CIDOC CRM and Epigraphy: a Hermeneutic Challenge

Achille Felicetti[1], Francesca Murano[2], Paola Ronzino[1], and Franco Niccolucci[1]

[1] PIN, VAST-LAB, Prato, Italy
[2] Università degli Studi di Firenze, Italy
{achille.felicetti,paola.ronzino,franco.niccolucci}@pin.unifi.it
francesca.murano@unifi.it

Abstract. This paper identifies the main concepts involved in the study of epigraphy and proposes the use of CIDOC CRM to encode epigraphic concepts and to model the scientific process of investigation related to the study of epigraphy. After analysing the existing CIDOC CRM entities and those provided by the CRMsci and CRMarchaeo extensions, we introduce more specific epigraphic classes to be used as the basis for creating a new extension, CRMepi, which is more responsive to the specific needs of this discipline.

Keywords: CIDOC-CRM extension, Epigraphy, EpiDoc, CRMepi

1 Introduction

Material sources bequeathed to us from antiquity represent an immense treasury of knowledge of a lost world. In addition to literary sources, numismatics, epigraphy and archaeology provide important information we would not otherwise possess, and often the objects they document constitute the only evidence we have of an ancient population. In particular, among these sources, inscriptions are particularly essential because they represent the most direct and resonant voice of the ancient people that have handed them down to us, as deathless words in stone. Generally, the preserved inscriptions were originally conceived to endure over time and to be transmitted to the future. For these reasons, these sources need to be digitised and integrated in some way, along with other cultural heritage information. Inscriptions are complex objects, and their study requires careful analysis from different points of view.

What characterises this class of objects is that they form a whole with their physical support. Indeed the meaning of an epigraph cannot be fully understood without the analysis of the object or monument or other archaeological object on which it appears, just as one cannot fully understand the nature of that particular archaeological object without thoroughly investigating the sense of the inscription or iconographic representation it hosts.

Furthermore, the inscription itself is, from a conceptual point of view, an element with physical characteristics that are themselves bearers of meaning and of valuable information going far beyond the inherent meaning of the text. The

shape of the letters, their spacing, the direction, technique and other similar characteristics provide precious clues to the times, makers and functions of the inscription in question.

The tools available today seem to be insufficiently flexible and efficient to implement a comprehensive and useful level of digitisation. Relational databases, for instance, are not fully functional for this goal because of the rigid structure they provide, which is quite inadequate to describe textual entities and their possible annotations. Databases are even less suitable to describe the complex web of relationships that links archaeological objects with the inscriptions they bear and their meaning. Since the beginning of the studies on digitisation of texts and other unstructured data, the XML family of technologies emerged as the tool that could free the information from the rigid structures of relational databases built around tables and records. Epigraphy has also largely benefitted from XML. It is sufficient here to say that EpiDoc [1], the standard currently used to encode epigraphic information in electronic format, is also based on XML, the guiding principles of which it is both the bearer and the beneficiary. However, in the modern digital world, where the imperatives are those of interoperability and integration, the use of more efficient tools such as ontologies and conceptual models seems to be of crucial importance. In this paper, after investigating the problem, identifying the key concepts involved and giving an overview of the existing solutions, we will try to give a coherent description of the new possibilities offered by semantic tools to deal with epigraphic entities. In particular, we will use the intellectual model of CIDOC CRM and its extensions (CRMsci and CRMarchaeo) to provide them with shape and consistency, and to try to sketch a new CIDOC CRM extension (CRMepi) to be used in epigraphic studies.

2 What is Epigraphy?

2.1 Epigraphs and Epigraphy

Although epigraphy is a scientific discipline with a centuries-old tradition, a single and fully accepted definition of its object, the epigraph (or inscription), has not yet been formulated. The definition most commonly found in manuals of the discipline defines the inscription more or less as a direct evidence of the past inscribed on stone or other durable materials. Guarducci, in particular, puts emphasis on the materiality of the support, which, among other things, makes it possible to distinguish the object of epigraphy, as opposed to papyrology. She identifies what epigraphy is concerned with, as opposed to papyrology, and she identifies two particular characteristics of epigraphic documentation that differentiate it from other historical primary sources: uniqueness and authenticity [2]. More recently, Panciera has proposed to define an epigraph as "any particular type of written human communication of the sort that we would today call unidirectional, in the sense that it does not anticipate that a response will be provided to the sender, and which has the characteristic of not being addressed

