A Formal Language for the Description of Historical Architectural Elements

Gonzalez-Perez, C.[1], Blanco-Rotea, R.[1], Mato, C.[1], Camiruaga, I.[2]

[1] Institute of Heritage Sciences (Incipit), Spanish National Research Council (CSIC), Spain
[2] Built Heritage Enhancement Office, Consorcio de Santiago de Compostela, Spain
{cesar.gonzalez-perez, rebeca.blanco-rotea}@incipit.csic.es, cristina.mato.fresan@gmail.com, idoia@coag.es

The study of heritage buildings and historical architectural elements in general can be approached from different perspectives, such as material, stratigraphic, and chronologic viewpoints. It is often hard to integrate different perspectives together and obtain a holistic understanding that encompasses heterogeneous, but connected, information. This paper introduces LDEA, a formal language for the description of historical architectural elements, which defines a collection of concepts, relationships and syntactic rules capable of describing any complex heritage building from an arbitrary combination of material, stratigraphic and chronologic perspectives. LDEA can help archaeologists, architects and heritage managers work together in an integrated and holistic fashion, achieving a multi-vocal understanding of the buildings that is rarely obtained by using conventional techniques that employ a single perspective and a non-formalised approach.

Keywords: Formal language, Archaeology of buildings, Ontology, Multi-vocality.

1. Motivation

The study of heritage buildings and other historical architectural elements usually engages professionals of different fields, such as archaeologists, architects, art historians and heritage managers. The dominant paradigm and tradition in each field usually shapes how one looks at a building, what kind of information is recorded, what kinds of abstractions are considered appropriate, and what the expected outcomes of a study are. As in any other endeavour where multiple fields cooperate, the synergy of their interaction can only be attained if the necessary communication mechanisms are in place. By “communication mechanisms” we mean not only the actual notations (graphical, textual, oral or otherwise) that are used to convey meaning, but also the shared set of conceptual building blocks that individuals use to compose models of the information being dealt with. These conceptual building blocks determine what is sometimes called the discipline- or domain-specific universe of discourse, which is, in turn, usually communicated via a highly specialised terminology (the notation).

As long as communication occurs within any particular field, misunderstandings are few, since the universe of discourse and terminology are well known. When different fields need to interact in a genuinely multi-vocal discourse, however, misunderstandings are frequent, not only because of terminological problems, but (and more importantly) because of conceptual differences and mismatches. The authors were victims of exactly this problem while working in the City of Santiago´s Heritage Information System, which involved architects, art historians, archaeologists, computer specialists, heritage managers and engineers, each of which uses its own universe of discourse to describe the same things.

This problem is often tackled in one of two ways: by adding formalism to the concepts and notation being used, so that ambiguity is reduced and misunderstandings minimised; or by introducing a shared or standardised set of concepts and notational artefacts that can be used seamlessly across multiple communities or fields. Formalisation and standardisation are thus two mechanisms which, when used in conjunction, can help decrease significantly the problems inherent to multidisciplinary work in the study of heritage buildings.

This paper introduces LDEA (*Lenguaje para la Descripción de Elementos Arquitectónicos* in Spanish; Language for the Description of Architectural Elements), a conceptual tool that uses formalisation and standardisation to facilitate the integration between the

