Ce qui est écrit et ce qui est parlé. CRMtex for modelling textual entities on the Semantic Web

Editor(s): Name Surname, University, Country
Solicited review(s): Name Surname, University, Country
Open review(s): Name Surname, University, Country

Achille Felicetti[a] and Francesca Murano[b]
[a]*PIN, VAST-LAB, Prato, Italy*
[b]*Università degli Studi di Firenze, Italy*

Abstract. This paper presents the new developments of CRMtex, an ontological model based on CIDOC CRM created to describe ancient texts and other semiotic features appearing on inscriptions, papyri, manuscripts and other similar supports. The model is also designed to describe in a formal way the phenomena related to the production, use, conservation, study and interpretation of textual entities. CRMtex was originally meant to detect the close relationship linking ancient texts with the physical objects they are carried by, the tools and writing systems used for their production, the various scientific investigations and readings carried out on the text by modern scholars. It eventually evolved to provide researchers with the fundamental concepts for the correct and complete rendering of textual objects, the events representing their history and the cultural and social environments in and for which they were created. The full compatibility of CRMtex with the CIDOC CRM ontology and its extensions ensures persistent interoperability of data encoded by means of its entities with other semantic information produced in cultural heritage and digital humanities. The new entities presented in this paper deal more closely with textual and intertextual structures and try to deepen the close relationships existing between fragments of text or sequences of signs and the underlying meaning they were originally intended to convey.

Keywords: Ancient texts, Ontologies, Linguistics, CIDOC CRM, CRMtex

1. Introduction

In recent times, we are witnessing an intense debate that is animating the world of epigraphists and papyrologists about the need to find or eventually develop conceptual models able to express the complex entities of their domains in a semantically rich encoding, and to establish interoperability of their data with those generated, for example, in the areas of archaeological, historical and linguistic studies. The gigantic integration effort put in place by Papyri.info [1] and Trismegistos [2] and the various attempts made by projects such as EAGLE [3] to develop a semantic model in the field of epigraphy, testify to a constant and growing interest in the use of advanced and efficient conceptual tools for the generation of standardized, integrated and interoperable information in these disciplines. In the epigraphic world, another important initiative, Epigraphy.info [4] aimed at establishing a collaborative environment for digital epigraphy, is trying to raise awareness in the community of epigraphists about the importance of publishing information in a uniform format and, possibly, in a Semantic Web fashion. This initiative has the merit of having placed all the major players in the epigraphic world around a table and having directed their efforts towards the development of ecosystems in which epigraphic data coming from different sources can be easily retrieved and analyzed.

In the same perspective, many European and international initiatives are also focusing their attention on

information concerning ancient texts and on the interoperability challenges in which they are involved. The recently completed PARTHENOS project [5] has placed interdisciplinarity at the center of its activities by designing a system in which historical, archaeological and linguistic data coexist in a single digital environment. ARIADNEplus [6], an initiative recently started as a continuation of the first and successful ARIADNE project, is also attempting an integration between archaeological information and data from other disciplines, with particular regard to the study of archaeological artefacts bearing inscriptions, such as amphorae, coins and other similar objects, and with the clear intent of creating an interoperable archive based on FAIR principles and international standards. ARIADNEplus is also looking for an ontology or application profile capable of consistently relating textual and archaeological data. This is one of the gaps that our work aims to fill.

In two of our previous works [7][8] we tried to give an account of what had been done in the field of epigraphy and what tools had been used to describe in semantic format textual entities created in antiquity. In those same works we also laid the foundations for the definition of a semantic model (CRMepi, later expanded to become CRMtex) centered on the semantic definition of the ancient text and the description of its multifaceted aspects. In the present paper, after a quick review of similar recent initiatives, we present the latest developments of the CRMtex model and the conceptual considerations that underlie its evolution.