to a person or to a group but to a collectivity, and which for this reason is made with the location, writing technique, graphic form and impagination, mode and register of expression chosen because they are most suitable to the attainment of its intended goal, and which differentiates itself in this manner from other forms of contemporary verbal communication (oral, literary, or documentary)" [3]. This definition is undoubtedly more complete than the previous ones, but is still not sufficiently exhaustive with regard to the huge variety of documents that are the object of epigraphic studies and may be included in epigraphic *corpora*.

The problem originates from the plethora of tools, techniques and purposes that can be employed in order to characterise an inscription. These constitute a complex and intertwined series of elements that do not, however, suggest a comprehensive definition. We must also consider that different cultures have given rise to different traditions of writing and thus also of epigraphy. Consequently, a consistent and comprehensive definition of the term "inscription" should consider not only epigraphic products from the major Western traditions (Latin and Greek), but also those relevant to "minor" traditions (e.g., the languages of ancient Italy) and non-Western ones (e.g., the heterogeneous legacy of the Semitic world). In addition, we should consider that there is an ambiguity of reference with respect to the term "epigraph", which can be used to indicate both the plain text and the combination of text and physical support, especially in those cases in which the support has been created expressly for bearing the text; in fact, inscriptions may also appear on artefacts made for different purposes (such as vessels) or even on non-artefacts or on natural surfaces (caves, cliffs, etc.). It is obviously not up to us to say the last word about this topic, or to define what epigraphy is and what distinguishes it, for example, from papyrology and other textual studies. Nevertheless, it is certainly very difficult to find a comprehensive ontological definition for classes of objects that have been from time to time assigned to one category or another, mainly depending on the discipline involved in their study.

2.2 The nature of an epigraph

From a logical point of view, and in accordance with the tradition of epigraphic studies, an inscription can be analysed according to three main aspects: the text-bearing object or monument (obviously involving archaeological topics), the text (and the obvious correlations with content and linguistic aspects *lato sensu*) and the feature engraved on the support in the form of letters or other symbols, which is the central element that characterizes and differentiates an epigraph from any other manifestation of written communication. Whereas in documents on papyrus or parchment (*et similia*), the materiality of the features is not relevant in comparison with the morphological and typological characteristics of the handwriting, investigated by palaeography, they are essential in inscriptions and their analysis precedes and prepares palaeographic studies. Such features represent the peculiar object of epigraphic studies, and analysis of them is as fundamental as that of the archaeological, palaeographic, linguistic

and historical aspects. Paramount importance must be ascribed to the communicative purpose of the epigraphic text, since this is exactly what distinguishes an inscription from any other feature of any kind that may be present on a given support; in other words, given the intentionality of a feature, a figurative decoration is distinguished from an inscription by both semiotic and linguistic purposes (the explicit will of communicating a message), on which a different communicative-informational structure also depends. According to a semiotic analysis, an intentional feature (i.e., one voluntarily created by man to convey a message) occurring on a given support can appear in the following forms (see [4] [5] [6]):

- Features not belonging to any writing system, i.e. the figurative decoration, even when it has value of icon or symbol (e.g., the sign of the Christian cross)

- Features belonging to 'non-linguistic' writing systems, i.e. signs of pure semasiographic systems of writing, 'language-independent', used to represent concepts and not related per se to a given linguistic structure

- Features belonging to a linguistic writing system, but not used per se: in this case we are in the presence of signs with a linguistic value but written for a purely decorative purpose or used as symbols (e.g., the A and Ω signs used as symbols of the beginning and end in the Christian tradition)

- Features belonging to a linguistic writing system and used per se, i.e. signs of a glottographic writing system, (variously) depending from a given linguistic system, thus taking on a real linguistic value

We can talk of inscriptions in the latter three cases, i.e., when we recognise signs belonging to certain writing systems; nevertheless, we can talk about written communication only in the presence of signs used *per se* as encoders of linguistic signs (and structures), and therefore, of a text.