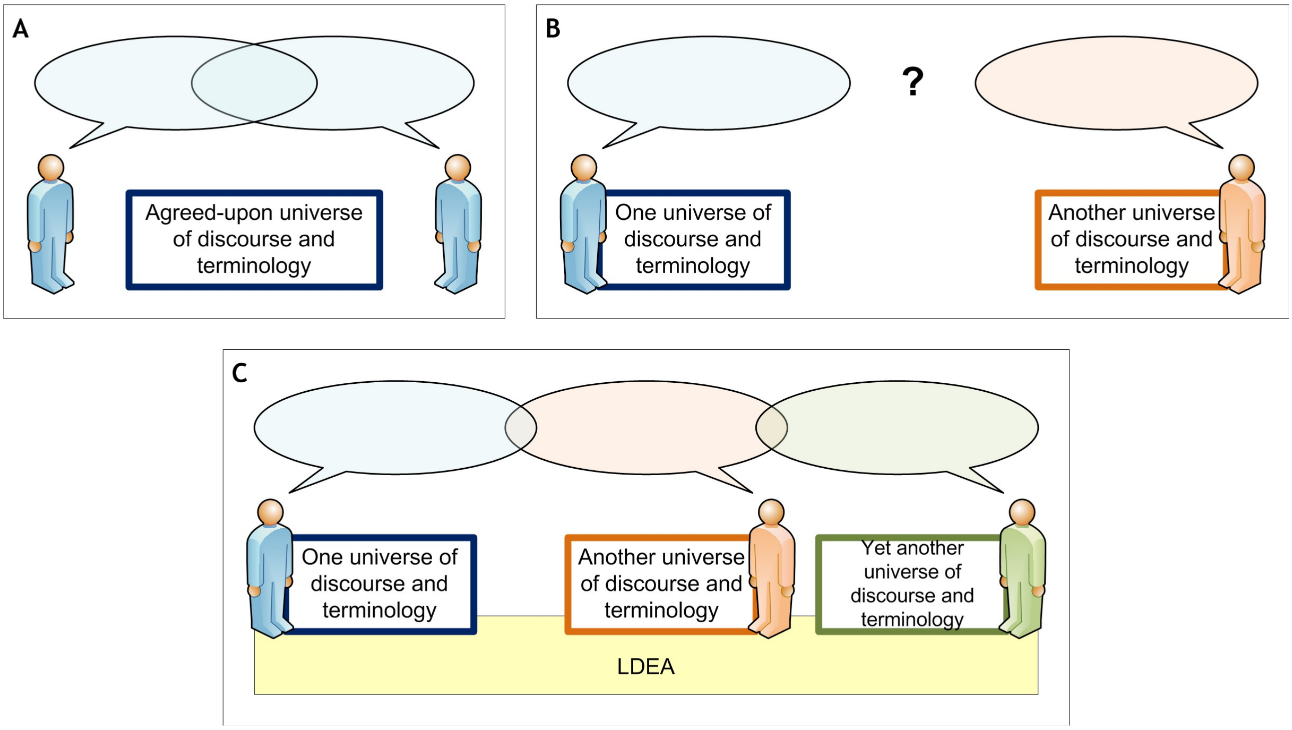

Figure 1: *LDEA goals and motivation. (A) shows communication taking place within a single discipline. (B) shows communication problems and lack of integration when two different fields or disciplines suffer from terminological issues and conceptual mismatches. (C) shows how LDEA can provide a formalised and standardised infrastructure that supports integration and communication between disciplines.*

different fields that usually work together on heritage buildings and related elements (Figure 1). Describing the complete LDEA would need more space than is available, so this paper offers an overview of the language followed by a detailed description of a subset of it. Also, and for the same reason, this paper focuses on the conceptual side of the work, leaving notational issues for a future publication.

LDEA describes information along three different axes: materiality, because this is how things are built in architecture; stratigraphy, because this is how materiality gets organised into temporal sequences; and chronological, because this is cross-cutting to the other two and gives them an absolute temporal context.

The next section describes the structure of LDEA, accounting for the three axes that compose the core of the language. This is followed by a case study that illustrates how the three axes work together. Then results are discussed and alternative approaches are taken into account. Finally, conclusions are presented and future work directions are described.

2. Structure of LDEA

LDEA describes information along three different axes: *materiality,* because this is how things are built in architecture; *stratigraphy*, because this is how materiality gets organised into temporal sequences; and *chronology,* because this is cross-cutting to the other two and gives them an absolute temporal context. The three axes are interconnected in LDEA so that information belonging to one axis can be reached from associated information in any of the other two axes.

2.1. Information modelling approach

LDEA uses a semi-formal approach for the definition of information. This entails two aspects:

- The conceptual building blocks that are often used by professionals working on heritage buildings are defined as concepts, and unambiguous terms are chosen to depict them.
- The syntactic rules that govern how these concepts can be combined in order to express complex constructs are also defined unambiguously as relationships.

Some of the concepts and relationships in LDEA belong to the material axis; some others belong to the stratigraphic axis; and some others belong to the chronological axis.

The technique used to define the concepts and relationships in LDEA is *class modelling*. This is a technique often used in software engineering to create information models that represent the structure of a certain subset of the observed reality from a static perspective, *i.e.* focusing on the entities (or concepts) that compose such a reality plus the relationships that exist between them. The terms "class" alludes to the fact that what are represented in the model are types of things that may exist in the perceived reality (*i.e.* classes) rather than specific instances. For example, a class model in the archaeology domain may contain classes such as “Site” or “Artefact” rather than specific sites or artefacts such as “Stonehenge” or “the Phaistos Disc”.

Class modelling is found as a part of the ISO standard ISO/IEC 19501, also known as Universal Modelling Language (UML) version 1.4.2 (ISO/IEC, 2005). In this paper we use ISO/IEC 19501 to depict class diagrams in order to express fragments of LDEA, although an alternative approach to this is described in Section 4. A deeper description of class modelling is out of the scope of this paper. A comprehensive and more technical treatment of classes, attributes and relationships can be found in, for example, ERIKSSON and PENKER, 1998, ch. 4; or RUMBAUGH *et al.*, 1991, ch. 3.

