 22).

4.3.4. Quality of the produced models

Evaluating the quality of the models is a very hard task. We cannot evaluate the five produced models in a quantitative way as we would have needed many more models. However, the results of this evaluation are very positive and we need to link these results with the quality of the produced models. The produced models were evaluated in a qualitative way according to four criteria: syntax of the model, completeness of the model, use of categories or not, and the quality of the valsters in the model. The four criteria were measured as low, medium or high.

Firstly, the syntax of the model measured whether the produced models respected the prescribed syntax: out of the five models, two did not respect the syntax; two models respected it and one was a borderline case (the syntax of the valster modelling technique was not formally conveyed to the subjects). Secondly, the completeness of the model

measured the degree to which a model contained all the information available about the pot. Two models were very precise but did not represent all the information available; two other models represented the information in an adequate level of detail and in the expected scope of the case study. One model lacked some information and was not very precise. Thirdly, the use of categories was forbidden in this exercise, due to the very nature of valster modelling. Still, three models included categories such as "replica", "artefact" or "structure". Two subject pairs did not use categories in their models. Finally, the quality of the valsters in the models was evaluated according to how properties were arranged in clusters in the models. Two models presented no clusters of properties, or very few. Two models contained relevant clusters, and one model was halfway. As a conclusion, we can say that the quality of the models was medium, as two models had a pretty low quality, two were very good, and one was medium.

4.4 Synthesis of the evaluation

Table X presents the synthesis questions that were asked to the subjects in the last questionnaire, after the discussion.

Table X. Items about the synthesis

Overall, what did you think of this exercise?
Concerning the exercise you did, you are: very satisfied, satisfied, rather not satisfied, no satisfied at all?
What was the easiest thing to do?
What was the hardest thing to do?
Which are the strong points of the technique?
Which are the weak points of the technique?
Which are the improvements that could be done?

All the subjects found the exercise interesting and they were all satisfied (four subjects) or quite satisfied (six subjects) with the models produced.

The easiest thing to do was to select the properties from the list of properties we gave them. The most difficult thing to do was to establish relations between the objects (four subjects), to look for other properties (two subjects), to avoid thinking in general but in one concept in particular (one subject) and to know whether a property was subjective or not (one subject).

The strong points of the valster modelling technique, as reported by users, are:
—The capacity to organize and modulate the information (two subjects) without any prior knowledge (one subject),
—The capacity to minimize ambiguities without classifying the information *a priori* (five subjects),
—"The ability to divide, fragment information to the fullest extent possible" (subject number 12),
—The ability to describe information in a spontaneous way.
—The ability to describe information in a disorderly manner, relying on a list of properties that allows specialists not to start from scratch and not to forget properties, and for non-specialists to consider properties that they have not thought of.
—The ability to compare models made by different people and based on common properties.

The weak points are the introduction of subjectivity when selecting the properties: "When you conceptualize a category, you are always making a subjective operation" (subject number 11); by "category", the subject means "property".

The subjects proceeded as follow: "We have gone from the basic to the more specific" (subject number 32), "we tried to individualize the largest blocks of information and then we drew relationships between the blocks" (subject number 31) "We wanted to be clear from the start and try to distinguish between what are the properties of the object that interested us and the other objects related to this object" (subject 51).

Regarding the level of details of the models produced, it can be explained by "the consequence of what entails a determination of discipline. As we are archaeologists, we know that in general all the material must be treated in the most detailed and objective way as possible." (Subject number 21)

5. CONCLUSION AND FUTURE WORK

In this paper we present two qualitative evaluations of information modelling techniques for cultural heritage conducted at the Institute of Heritage Sciences (Incipit) of the Spanish National Research Council (CSIC). In the first experiment we evaluated the perceived ease of use, perceived usefulness and intention to use of two modelling techniques: UML class diagrams and conceptual graphs. We concluded that UML class diagramming was more adequate to cultural heritage modelling. Even though it is more complex than conceptual graphs, it allows a more precise representation of the cultural heritage information, which is fundamental to this domain.

The second experiment concerned the evaluation of the valster modelling technique, which allows users to create conceptual models without classifying concepts *a priori*. The results of this experiment are quite positive and let us think that the valster modelling technique is useful, particularly in cultural heritage, where classification made in advance should be reduced to a minimum and kept under strict methodological control.

From these experiments, we could recommend the use of the valster technique combined with the use of UML class diagramming. The valster technique could be used as a first step for cultural heritage modelling: specialists can define the concepts and their properties of their domain without classifying the concepts *a priori*. They can also define relations between the concepts. The model can be refined as long as concepts and properties have to be represented. Once the obtained valster model is satisfying, cultural heritage specialists can translate it to a UML class diagram. Indeed, the produced valster model should represent clusters of properties and relations between the clusters. By abstracting the clusters, cultural heritage specialists should then be able to define the appropriate corresponding classes. In the class diagram, they can specify the name of the concepts, the type of the properties and the name of the associations and their cardinalities. This two-step method has to be properly defined, tested and evaluated in order to measure if it allows cultural heritage specialists to define the concepts of their domain more pertinent and more easily than with other traditional information system development methods. We should evaluate this proposition through different archaeological projects lead at Incipit in order to measure its actual efficiency and actual effectiveness.