Finally, we can summarize as follows:

- not-glottographic feature, a figurative decoration, but also a sign of pure semasiographic systems of writing

- glottographic feature, not necessary codifying a linguistic expression, since a sign can be used with different purposes than a linguistic one

3   Standards for (Digital) Epigraphy

The edition of ancient texts boasts one of the earliest and more consistent systems of standardisation in the field of Humanities: the Leiden Conventions [7].

This standard, which arose from the need to publish texts using a shared notation to describe the various observable phenomena they show, was created by an international group of scholars gathered in Leiden in 1931 and is the standard still adopted in modern epigraphy. Many of the well-established and growing database-based epigraphic *corpora*, including the Epigraphische Datenbank Heidelberg, the inscriptions section of the Deutsches Archäologisches Institut and the Epigraphic Database of Rome, also provide an extensive text field containing the text of the inscription in Leiden format, besides the typical descriptive fields used for metadata, such as find location, date, dimension and so on. The Leiden Conventions specify how features of an inscription besides the text itself should be represented in print, by using a set of standard symbols and text decorations to reproduce the state of the original document and to report the editors' interpretations.

However, with the advent of the digital era, epigraphists had to face a set of problems very similar to the ones brilliantly solved by the Leiden Conventions. An electronic format that could allow digital publishing, storage and exchange of epigraphic information in a consistent and shared format was needed. From this need arose EpiDoc, a collaborative format designed to transcode in digital format the Leiden-encoded printed editions. The initiative began in the 1990s in response to the request for a free and unrestricted set of tools supporting the creation of online epigraphic archives, which was expressed during the same period in the course of a series of conferences on epigraphy and IT. The XML format was identified as being the most suitable for this purpose and the first EpiDoc DTD was released and quickly adopted by a relatively wide community of researchers. Basically, EpiDoc is an application profile of TEI specifically adapted to the needs of epigraphy. This profile has extensively evolved from its first draft. As of today, EpiDoc provides features for the recording of the materiality and history of text-bearing objects, as well as features for scholarly editions of the text, such as commentary, illustrations, bibliography, and publication data. EpiDoc also offers facilities for the detailed description and editorial representation of the texts themselves, including transcription in the technical sense of reporting readings and representing the writing system, form, appearance, layout and editorial interventions in the text.

The EpiDoc system, despite its undoubted merits, still presents some issues, especially with respect to the inline text encoding features, arising from the fact that there are no native tools fully able to support the EpiDoc format for sessions of text editing, and thus to simplify the encoding operations. Essentially, EpiDoc-based text edition still remains a manual task, which greatly complicates the digitisation of large *corpora* of inscriptions. From a technological point of view, the choice of reproducing the paper publication format by means of XML mark-ups also raises an issue related to the style sheet necessary for the optimal rendering of the XML encoded text and therefore inseparable from it. This could represent a further portability issue for the sharing of the information from one archive to another and on the Web. EpiDoc is also unable to guarantee the typical "relational" features offered by a database, since it is lacking in

all the paraphernalia necessary to describe the complex web of relationships that characterize the various aspects of epigraphy. Only ontologies and similar semantic tools seem to be able to merge the advantages and flexibility typical of XML with the characteristic "relationality" of databases.

4 A tentative CIDOC CRM representation

4.1 Defining the concepts

In past years, some attempts have been made to use CIDOC CRM for the description of epigraphic entities. One of the first such projects was VBI-ERAT-LVPA [8], the aim of which was to use CIDOC CRM for the integration of epigraphic digital archives. This project had the merit of having given a first reply to the question of how to describe an epigraph and its various components using conceptual tools, but it did not provide definitive conclusions on the subject.

More recently, some methodological proposals have been put forward to combine EpiDoc and CIDOC CRM features and to harmonize the features they provide [9]. In particular, the EAGLE project [10], which aims to create a portal for the integration of some of the most important existing epigraphic archives, is currently engaged in the mapping between EpiDoc and CIDOC CRM, a task of great interest that will surely provide excellent suggestions for the definition of a possible extension and for the convergence of the two models. The ARIADNE project [11], although its main focus is the integration of archaeological archives, is also involved in the study of inscriptions, not only as archaeological objects but also as regards interoperability between archaeology and epigraphy. During several workshops and summer schools, the issue has been extensively discussed and outlined in its main aspects. This paper is also one of the results of these activities.