2.2. High-level concepts

Figure 2 shows a very abstract view of LDEA, including the top-level concepts that compose the foundations for the rest of the language. Each box in the diagram (except those on the right-hand side, labelled with the “enumeration” keyword; we discuss these later) represents a particular class of thing. For example, the diagram states that, according to LDEA, every construction must have an identifier, a name and a description, all of which are texts. Furthermore, every construction must have a usage, which is of type “Construction Usage”. This type is represented as an enumeration of possible values on the right-hand side of the diagram as a separate box. Each of the individual entries inside a class box (such as “Identifier”, “Name”, “Description”, etc. is called an attribute of that class.

The lines that join in a triangular arrowhead from “Construction Element” and “Construction Set” and into “Construction” represent a generalisation; this means that “Construction” is a generalisation of the more specific concepts “Construction Element” and “Construction Set”. In other words, “Construction Element” and “Construction Set” are specialised concepts of “Construction”. From a conceptual point of view, triangular arrowheads depict a subtyping relationship, in which the class pointed to by the arrow is a “supertype” or more abstract entity, and the source classes of the arrows are the “subtypes” or more concrete entities. Thus, generalisation/specialisation hierarchies can be expressed by concatenating multiple levels of abstraction, as shown in Figure 2 by the further specialisation of “Construction Element” into the even more concrete “Building” and “Open Space”.

An interesting consequence of specialisation hierarchies is that, by the very meaning of specialisation itself, attributes are “inherited” from any one class to every class specialising from it. For example, and according to Figure 2 again, since “Construction Element” is a specialisation of “Construction” and “Construction” has “Name” and “Usage” attributes (among others), then “Construction Element” automatically inherits these attributes and can be considered to have these properties as well. Now, since “Building” specialised from “Construction Element”, it inherits the attributes in turn. Specialisation hierarchies thus propagate attributes downstream by means of inheritance. This is extremely convenient to specify information at a high level of abstraction (usually shared between disciplines) without needing to worry about the specific details, which tend to differ between fields.

The last important aspect to be highlighted in Figure 2 is the lines between classes that start with a diamond-shaped arrowhead. These represent whole/part relationships, *i.e.* relationships between classes where one class (the one pointed at by the diamond-shaped end) is a whole and the other class is a part. Whole/part relationships are ubiquitous in most systems (trees in a forest, pages in a book, bricks in a wall, arguments in a thesis), and Figure 2 contains two good examples. On one hand, it is stated that every construction set must contain one or more constructions; on the other hand, it is stated that every built nucleus may be made of multiple constructions. The small numbers placed next to the end points of the lines represent the multiplicity of the whole/part relationships, i.e. the minimum and maximum numbers of specific instances that may be involved in the relationship at any point in time.

2.3. Some choices taken

Since transcending boundaries of conventional disciplines is a major goal of LDEA, there is no use in defining concepts such as “street” or “city”. If this had been done, too many exceptions would have been found to these concepts, making the organisation and the relations between classes too complicated. In order to satisfy its objectives, LDEA demands general considerations, and with this premise in mind, it moves from very abstract statements to more concrete ones.

To start with, the “Built Nucleus” class refers to a group of constructions that cluster together around a particular geographical location, but can be applied in many different contexts: a city, a street, a hillfort, a village, a castle, etc. A built nucleus is composed of constructions (Figure 2), each of which may be either elementary or decomposable. A house, for example, is typically an atomic construction, as it cannot be divided into smaller constructions and has a well-defined sense of unity. On the contrary, a monastery is a complex system, i.e. “a group of interacting, interrelated or interdependent things or parts, forming a complex or unified whole, especially to serve a common purpose” (CHING, 1996: 21), and is therefore, better represented as a construction set composed of other constructions (each of which may be, in turn, elementary or complex). The “Construction Set” class is a generic grouping mechanism that can be used to arbitrarily collect together constructions, depending on the needs of the study or work being done.

The most typical construction elements that may come to mind are buildings and, consequently, “Building” specialises from “Construction Element”. There are, however, other kinds of construction elements that must