2. Looking for new semantic tools

2.1. Ontologies and application profiles: a work in progress

Despite the great interest of many scientific communities for the tools proposed by the world of the Semantic Web, it is interesting to note that some wide-ranging initiatives such as EPIDAT [9], whose purpose is to publish epigraphic data in LOD format, lament the absence of an ontology able to confer semantic value to epigraphic information. On the other hand, though, an increasing number of activities conducted by groups interested in the subject of ontologies for ancient texts has flourished in recent years.

The ontological approach is also pursued by some major players in the field of epigraphy with alternate success: the Epigraphic Database Heidelberg, for instance, has released a very basic ontology for the encoding of his vast digital *repertoire* in Linked Open Data format [10]; however, its model still seems less suitable to be conceived as a tool for a deep integration.

A very interesting initiative, the Economics and Political Network project (EPNet) [11], is building an ontological model based on CIDOC CRM to deal with the events and objects connected with the distribution of food in the Roman world. The EPNet ontology looks very promising and has already been investigated by the ARIADNEplus project as a candidate to become part of the application profile for epigraphic data. With the same intent, The Epigraphic Ontology Working Group (EpOnt) [12] is trying to establish an application profile based on concordance of ontologies, for recording epigraphic editions. The initiative is extremely interesting and we believe will produce excellent results very soon.

It should be noted that all these initiatives aim at developing very specific tools for solving the problems of the disciplines in which they are born. None of them aims to give a common conceptual basis or to look for points of contact, which also exist between these various disciplines keeping in mind the objective of interoperability. The notions of “text” as an object and of "writing" as an event, for example, allow, once defined, to similarly model inscriptions on vases or coins, graffiti or texts on papyrus using an identical format.

2.2. EpiDoc: a de facto standard for ancient texts

It deserves to be emphasized that epigraphists and papyrologists have long since elected TEI EpiDoc [13] as their own metadata standard, as this tool is extremely versatile for representing texts and the phenomena that typically characterize them, with a particular attention to the needs of a rich and well-rendered visualization.

EpiDoc provides a series of tags for detecting specific elements, since the text itself may contain semantically relevant information that needs to be captured in some way. Interesting examples in this sense are the tags that identify temporal entities, actors and place names, which give EpiDoc the ability to bind external semantic elements starting from identifiable textual fragments.

Nevertheless, it should be also noted that EpiDoc does not offer the typical descriptive tools used by ontologies to capture the conceptual *nuances* of the text as a material phenomenon framed in time and

space, and to define metadata that can describe its structure, history and the events and people who determined their existence and life. In this context, it becomes essential to use models that can put the TEI ecosystem in touch with the universe of the ontologies and to act as a link between these different worlds. In the field of numismatics, for instance, the Nomisma.org project [14] has successfully attempted to act as a link between different numismatic resources by integrating specific vocabularies, models and ontologies. CRMtex tries to propose a similar solution by establishing a solid conceptual basis for bridging knowledge of different types and implementing interoperability for textual data in an effective way.

3. CRMtex: an ontology for ancient texts

The need of creating a new ontology for ancient texts started from the assumption that, unlike printed texts, non-mechanised written texts (including inscriptions, papyri and manuscripts) have typical peculiarities to be taken into account for their study.

We based our model on the solid foundations of CIDOC CRM [15] because it constitutes one of the most used ontologies in the field of Cultural Heritage. CIDOC CRM provides, already in its core version, most of the entities necessary to model common elements such as actors, objects, places, events and their mutual interrelations on a chronological basis.

The core concept of our model is the notion of "text" as the product of a semiotic process, involving an encoding ("writing") and a decoding ("reading") process. Writing is in turn a particularly sophisticated human technology allowing the encoding of a linguistic message through a series of signs specifically selected for this purpose.

Investigating in detail the close relationship that links the text with the writing event, some considerations to clarify its nature deserve to be exposed.