As mentioned, the purpose of this work is to lay the groundwork for a possible epigraphic extension of CIDOC CRM (CRMepi) through the conceptual analysis of the specific entities and problems with which epigraphy is concerned. We are aware that the lack of a comprehensive conceptual definition of the identity of an epigraph obviously complicates the formal definition of epigraphic entities using tools provided by ontologies and conceptual models at our disposal. Some general observations may however be made and the conceptual model of the CIDOC CRM can certainly be used as an intellectual guide to reasoning on the concepts involved in this process (see Figure 1).

4.2 The physical support

In the case of epigraphy, an essential element that emerges from the above discussion is the close relationship that tightly binds the inscription with its physical support. This close cohesion between the support and the text, as already mentioned, is, for instance, what distinguishes an epigraph from a papyrus, the study of which mainly concerns textual analysis. For some inscriptions, however, the

shape, the materials, the production techniques and all the attributes of the physical object that hosts the epigraph can become fundamental not only for their understanding but also for the definition of their nature. If we focus on the support on which the inscription was engraved, we note that CIDOC CRM offers plenty of concepts with which we could describe it. The physical support is in fact very often an archaeological object, a class of objects which has frequently been investigated in a CIDOC CRM perspective. In terms of integration and interoperability it is also important to note that, thanks to its nature, the support constitutes one of the main points of contact between epigraphy and archaeology. The specific archaeological aspects (discovery, provenance, archaeological context etc.) relating to the physical support can be documented using the CRMarchaeo extension [12].

In relation to epigraphy, it should be noted that very often the physical support has been designed and built specifically to accommodate the inscription. In this case the CIDOC CRM *E84 Information Carrier* entity seems definitely to be an optimal choice. However, this condition does not always occur: certain inscriptions may in fact even have been placed on objects not specifically designed to accommodate an inscription, as in the case of buildings, vessels or other objects of daily use on which the inscription may have been placed at a later time. In this case, the use of a more generic class, like *E22 Man-Made Object*, seems more appropriate. There are also cases in which the inscription is placed on natural surfaces not created by human activities, such as inscriptions on rocks, in caves or other similar natural places. The use of the superclass *E19 Physical Object* sufficiently broad so as to include every possible kind of physical support and would be in this case a more suitable choice. Each of these classes can still be linked with the physical features they bear, via the *P56 bears feature* property, having the *E19* class as domain and thus being inherited by all its subclasses. The EpiDoc elements used to mark archaeological information concerning physical objects or monuments (such as the *supportDesc*, *material*, *objectType* and *dimension* tags) can easily be mapped using these CIDOC CRM entities.

4.3 The inscription

CIDOC CRM provides a specific class to model the concept of inscription: *E34 Inscription*. The scope notes of this class state that "this class comprises recognisable, short texts attached to instances of *E24 Physical Man-Made Thing*". We need at this point to make sure that this class is consistent enough with the concept of inscription in the epigraphic sense, so as not to risk incurring conceptual ambiguities. Although many inscriptions bear short texts, the brevity or length of an inscription is not among its main characteristics. In fact, there are inscriptions occupying entire walls (the Gortyn Law Code or the *Res Gestae Divi Augusti*, for example) and in any case the "short text" of the *E34* class remains too vague and undefined for the purposes of our investigation. The *E34* class also belongs to the classes of conceptual objects which in turn are defined as "non-material products of our minds and other human produced data",

something that renders only in part the essence of an epigraph, not taking into any account its "materiality" which is a fundamental component of its identity. The study of epigraphy typically moves from the analysis of the physical characteristics of inscriptions before getting to their archaeological, palaeographic, linguistic and historical characteristics. In this sense, an inscription intended only as a conceptual object does not seem to capture fully the very nature of the epigraph itself. Thus, the etymology of the word "epigraph" indicates as a fundamental condition of its identity its being written on something. In all these ways it seems to present a much closer resemblance to the classes created for the description of physical features, and more specifically the *E25 Man-Made Feature*. We have managed to create some new and more appropriate classes to be used in documenting epigraphic concepts, and in particular:

– **EPI1 Epigraph**. Subclass of *E25 Man-Made Feature* intended to describe a particular feature created by humans, in various ways and on various kinds of support, mostly rigid ones, with the declared purpose of conveying a specific message towards a given recipient or group of recipients.