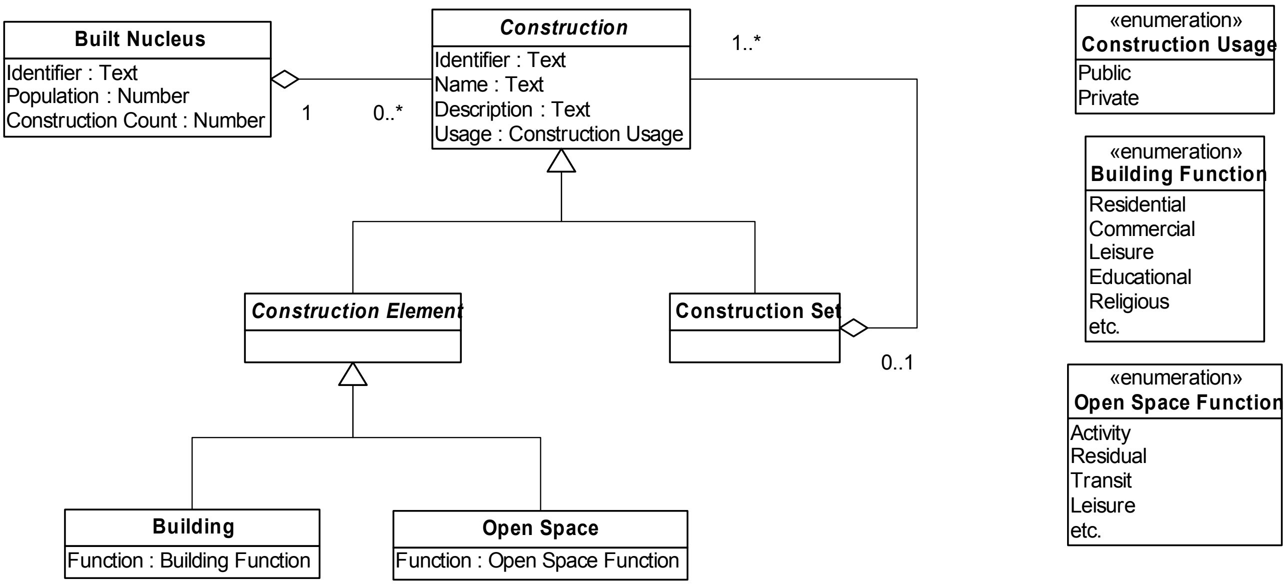

Figure 2: *A class diagram expressed using UML (ISO/IEC 19501) notation. Boxes depict classes (i.e. concepts in the universe of discourse) whereas lines and arrows depict different kinds of relationships between those concepts. Text inside the boxes may represent class names (centred, on the top) or attributes (lower section of the boxes).*

be borne in mind. The concept of "Open Space" is used in LDEA to account for all those spaces that, without a material reality, actually shape a place, where "space" implies the possibility for any piece of land or reality to become a place; in this respect, we follow Tuan´s point of view (TUAN, 2001). A town square, for instance, is formed by buildings or structures that surround and enclose a given area. There is an absence of material but the existence of the square cannot be denied. The "Open Space" class in LDEA fulfils this need.

The few classes described so far set the foundation for the 120-odd classes that make up the complete LDEA specification. The next sections present the details of some of these classes along each of the three axes.

2.4. Materiality aspects

Describing the materiality of a heritage building to a degree of detail that is of value is a daunting task, and involves a large number of classes and relationships. For the sake of simplicity, this paper portrays only a small subset of the classes that LDEA uses to implement the material aspects of heritage building description.

Figure 3 shows that buildings in LDEA are composed of elements, which may be horizontal, vertical or transit. "Wall" is a specific type of "Vertical Element", together with "Opening" and "Pillar". "Opening" represents any opening through a vertical element, such as a window or a door. The "Wall Leaf" class, representing each of the leaves of a particular wall, is also considered a particular subtype of "Vertical Element". Notice that whole/part relationships are used to express that every wall may contain multiple openings (but every opening belongs to a specific wall), and every wall must be composed of one or more leaves (and each leaf belongs to one particular wall). Each leaf in a wall may be interior, exterior or indifferent. This is just a small sample of the class hierarchy in the material axis.