Moreover, we should try to relate Gardin (Gardin [1958], [1980] and Clarke (Clarke [1978]) theories to the use of UML class diagrams and the value cluster models in order to improve the latter. We should provide a way of modelling these concepts theories within UML class diagrams and value cluster models to help cultural heritage specialists during the modelling task.

As a further way to assist cultural heritage specialists in conceptual modelling tasks, a simple, affordable object-oriented conceptual modelling language, named ConML (ConML [2011]), was created with the humanities and social sciences in mind. ConML is defined through a simple metamodel containing the basic building blocks for static modelling, such as class, attribute, association, specialisation, enumeration and composition/aggregation. A 6-hour training course was organised at Incipit for cultural heritage specialists, where these concepts were presented. At the end of the course, attendees were able to understand ConML models and discuss about them. From a textual description of the main concepts of cultural heritage domain that they had written, they were able to create a model using ConML and criticise the different modelling solutions. Also, ConML has been used as a vehicle for a number of external training courses such as an annual post-graduate course run by Incipit through the CSIC's Tertiary Education Programme, and the Hands-on Archaeological Conceptual Modelling workshop at the Computer Applications and Quantitative Methods in Archaeology (CAA) conference.

Although a stable version of ConML is available, work is under way to add support to ConML for the modelling of temporality, subjectivity and perhaps additional "soft" issues such as vagueness and fragmentation. We hope that future versions of ConML will help bridge the gap between conceptual modelling and cultural heritage by removing perceived complexity while adding expressive power.

REFERENCES

CAMBRIDGE, 2010. Cambridge Advanced Learner's Dictionary, http://dictionary.cambridge.org/
CONML, HTTP://WWW.CONML.ORG/, 2011.
CHEN, P.P., 1976. The Entity-Relationship Model - Toward a Unified View of Data. *ACM Transactions on Database Systems*. 1, 9-36.
CROFTS, N., DOERR, M., GILL, T., STEAD, S. AND STIFF, M. (EDS). 2010. Definition of the CIDOC Conceptual Reference Model. Version 5.0.2, ICOM/CIDOC.
CLARKE, D. L. 1978. Analytical Archaeology, Second Edition, Methuen & Co Ltd.
DAVIS, F. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13 (3), 319-340.
EDMUNDS, H. 1999. The Focus Group Research Handbook, American Marketing Association.
GARDIN, J.C. 1958. Four Codes for the Description of Artifacts: An Essay in Archeological Technique and Theory. *American Anthropologist*, 60 (2), 335–357.
GARDIN, J.-C. 1980. Archaeological Constructs: An Aspect of Theoretical Archaeology. Cambridge University Press.
GONZÁLEZ PÉREZ, C, 1997. SIA+: Manual del Usuario, Grupo de Investigación en Arqueología del Paisaje, Universidade de Santiago de Compostela
GONZÁLEZ PÉREZ, C. 1999. Tecnologías de la Información y Patrimonio Cultural 1: El Paradigma Orientado a Objetos. *CAPA: Cadernos de Arqueoloxía e Patrimonio*, 10, 1-34.
GONZÁLEZ PÉREZ, C. AND HENDERSON-SELLERS, B., 2007. Modelling Software Development Methodologies: A Conceptual Foundation. *Journal of Systems and Software,* 80, 1778-1796.
GONZÁLEZ PÉREZ, C., 2002. Sistemas de Información para Arqueología: Teoría, Metodología y Tecnologías. *BAR International Series*, 1015. Archaeopress, Oxford.
GONZALEZ-PEREZ, C. 2012. Typeless Information Modelling to Avoid Category Bias in Archaeological Descriptions. In Chrysanthi, A., Murrieta Flores, P. & Papadopoulos, C. (eds.) *Thinking Beyond the Tool: Archaeological Computing & the Interpretive Process*. Archaeopress. In press.
GOULD, S. J., 1990. Wonderful Life: The Burgess Shale and the Nature of History. Norton & Company.
HUG, C., 2009. Méthode, modèles et outil pour la méta-modélisation des processus d'ingénierie de systèmes d'information, PhD thesis, Joseph Fourier University, Grenoble, France.
HUG, C., FRONT, A. AND RIEU, D., 2010. A Method and a Tool Based on a Conceptual Graph for Information Systems Engineering Processes. In *Proceedings of the 5th International Working Conference on Evaluation of Novel Approaches to Software Engineering* (ENASE 2010), Athens, Greece.
HUG, C., MANDRAN, N., FRONT, A. AND RIEU, D., 2010b. Qualitative evaluation of a method for information systems engineering processes. In *Proceedings of the Fourth International Conference on Research Challenges in Information Science (RCIS 2010), Nice, France, IEEE,* 257-268.
ISO/IEC 19501, 2005. Unified Modeling Language (UML) Version 1.4.2.