– **EPI2 Engraving**. Subclass of *E12 Production* indicating the activity of creating inscriptions in an epigraphic sense by using various techniques (painting, sculpture, graffiti etc.) and by means of specific tools on a given physical carrier. The definition of this activity allows us to make a better distinction between the creation of inscriptions and the production of the physical carriers that host them (two activities that are not always and not necessary contemporary), and to distinguish more accurately the creation of an epigraph from that of a story or poetry or other literary texts written, for example, on a papyrus.

– **EPI3 Epigraphic Field**. Subclass of *E25 Man-Made Feature*. This represents another important element of epigraphy, usually understood as the surface or portion of the physical carrier reserved, delimited and arranged for the purpose of accommodating an inscription, to highlight it and to isolate it from the other parts of the object or building to which it belongs. There are various types of epigraphic fields, and among the most important of these epigraphists usually distinguish those created during the same production event of the carrier and those added to it at a later time. From a conceptual point of view, the epigraphic field is a feature designed to accommodate another feature (the inscription). EpiDoc also provides specific entities for the description of these elements (the tag *layoutDesc* for example) that can be easily mapped on this class. To define the relationship occurring between *EPI3 Epigraphic Field* and *EPI1Epigraph*, the new property *EPP2 is included within* has been proposed as a sub property of *P56 bears feature*, which is in turn a shortcut of the full path relating *E19 Physical Object* through *P59 has section (is located on or within), E53 Place, P53 has former or current location (is former or current location of)* with *E26 Physical Feature* as described in the related scope notes.

4.4 The text of the inscription

As already noted, conceptually an epigraph significantly differs from the definition of the class *E34 Inscription* of CIDOC CRM. To avoid semantic and linguistic ambiguities we decided to use its superclass *E33 Linguistic Object* to describe the text (intended as a linguistic production), which the inscription records. Before proceeding any further, it would be appropriate to clarify the relationship between the epigraph, intended as a feature consisting of a set of signs, and its text as obtained through observation and decoding of those signs and the interpretation of the linguistic signs they refer to.

One of the most important operations carried out by epigraphists for their study, and especially for publication, is the so-called "reading" of the epigraph, consisting of a deep and accurate analysis and study of the surface and the signs followed by establishing as faithful as possible what is shown by the physical feature. This scientific process can be modelled by means of the concept of observation documented in the scientific extension of CIDOC CRM (CRM-sci), and more precisely by means of the *S4 Observation* class. This class seems particularly appropriate to document this kind of scientific analysis. It could also avail itself of specific instrumentation to assist reading, such as microscopes or magnifying glasses, especially in case of inscriptions of reduced dimensions. It is also possible, thanks to this class, to document the different processes of analysis and study carried out over the years on a particular object by various scholars, and to report details on the tools and methods used. The observation (*S4*) class is used to define and identify (*O16 observed values*) the graphemes (i.e., the symbolic object *E90* used to encode, on an abstract level, the linguistic units in the text) of which the engraved signs on an *EPI1 Epigraph* are the concrete graphical manifestation; this relation is made explicit by means of the *P128 carries* property. The opportunity to instantiate a more specific subclass of *E90 Symbolic Object* in order to provide a better description of these ideal graphemes requires a more thorough discussion.

The graphemes inferred from the observation of the epigraph represent the level of the intellectual decoding and understanding of the signs and constitute the basis for the subsequent operations of transcription usually carried out by epigraphists, in particular for so-called *diplomatic* transcription (i.e., a specific transcription recording only the characters as they appear on the support, without any editorial intervention or interpretation), which is also of great importance from the point of view of publication. The publishers of an epigraph, for practical reasons, generally those of typeface, perform these transcriptions using Latin or Greek characters, even in case of non-Latin and non-Greek inscriptions (Etruscan inscriptions, for example). To document this process we have created some specific classes:

– **EPI4 Transcription**. Subclass of *E7 Activity* describing the specific operation of transliteration that, starting from the symbols observed on the epigraph (*E90 Symbolic Object ->P16 was used for*), leads to the creation (*P94 created*) of a set of instances of *E73 Information Object* recording the

transcription(s) performed.