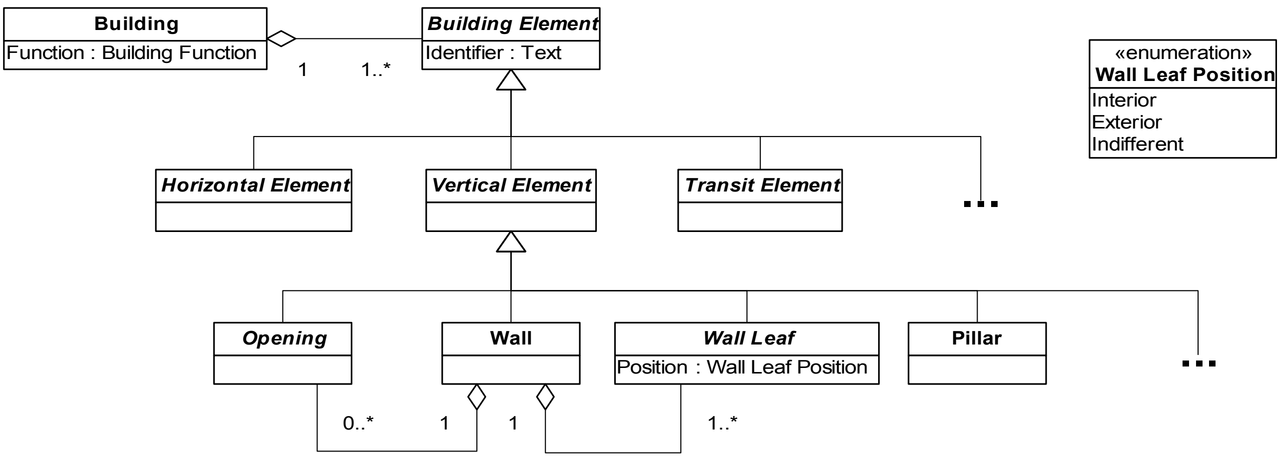

Figure 3: *Main classes related to building elements in LDEA. The "Building" class on the top left of the diagram is the same class that appears on the bottom left in:Figure 2.*

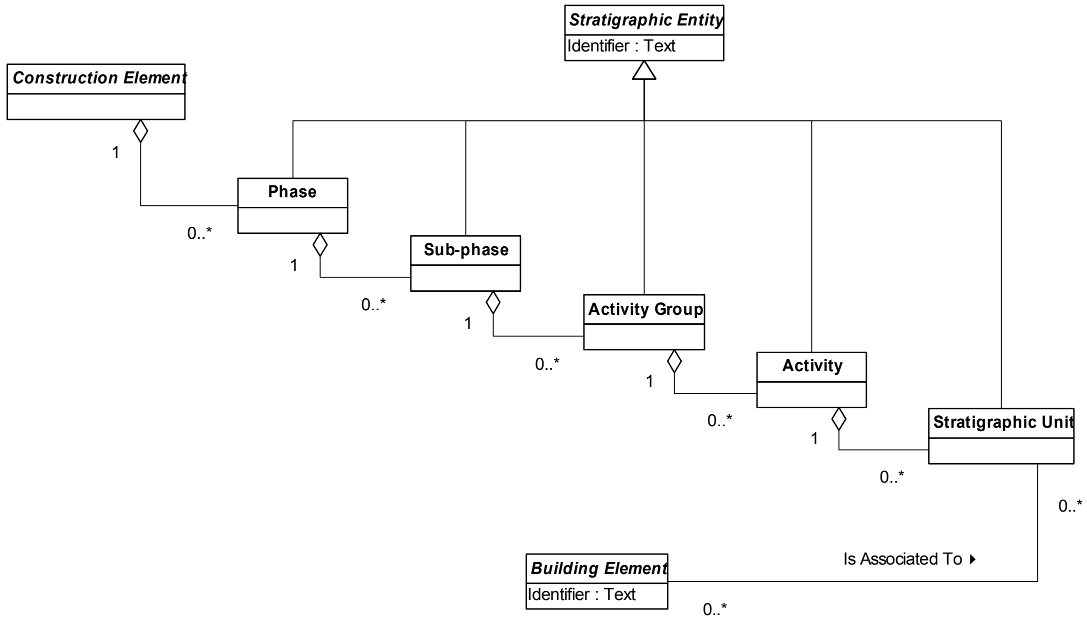

Figure 4: *Main classes in the stratigraphic axis of LDEA. The "Construction Element" and "Building Element" classes, also present in Figure 2 and Figure 3, relate stratigraphic information to material information.*

2.5. Stratigraphic aspects

From the point of view of the stratigraphy, LDEA tries to deconstruct the materiality of the building in order to comprehend the same matter within a completely different conceptualisation. In fact, architects and archaeologists often look at buildings in different ways and, as a result, see different things, the matter being the same.

The classes that LDEA uses to describe the constructive sequence, as shown in Figure 4, are: "Stratigraphic Unit", "Activity", "Activity Group" (CARANDINI, 1997:139; PARCERO OUBIÑA *et al.*, 1999: 34), "Sub-phase", "Phase" and "Period" (HARRIS, 1979). The "Stratigraphic Unit" class represents the smallest units that can be recorded within a building that have an individual nature and are homogenous in stratigraphic terms. Stratigraphic units, corresponding to "upstanding strata" according to Harris (HARRIS 1979: 37-38), can be classified into stratigraphic elements and interfaces. They are formed by constructive or destructive actions, and have materiality and volume in the case of the stratigraphic elements, and only volume in the case of the interfaces (HARRIS, 1979: 43-47).

The "Activity" class represents any aggregate of stratigraphic units that have, or once had, the same function and belong to the same chronological period, thus representing a particular stage within the stratigraphic series of the building. The "Activity Group" class represents any homogenous collection of activities that corresponds to a complex context, having constructive, spatial, temporal and functional cohesion. In the case of simple structures, the whole building may comprise a single activity group, whereas in the case of more complex structures there may be various activity groups composed of activities related to the same structural function.