Although every speech can be transposed into an equivalent written message, and *vice versa*, speech has a priority over writing, at least in four respects: phylogenetic, ontogenetic, functional and structural [16]. In fact, all languages are spoken but not necessarily written; every human being learns to speak naturally spontaneously, the ability to write coming only later and through specific training; the spoken language is used in a wider and differentiated range of uses and functions; writing originated as a representation of speech. According to Ferdinand de Saussure [17], in fact, «a language and its written form constitute two separate systems of signs. The sole reason for the existence of the latter is to represent the former».

In this semiotic perspective, it is worth considering that even in writing, as in the analysis of the linguistic system, it is necessary to distinguish the concrete level of the personal execution (i.e. the real act of tracing signs on a surface) from the abstract level which all the single occurrences must be took back to, on the basis of a sameness principle (e.g. the identification of an "A", independently from the peculiar shape somebody gives to it).

Thus, a "text" is constituted by a number of signs physically traced (i.e. *written*) on a specific support and intended to encode a linguistic expression.

Because of their non-mechanised origin, ancient texts are unique and unrepeatable entities; in addition, they form, along with their support, an inextricably linked unique object of study. From a conceptual point of view, therefore, whether it is painted, written in ink or engraved, a text preserves its physical nature, being it a feature deriving its existence from its strict dependence on the support on which it is located.

Following this approach, CRMtex provides the *TX1 Written Text* class to describe the physical signs composing a text, engraved or incorporated on or into some kind of physical support, having semiotic significance and the intentional purpose of conveying a linguistic message.

Simultaneously, the text production event is represented by the *TX2 Writing* class, indicating the activity of creating permanent marks on a physical support using various techniques (painting, sculpture, etc.) and by means of specific tools. The *TXP5 was written by* property (subproperty of *P108 was produced by*) renders in a more clear way the link between the text (*TX1*) and its production (*TX2*).

The *TX1* and *TX2* classes**,** which represent the backbone of our conceptualization work on the text production side, are flanked by another class, that better specifies how the text production process took place: the *TX3 Writing System* class, representing the conventional set of signs and the related rules used to codify and represent (i.e. to *write*) utterances meant to be recovered at a distance of time and/or space by those who have the knowledge of the same code (i.e. the same linguistic system). The *TXP9 is encoded by* property provides a direct link between the text (*TX1*) and the writing system (*TX3)*, thus offering the pos-

sibility to describe this relation in more generic terms (see Fig. 1).

In addition to dealing with text as an object, our model also focuses on the aspects of the research and provides classes and relationships to describe the typical operations that scholars from different disciplines put in place to gain knowledge about textual entities. It is evident, in this perspective, that the study of ancient texts typically starts from the analysis of the physical characteristics of the text itself before moving to the investigation of their archaeological, paleographic, linguistic and historical features. In this regard, the *TX5 Reading* (subclass of the *S4 Observation* CRMsci class) and *TX6 Transcription* (subclass of the *E7 Activity*) classes, with the related properties, are provided (and extensively described in [8]).

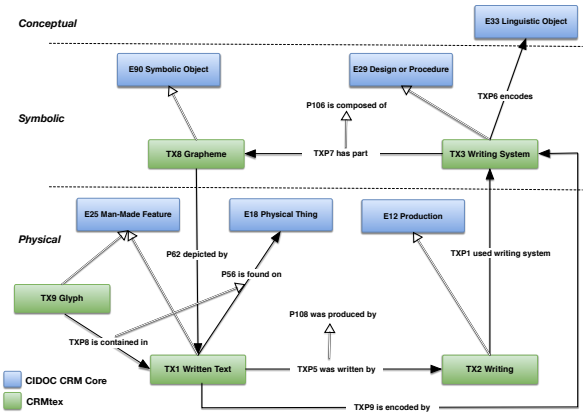

Fig. 1. General overview of the new CRMtex model.