– **EPI5 Writing System**. Subclass of *E29 Design or Procedure*, which refers
to a specific sequence of characters (graphemes, *E90 Symbolic Object*) used
both to write and to transliterate the text of the epigraph (e.g. Latin letters).

As mentioned, EpiDoc provides an entire section of tools for the encoding of the
diplomatic transcriptions, implementing the Leiden Conventions in the form of
specific inline XML tags designed to mark and describe each part of the text. An
instance of the *E62 String* class, in association with one of the *E73 Information
Object* (with type = *diplomatic*), might accommodate the EpiDoc-encoded text
by means of a *P3 has note* property, to strengthen integration. The *E33 Lin-
guistic Object* class, as we already mentioned, was used instead to represent the
text of the inscription, i.e., the intended linguistic entity as resulting from an in-
tellectual and linguistic process of creation, witnessed by the epigraph and also
inferred from the various transcriptions (*E73 Information Object*). The latter
relation is established via the *EPP2 has transcription* property, a sub property
of *P130 shows features of*. The connection between *E33 Linguistic Object* and
EPI1 Epigraph is instead represented by using the *P62 depicts* property, which
perfectly renders the nature of this relationship. It should be noted that the
graphemes (*E90*) also remain in constant and close connection with the *E33
Linguistic Object*. We have used the *P67 refers to* property to describe this con-
nection. In future developments, however, we shall assess the question of whether
it is opportune to create a new sub property that better expresses this type of
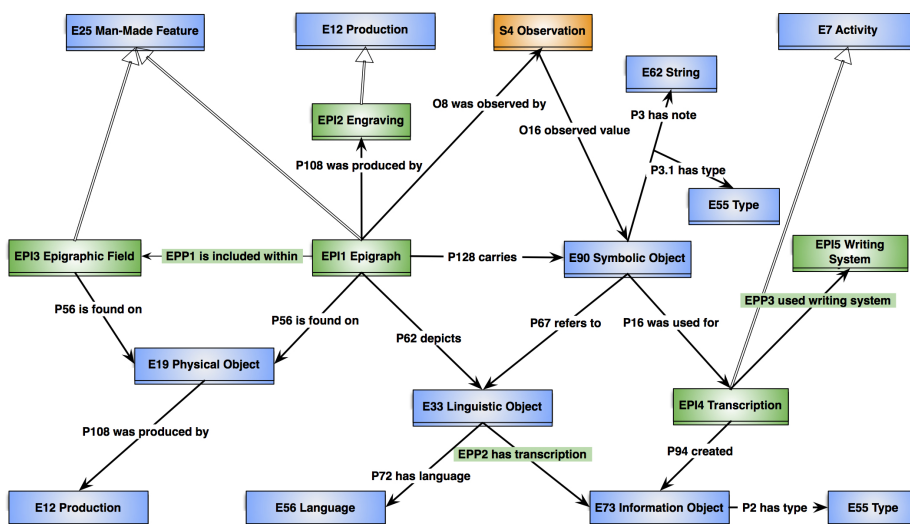link.

Fig. 1. The CIDOC CRM epigraphic model and extension

5 An example: the Oscan inscription VE 150

To test and demonstrate the potentialities of the proposed model, we chose an epigraph in the Oscan language from Pietrabbondante (Isernia, Italy), stored at the Archaeological Museum of Naples (Italy). It is a *cippus* (MANN.2525) bearing a dedicatory inscription (VE.150) [13] of an Italic sanctuary in the Oscan alphabet. Both the artefact and the inscription have been dated to around the second century BC. Many scholars have studied the inscription, but for the purposes of the example presented here we have taken into account the analysis and interpretation provided by Mommsen in 1850 [14] (see Figure 2).

The *cippus* was created specifically to host the epigraph. In this case, it seems more appropriate to use the *E84 Information Carrier* class rather than its *E22 Man-Made Object* superclass to encode it. The use of *EPI1 Epigraph* and the related production event (*EPI2 Engraving*) allows us to distinguish the event of creation of the epigraph from that of the archaeological object, although in this case the two events happened simultaneously, since the inscription was in fact sculpted (*P32 used general technique ->E55 Type = sculpture*) contextually to the production of the *cippus*. The definition of the same *E52 Time Span* for both the events indicates their contemporaneity.