The "Phase" class represents each episode of construction, use and abandonment of a building, including a number of stratigraphic units and activities, and the stratigraphic relationships between them. This concept has a strong interpretive nature, and may be associated to one or multiple activities or activity groups depending on the complexity of the building. Also, within the same moment of construction, use or abandonment identified in a building, other constructive actions may have occurred on a lesser scale, and connected to a homogenous constructive phase. These are represented by the "Sub-phase" class.

This overall classificatory approach is commonly known as stratigraphic analysis, stratigraphic method or Harris method (BROGIOLO, 1988: 35; HARRIS *et al.*, 1993: 87). LDEA adopts this approach and organises stratigraphic information accordingly. The "Construction Element" and "Building Element" classes serve to link information in the stratigraphic axis to information in the material axis; in fact, these classes appear both in Figure 4 and in Figure 2 and Figure 3, thus bridging the material and stratigraphic realms.

2.6. Temporality aspects

As time goes by, changes that occur to the observed reality are supposed to be reflected on the associated information model. For example, if a new doorway is opened through an old wall as part of the refurbishment works on a heritage building, the changes to the material and stratigraphic perspectives, plus the interconnections between them, must be duly recorded. This could mean, for example, adding new data entities to the model (such as an instance of the class "Opening" in Figure 3 and a few instances of the class "Stratigraphic Unit" in Figure 4). It could also entail the modification of some attribute values to existing data entities; for example, an existing

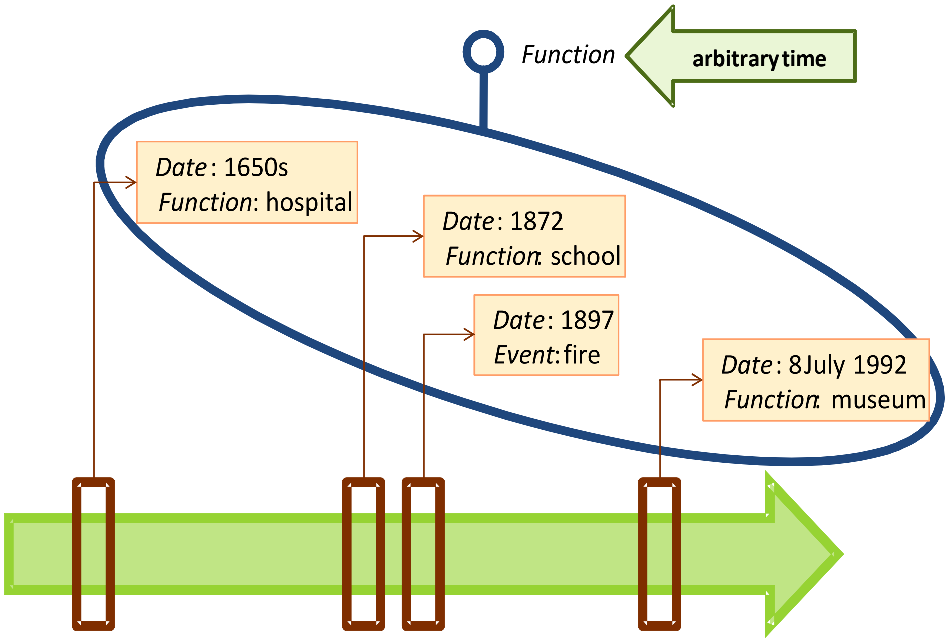

Figure 5: *An attribute of an information entity as a temporal aggerate of a sequence of versions. The horizontal green arrow represents time. The vertical brown bars depict snapshots with associated data.*

wall leaf that was so far interior might become exterior after the new doorway is opened.

In general, changes to the information model over time are captured by LDEA by assuming that *time is an extra dimension of information*. This means that the information model that we have described so far represents a snapshot of the observed reality (namely, its materiality and stratigraphy) at a particular point in time, and that only by "stacking" multiple snapshots along a timeline a complete, living history of the represented elements can be obtained. LDEA takes on a "timeless" approach to object identity that meshes well with class modelling, as posited by (PARTRIDGE, 2005: ch. 7). By "timeless" here we do not mean that time is absent from the information model; rather, we mean that time is so pervasively present that it is not explicitly visible as part of the domain-specific concepts that are captured as classes. In fact, LDEA includes no classes to represent time itself; to the contrary, every class, attribute and association is considered to be "versionable" over time so that their instances can vary as additional snapshots are added to the information model.

Figure 5 shows the time line for a fictitious building of which four snapshots have been taken. Each snapshot is depicted by a date plus some information of value, such as the function of the building at that time or a significant event that happened to the building. From this perspective, the "Function" attribute of the building (introduced in Figure 2) is not an atomic property in the information model, but a whole dimension the value of which changes depending on the arbitrary value that the time variable may take; in other words, and from this timeless perspective, the function of the building is only determined when a particular time is selected.