4. Between Semiotics and Linguistics: new entities in CRMtex

4.1. Written Text Segments

In designing the new entities of our model, we began by thoroughly investigating the interconnections existing between the text and its various components. We have also tried to establish a complete chain of connections to link these components and the whole text with the linguistic level they encode. Some elements have proved to be absolutely essential for this purpose. On the side of the reading process (i.e. the decoding of the text), and therefore of the investigation of the text by the scholars, one in particular has strongly claimed the scene: the text segment element.

Thus, we have introduced the new *TX7 Written Text Segment* class, a subclass of *TX1 Written Text* intended to identify portions of text considered to be of particular significance as witnesses of a certain meaning, or bearers of special phenomena relevant to the investigation, study and understanding of the text (see Fig. 2). Examples of text portions are: text columns, text fragments, sections, paragraphs, single words or letters, or other specific components of the written text.

Scholars of different disciplines need to identify such segments, based on the requirements of their study, and to focus their attention on them in order to describe their physical conditions (form, layout, etc.), to verify their legibility or to identify particular phenomena (e.g. linguistic or paleographic) connected to them. It is important, at modeling time, to unambiguously define such segments and their relationship with the text in its entirety, so as to be able to assign specific issues to the individual segments, independently of the text as a whole. Particular production (*TX2*) or destruction (*E6*) events, can be associated to each fragment as in the case of letters or words damaged or worn by atmospheric agents or human interventions. Specifications about conditions (*E3*) for documenting its status during the observation process (*S4*) can be easily stated as well. This allows scholars to document different events for the investigated segments in a more precise and punctual way and to assign observations and interpretations to them.

The relationship between a written text (*TX1*) and its components is documented through the *TXP4 has segment* property.

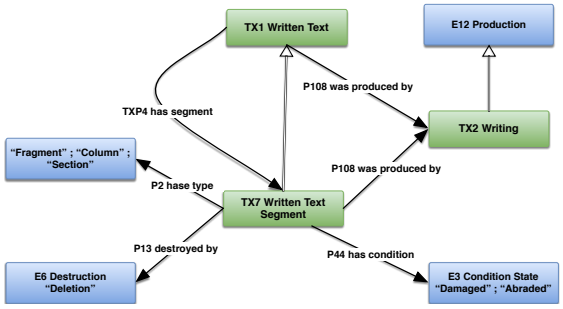

Fig. 2. Written Text and Written Text Segments in CRMtex.

4.2. Glyphs and Graphemes

The physical signs composing a *TX1 Written Text* constitute the material manifestations (glyphs) of writing system units, i.e. the graphemes, the minimal

functional distinctive units of writing. Ernst Pulgram stated that «in reducing a language to writing, that is, in making visible marks that evoke or recall linguistic performance, it would seem that each mark must represent a syntagmeme or a lexeme or a morpheme or a phoneme or whatever other kind of unit the inventor of the system may chose as his basis» [18]. For instance: in a Latin inscription, single alphabet letters (glyphs) represent graphemes, a grapheme corresponding to a letter only in alphabetic system of writing. In Mycenaean Linear B inscriptions and in Old Persian cuneiform inscriptions glyphs represent syllabograms (graphemes representing a syllable, not a single sound); in an Egyptian hieroglyphic text, glyphs represent syllabic, alphabetic and also ideographic elements, i.e. elements standing for lexical/semantic units.

Phonographic writing systems [19][20] represent phonological units of one size or another, but the 1:1 correspondence between sound (phoneme, syllables, *etc*.) and sign (grapheme) is lost in diachrony, obscured by spelling conventions and phonetic changes to which linguistic systems are subjected in history. Think of the spelling discrepancies in English between writing and reading: for example, the <i> grapheme stands for various phonemes: /ɪ/ (as in *him*), /ʌɪ/ (as in *time*), /i/ (as in *police*), /a/ (as in *timbre*); *vice versa*, the /f/ phoneme can be represented by <f> (as in *film*), <ph> (as in *philology*), <gh> (as in *enough*).