The *E90 Symbolic Object* is represented by a set of characters of the Oscan alphabet evidenced through the property *P3.1 has type*, a sub property of *P3 has notes* usually used to record specific notes concerning peculiar aspects of a given entity such as the writing system used, as happens in this case. The alphabet to which the graphemes of the *E90 Symbolic Object* belong is an element of capital importance, especially in the study of non-Latin and non-Greek inscriptions, which needs to be expressed in a richer and more specific way. We will consider the definition of adequate classes and properties for modelling this concept in future revisions of the model. Through analysis of the inscription (*O4 Observation*), in 1850 Theodor Mommsen provided a reading and a transcription of it (*EPI4 Transcription ->P94 created ->E73 Information Object*) using the Latin alphabet (*EPP3 used writing system ->EPI5 Writing System*) to create both the diplomatic and interpretative transcriptions of the text. This same set of classes and properties can be instantiated several times in case of new or different readings, transcriptions and interpretations of the same epigraph by other scholars, in order to create a chain of events able to represent the history of studies of the object.

Diplomatic transcription constitutes the basis for the interpretative edition of the *E33 Linguistic Object*, i.e., the text intended as a linguistic production, encoded by means of a given writing system in the epigraph. The *E33 Linguistic Object* is therefore linked both to the *E90 Symbolic Object* (i.e. the Oscan graphemes, units of writing system on an abstract level) and to the *EPI1 Epigraph* (the concrete manifestation of such Oscan units as physical features), through the *P67 refers to* and the *P62 depicts* properties respectively. Since the *E33 Linguistic Object* is an expression of the Oscan language (*P72 has language*), it can be provided with a translation into any other language (*P73 has*

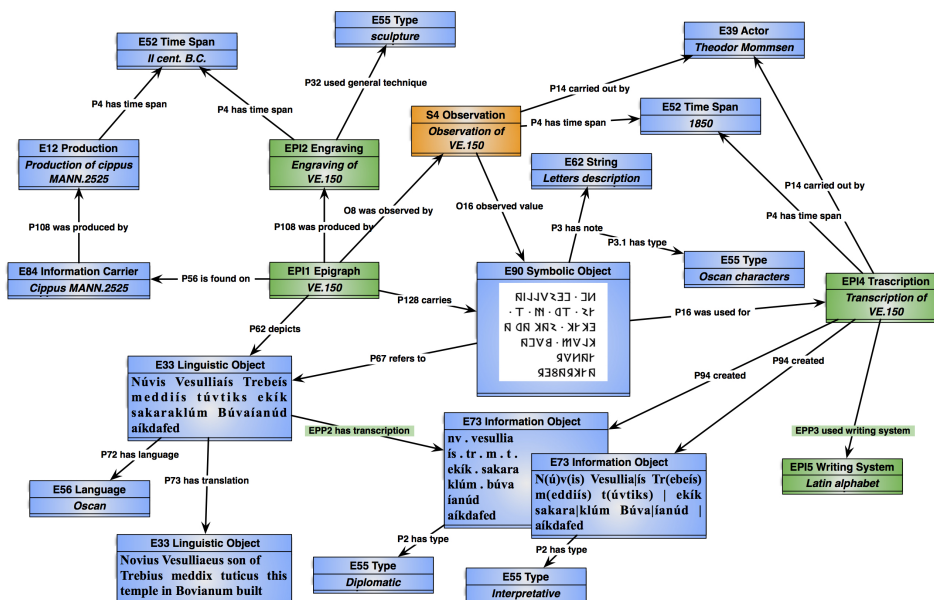*translation*), for example, into English, in order to make the content more understandable.