3. Case study

As an example of the use of LDEA, the façade of the house situated in number 46 of Rúa do Vilar in Santiago de Compostela (Galicia, Spain) was analysed using the three proposed axes: material, stratigraphic and chronological.

In terms of materiality, the built nucleus corresponds to the historical centre of Santiago de Compostela. Within this we have singled out a particular building (number 46 of Rúa do Vilar). Looking at its façade, we have selected a vertical element – the wall – and from it the exterior leaf. The constructive system used for this leaf is masonry, and, in particular, square-faced stone blocks. From this point we could continue to define the components of the stone masonry to arrive to the specific materials (granite), the finish used, the type of pointing, the mortar used, and other particular details captured by the LDEA class model.

Figure 6 shows a photograph of the chosen building, on which two building elements (a wall leaf and a window) have been highlighted using coloured shades. In addition, the stratigraphy of the façade is displayed both as a Harris matrix and on the façade itself. It is very clear from the figure that the mappings between the material axis (i.e. building elements) and the stratigraphic axis (i.e. stratigraphic units) are many-to-many; this means that each individual building element may correspond to multiple stratigraphic units, and each individual stratigraphic unit may correspond to multiple building elements. An example of the first is the wall leaf highlighted in yellow, which is likely to be treated as a single whole from a material perspective, but is seen as a complex composite of multiple stratigraphic units when observed from the stratigraphic point of view. An example of the second case is stratigraphic

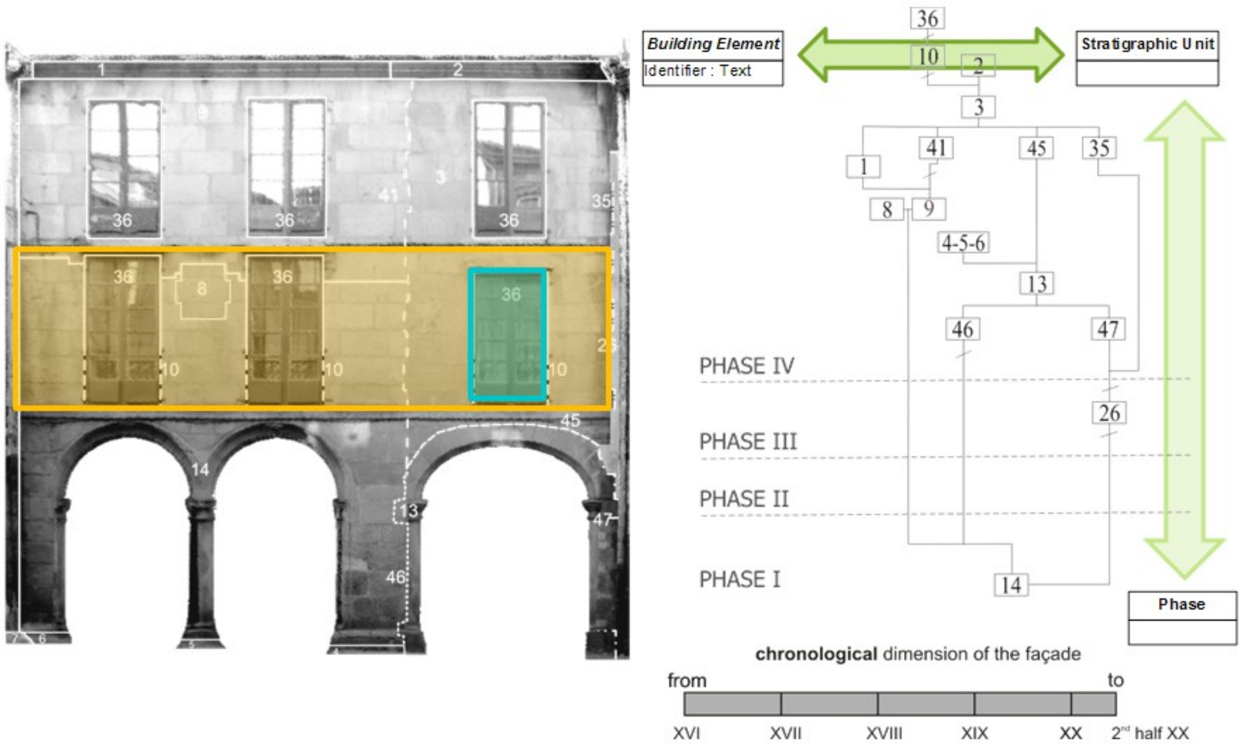

Figure 6: *46 of Rúa do Vilar showing two building elements (highlighted in yellow and green over the photograph) side by side with the stratigraphic matrix of the façade. The mapping between the material (building elements) and stratigraphic (stratigraphic units) axes is a many-to-many one, achieved through the associations shown in Figure 4.*

unit number 36, which appears associated to various material instances.