On the side of the message retrieval, since each grapheme is bound to a given linguistic unit of specific languages, reading the written message presupposes the ability to read the language of the writer.

In this view, the model provides two new classes to represent the units the scholars deal with: *TX8 Glyph*, subclass of *E25 Man-made Feature* and *TX9 Grapheme,* subclass of *E90 Symbolic Object* (see the general schema in Fig. 1). Specific properties are used to settle the strict correspondence between graphemes and glyphs and their typical parthood relationships like, for instance, the *TXP7 has part*, used to state the (conceptual) belonging of a grapheme to a given writing system, and the *TXP8 is contained in* to state the (physical) belonging of a glyph to a given text or segment of text. The *TXP9 is encoded by* is used to state in a more general way that the graphemes used to compose the text (*TX1*) belong to a given writing system (*TX3*).

Moving to the level of the linguistic sounds, it will be the decoders (readers, including scholars), who from time to time, on the basis of the knowledge of the linguistic system, will attribute to each sign or group of signs the adequate phonetic value, also on the basis of spelling conventions in place in a given graphic system at a given historical moment, since the orthographic rules can change over time, even if less quickly than the linguistic system does. The ontological description of the link between linguistic and graphic units is under preparation by the authors.

5. Application scenarios

5.1. CRMtex and EpiDoc

In designing our model, we have always tried to maintain the compatibility of our entities with those of EpiDoc. We have also made CRMtex classes and properties particularly suitable to describe also peculiar phenomena of the text and its conditions for which EpiDoc tags are usually used. For instance, in presence of characters erased in antiquity but still legible in a more or less clear way, EpiDoc employs the following syntax:

```
<del rend="erasure">
 <orig>abc</orig>
</del>
```

The same information can be expressed in CRMtex by combining the *P43 has dimension* and the *P44 has condition* properties the following way:

```
<http://crm.tx/text102/fragment5>
   a   crmtex:TX7_Written_Text_Segment ;
   crm:P44_has_condition   "erasure" ;
   crm:P43_has_dimension   <frg5_dim> ;

<frg5_dim>
   a   crm:E54_Dimension ;
   crm:P90_has_value   "3" ;
   crm:P91_has_unit    "character"
   crm:P3_has_note     "abc"
```

More details can be specified, if necessary, for each of the 3 mentioned characters by instantiating a *TX8 Glyph* class for each of them in order to describe, for example, the specific circumstances under which they were damaged.

On the other hand, an erasure indicating a text lost and totally illegible, usually encoded in TEI EpiDoc (XML) as:

```
<del rend="erasure">
 <gap reason="lost" quantity="4"
  unit="character"/>
```

```
</del>
```

implies, according to CRMtex, the use of an *E6 Destruction* class, indicating an event that took out of existence (and out of the support) the signs of the original text making them unrecoverable. CRMtex describes such a phenomenon the following way:

```
<http://crm.tx/text102/fragment13>
   a   crmtex:TX7_Written_Text_Segment ;
   crm: P13i_destroyed_by   <frg13_dest> ;
   crm:P44_has_condition   "erasure" ;
   crm:P43_has_dimension   <frg13_dim> .

<frg13_dest>
   a   crm:E6_Destruction .

<frg13_dim>
   a   crm:E54_Dimension ;
   crm:P90_has_value   "4" ;
   crm:P91_has_unit   "character" .
```

RDF notation is certainly less concise than that provided by EpiDoc, but it is semantically richer in details and open for future enrichment. The use of the *E54 Dimension* class, for example, offers the opportunity to specify the events, times and circumstances in which the text was lost if such information comes to light during the research work.

Encodings of this type allow us to establish points of contact between the CRMtex and EpiDoc, thus extending the application scenarios for future integration and interoperability imagined at CRMtex design time.