LAKOFF, G. 1990. *Women, Fire, and Dangerous Things*. University of Chicago Press.

LIKERT, R., 1932. A technique for the measurement of attitudes, *Archives of Psychology*, 22, 140, 1-55.

MOODY, D. L. 2003. The method evaluation model: a theoretical model for validating information systems design methods. In *Proceedings of the 11th European Conference on Information Systems* (ECIS 2003), Naples, Italy.

NACHMAIS, D, NACHMAIS FRANKFORT C. 2008. Research methods in the Social Sciences: Seventh Edition. Worth Publishers.

OMG. 2010. UML Version 2.3, Superstructure.

RESCHER, N. 1997. Methodological Pragmatism: Systems-Theoretic Approach to the Theory of Knowledge, Basil Blackwell, Oxford.

SEARLE, J. 1990. The Mystery of Consciousness. New York Review Books.

APPENDIX

Figures 4 and 6 were originally produced in Spanish; they are translated here.
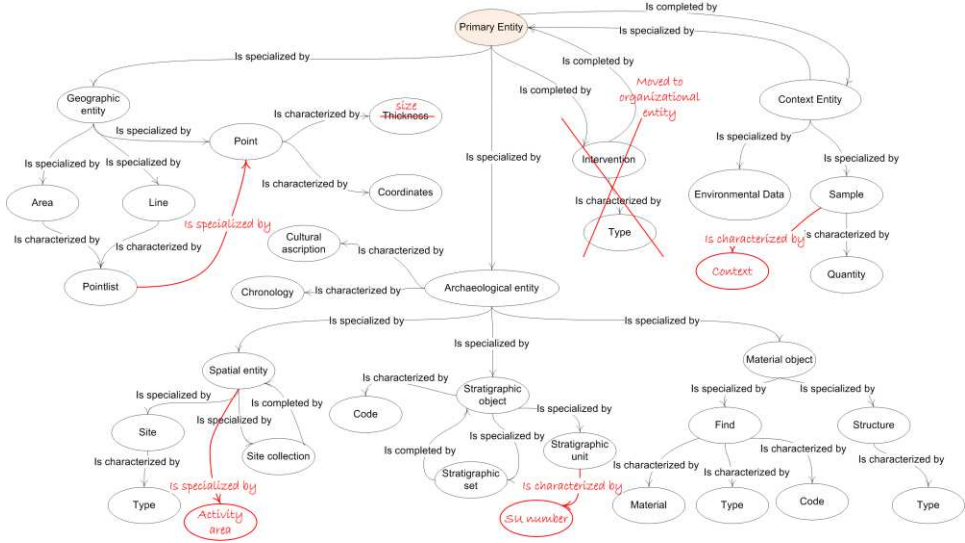

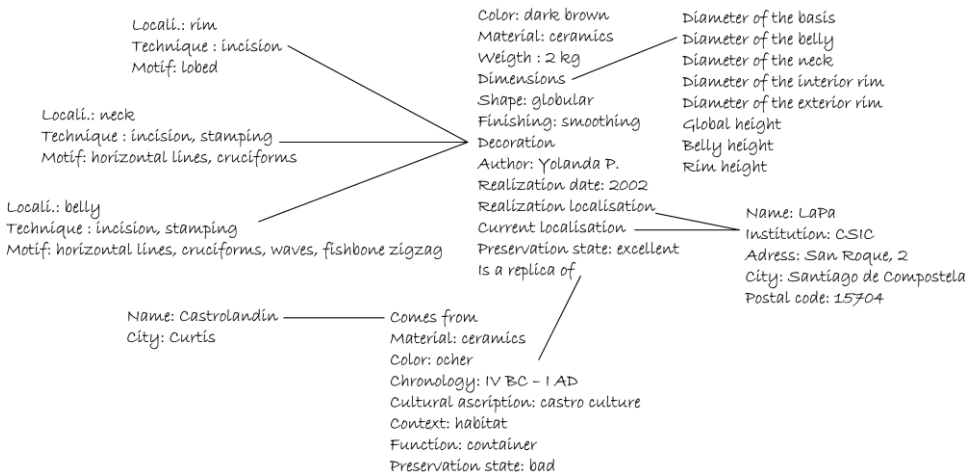Fig. 4. Extract of the solution proposed by one subject (trad.)

Fig. 6. Valster model produced by a pair of subjects (trad.)

ACKNOWLEDGMENTS