Fig. 2. CIDOC CRM modelling of the MANN.2525 *cippus* and the VE.150 inscription

6   Conclusions and further work

This study has just scratched the surface of what is certainly a complex problem. Although the bases for creating an epigraphic extension of CIDOC CRM (CRMepi) have been laid, much of the work still remains to be done. If we consider that EpiDoc not only provides entities for the description of the text and its structural characteristics but also provides a series of tags for the identification of actor and place names inside it, for example, we must notice that the text itself may contain semantically relevant elements that need to be captured in some way. Actor appellations can for instance relate to the commissioners of a given monument or to the people to whom a certain epigraph was dedicated; furthermore, place appellations could also refer to places where the inscription was located in the past, or to which the text refers in various ways. These appellations may also evoke and prescribe the use of thesauri and gazetteers (like Pleiades, for example) to operate a further enrichment of the descriptions or for the creation of terminology tools starting from the very texts of the inscriptions. A special case, but one that is very frequent in epigraphy, is that of the so-called "talking objects", in which the inscription becomes a "declaration" made by the

object itself ("I am the cup of Aphrodite", "I was made by Ergotimos") that in this particular case acts as an "actor", as though it were a living entity. This type of data allows us to enrich our information network about the object and to expand our archaeological and historical knowledge.

Further developments will start from these bases to plan cycles of semantic enrichment of epigraphic information from textual data. For example, it will be possible to deduce and instantiate *E39 Actor* and *E53 Place* classes from the appellations found in the text, a process already attempted for discovering relevant semantic entities in ancient literary sources [15]. The use of XLink/XPointer technologies, also based on XML as EpiDoc, would make it possible to establish cross-references between semantic entities and specific portions of the epigraphic text.

7   Acknowledgements

References

1. EpiDoc: Epigraphic Documents in TEI XML, `http://sourceforge.net/p/epidoc/wiki/Home/`
2. Guarducci, M.: Epigrafia greca. Caratteri e storia della disciplina. La scrittura greca dalle origini all'età imperiale. Istituto poligrafico e Zecca dello Stato, Roma (1967)
3. Panciera, S.: What Is an Inscription? Problems of Definition and Identity of an Historical Source. Zeitschrift für Papyrologie und Epigraphik 183, pp. 1–10 (2012)
4. Pulgram, E.: The typologies of writing systems, in Haas W. (ed.) Writing without letters, Mount Follick Series, vol. 4, Manchester University Press, pp. 1–28 (1976)
5. Pierce, C. S.: (1931-1958), Collected Papers of Charles Sanders Pierce, Burks A W (ed.), Harvard University Press, Cambridge, MA
6. Harris, R.: The Origin of Writing (1986), London, Duckworth
7. Krummrey, H., Panciera, S.: Criteri di edizione e segni diacritici, Tituli 2 (1980), pp. 205-215.
8. Doerr, M., Schaller, K., Theodoridou, M.: Integration of Complementary Archaeological Sources. In Nicolucci, F., Hermon, S. (eds.) Beyond the Artifact. Digital Interpretation of the Past. Proceedings of CAA 2004, pp. 64–69. Archaeolingua, Budapest (2010), `http://proceedings.caaconference.org/files/2004/09_Doerr_et_al_CAA_2004.pdf`
9. Lamé, M., et al.: Asking Text-Bearing Objects: Contribution of Epigraphical Theories to Digital Representation of Text, paper presented at SDH 2011, `http://crdo.up.univ-aix.fr/SLDRdata/doc/show/copenhagen/SDH-2011/submissions/sdh2011_submission_48.pdf`
10. EAGLE Project, `http://www.eagle-network.eu/`
11. ARIADNE Project, `http://www.ariadne-infrastructure.eu/`

12. CRMarchaeo, `http://www.ics.forth.gr/index_main.php?l=e&searchtype=st_p&c=711`
13. Vetter, E.: Handbuch der italischen dialekte, C. Winter, Heidelberg (1953)
14. Mommsen, Th.: Die unteritalischen Dialekte, G. Wigand, Leipzig (1850)
15. Andreussi, M., Felicetti, A.: The CIDOC CRM Encoding of the "Fontes ad Topographiam Veteris Urbis Romae Pertinentes" by Giuseppe Lugli. In: Posluschny, A., Lambers, K., Herzog, I. (eds.) Layers of Perception. Proceedings of CAA 2007. Dr. Rudolf Habelt GmbH, Bonn (2008), `http://proceedings.caaconference.org/files/2007/66_Andreussi_Felicetti_CAA2007.pdf`