5.2. The inscription on the Arch of Constantine

To illustrate the features of the new version of the CRMtex, we propose an example from the epigraphic world: the inscriptions on the Arch of Constantine, one of the most famous ancient monuments in Rome. The monument, still located in its original position between the Colosseum and the Roman Forum, is a triumphal marble arch (the largest monument of this kind in Roman era) dedicated in 315/316 A.D. by the Roman Senate to the emperor Constantine after his victory over Maxentius in the Battle of the Milvian Bridge in 312 A.D. Among the other decorations (including statues, panels, reliefs and similar decorative material), the arch carries, on its attic, two identical inscriptions [21], originally inlaid with gilded bronze letters, explaining the reason of its construction.

As of today, the bronze letters are lost and only the large cuttings in the marble, in which the bronze letters sat, remain. The text is repeated, identically, on the South and North faces of the arch. A transcription and a translation in English of the same inscription is presented below.

- *Inscription Transcription*

IMP(ERATORI) · CAES(ARI) · FL(AVIO) · CONSTANTINO · MAXIMO · P(IO) · F(ELICI) · AVGUSTO · S(ENATUS) · P(OPULUS) · Q(UE) · R(OMANUS) · QVOD · INSTINCTV · DIVINITATIS · MENTIS · MAGNITVDINE · CVM · EXERCITV · SVO · TAM · DE · TYRANNO · QVAM · DE · OMNI · EIVS · FACTIONE · VNO · TEMPORE · IVSTIS · REMPVBLICAM · VLTVS · EST · ARMIS · ARCVM · TRIVMPHIS · INSIGNEM · DICAVIT

- *Inscription Translation*

"To the Emperor Caesar Flavius Constantine, the Greatest, Pius, Felix, Augustus: inspired by (a) divinity, in the greatness of his mind, he used his army to save the state by the just force of arms from a tyrant on the one hand and every kind of factionalism on the other; therefore, the Senate and the People of Rome have dedicated this exceptional arch to his triumphs"

From the CIDOC CRM point of view, the Arch is an archaeological object (i.e., an *E22 Man-made Object*) made of marble, mainly intended to commemorate the emperor. Two distinct writing events (*TX2*) can be assigned to the inscriptions, to describe the different production phases of each of them and to distinguish them from the production of the monument.

CRMtex can be used to describe the inscriptions appearing on the arch and relate them to the monument via the *P56 bears feature (is found on)* property. Each of the two inscriptions can be rendered as a *TX1 Written Text*, being the physical features intended to carry a particular significance. A *TX2 Writing event* can be specified for each *TX1* via the *TXP5 was written by* property to render the production of the cuttings made to host the bronze letters.

A *TX4 Writing Field* class can be used to describe the portion of the surface of the arch reserved by the builders and appositely arranged for accommodating

the inscription in order to highlight it from the other parts of the object and to enhance its readability. Thus, the CRMtex encoding in this case will include two *TX4*s instances.

The linguistic message conveyed by the inscriptions (*E90 Symbolic Object*) is encoded by means of the writing system this language uses. It follows that the *TX1 Written Text* class is the concrete graphical manifestation (i.e. the signs – in this case the Latin letters – we can read on the stone) of the conceptual level of encoding a linguistic expression through the semiotic activity of writing (*TX2 Writing*) by means of a *TX3 Writing System* (in this case, Latin alphabet) and of the graphemes (*TX8*) composing it.

Over the centuries, the arch of Constantine has been investigated thousands of times by scholars from all over the world and also reproduced by famous illustrators such as Giovan Battista Piranesi. Also, the inscriptions have been studied and transcribed several times in order to understand its nature, clarify the meaning of each section and improve its historical comprehension so as to put it in direct relation with the events that determined its creation.