The class model of LDEA, as shown in Figure 4, bridges the gap between the stratigraphic and material perspectives by formally linking classes in both realms but, at the same, preserving the language that is particular to each of them and allowing architects, archaeologists and other professionals to approach their work from within a familiar conceptual framework.

4. Discussion

There are several aspects of LDEA that are worth considering from a self-critical perspective. To start with, a large part of the effort that was put in the development of LDEA was directly aimed to the construction of what we could call an ontology for historical architecture. The word "ontology" is nowadays often used in the context of software and knowledge engineering as closely related to "model" (ATKINSON *et al.*, 2006); the concept of ontology has emerged from the work in artificial intelligence and is often associated with the capability of machine reasoning on the information model, whereas "plain" conceptual models, such as the one presented here, are intended only for human consumption. WE must indicate that LDEA is not an ontology, but works in the same area, and if directed to the development of ontologies, might yield results that are not too different in the range of concepts that they contain and the way they are structured.

A second aspect of LDEA of which we are critical is the fact that it is currently expressed in UML, as described in Section 2.1. UML was chosen because it is an ISO standard and because it is well known, despite being too oriented towards programming and being overly complex. After the LDEA project was well under way, the authors developed an alternative language, called ConML (Conceptual Modelling Language) (INCIPIT, 2011a; INCIPIT, 2011b), which is of much higher level of abstraction than UML and has been purposely designed to be affordable and easily learnable by people with no previous exposure to information technologies. ConML has since been tested with excellent results, which are in the way of being reported through publication. We are considering the migration of the specification of LDEA from UML into ConML for the sake of understandability and ease of use by non-IT specialists.

Conclusions and future work

By providing a semi-formal approach to the definition of concepts and relationships spanning three different axes (material, stratigraphic and chronological), LDEA can inter-relate information entities that are not always put together, enhancing the collaboration of professionals in different fields related to historical architecture.

At present, LDEA has become part of a larger project that aims to develop an integral methodology for the description, assessment and interpretation of cultural heritage elements; this methodology will include a formal language, composed of concepts and relationships, plus a collection of methodological guides about its usage. LDEA will become embedded in the described language, together with other conceptual areas that are of interest to cultural heritage. In particular, multi-vocality and temporality management, which have been presented in this paper in the context of LDEA, are also core aspects of this larger project.

Acknowledgements

The work described in this paper has been partially funded by project MIRFOL under grant number 09SEC002606PR, INCITE Programme, Xunta de Galicia, Spain. The authors are grateful to the participants in the project for the City of Santiago Heritage Information System, which originated some aspects of the work presented in this paper.

References

ATKINSON, C., GUTHEIL, M. & KIKO, K., 2006. On the Relationship of Ontologies and Models. In *Proceedings of the 2nd International Workshop on Meta-Modelling* (WoMM). LNI 96. Karlsruhe, Germany. 47-60.

BROGIOLO, G. P., 1988. *Archeologia dell'Edilizia Storica*, Como, Edizioni New Press.

CARANDINI, A., 1997. *Historias en la Tierra. Manual de Excavación Arqueológica*, Barcelona, Crítica.

CHING, F. D. K., 1996. *A Visual Dictionary of Architecture*, Wiley.

ERIKSSON, H.-E. & PENKER, M., 1998. *UML Toolkit*, New York, John Wiley & Sons.

HARRIS, E. C. (1979) *Principles of Archaeological Stratigraphy*, London, Academic Press.

HARRIS, E. C., BROWN III, M. R. & BROWN, G. J., 1993. *Practices of Archaeological Stratigraphy*, London, Academic Press.

INCIPIT, 2011a. *ConML Technical Specification. ConML 1.1.* http://hdl.handle.net/10261/36481

INCIPIT, 2011b. *ConML Web Site* (web site). Accessed on 24 June 2011. http://www.conml.org

ISO/IEC, 2005. *Unified Modeling Language (UML)* version 1.4.2. ISO/IEC 19501: 2005.

PARCERO OUBIÑA, C., MÉNDEZ FERNÁNDEZ, F. & BLANCO ROTEA, R., 1999. *El Registro de la Información en Intervenciones Arqueológicas*, Santiago de Compostela, Spain, Grupo de Investigación en Arqueología del Paisaje.

PARTRIDGE, C., 2005. *Business Objects: Re-Engineering for Re-Use*, The BORO Centre.

RUMBAUGH, J., BLAHA, M., PREMERLANI, W., EDDY, F. & LORENSEN, W., 1991. *Object-Oriented Modeling and Design*, Englewood Cliffs, NJ, Prentice-Hall.

TUAN, Y.-F., 2001. *Space and Place: The Perspective of Experience*, University of Minnesota Press.