For this type of activity, aimed at studying and processing the inscribed text, CRMtex provides specific classes and properties. The transcription of the text(s) present in *Corpus Inscriptionum Latinarum*, for instance, can be represented via the *TX6 Transcription* class while the analysis of the same inscription(s), carried out by Rodolfo Lanciani in 1892 [22], can be documented using the *TX5 Reading* class, underlying the scientific nature of the investigation. Reading (*TX5*) and transcription (*TX6*) activities can be related via the *P20 has specific purpose* property, inherited by CIDOC CRM core.

The *TX7 Written Text Segment* class can be used to highlight portions of text on which the study focuses, on which peculiar phenomena appear or from which special meanings are derived. Rodolfo Lanciani, for instance, investigated the "INSTINCTV DIVINITATIS" phrase, making hypothesis on its real meaning in the framework of the message Constantine intended to transmit to inhabitants of Roman Empire (both Christians and Pagans). Figure 3 only shows a CRMtex general rendering of one of the inscriptions on the Arch of Constantine: more detailed descriptions of the text and the way it was investigated, thus expanding the semantic knowledge graph concerning this monument, can be defined when required.

The *TX8 Glyph* class in combination with the *TXP3 is rendered by* property, for instance, can be used to model one of the typical phenomena of Roman epigraphy, i.e. the use of specific signs as abbreviations, also present in this text (e.g. "S" for "SENATUS", "P" for "POPULUS" etc.). Associating abbreviation expansions to these glyphs would be ideal to document the choices made by scholars for rendering abbreviations during the transcription phases. The considerations that motivated these interpretative choices can be expressed by means of CRMinf [23], the extension of the CIDOC CRM developed to support argumentations and to document inferences and hypotheses formulation.

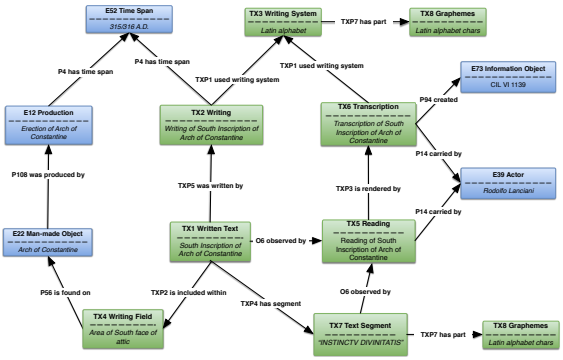

Fig. 3. CRMtex modelling of the inscription on the South side of the Arch of Constantine.

6. Conclusions

CRMtex was developed by adopting the best modeling principles of the ontological world and the fundamental paradigms of the linguistic research: this makes it a tool capable of conferring ontological value to textual entities, offering innumerable benefits for research in many humanistic disciplines. The possibility to provide representation of cultural data on the Semantic Web, to publish them in standard formats (such as LOD) and to make them easily available, interoperable and reusable in an infinite number of contexts, certainly represents one of the most relevant features of the model.

CRMtex native ability to describe relationships between text and artefacts by efficiently placing the text in the context of the life and history of ancient objects, also makes it ideal to be employed in projects like ARIADNEplus or in initiatives like Epigraphy.org. The perfect compatibility with EPNet, the model used by some ARIADNEplus partners to codi-

fy epigraphic information, will foster the possibility for CRMtex to become part of the Application Profile for epigraphy under definition within this project.

Nevertheless, a lot of work still remains to be done for the ontology to reach its maturity.

In 2018 CRMtex was accepted as part of the CIDOC CRM family [24] [25], thus becoming a new tile of the CIDOC CRM mosaic of models. A fine-tuning work to make CRMtex perfectly integrated and consistent with the other extensions of this ecosystem is already in place. In particular, we will need to plan harmonization with CRMinf, of which we have already stressed the importance for the interpretation of the text (see Chapter 5) and with FRBRoo, a CIDOC CRM compatible model aimed at representing the semantics of bibliographic information. Many FRBRoo classes (such as the *F2 Expression*, *F12 Nomen* and *F23 Expression Fragment*) actually present interesting points of contact with CRMtex and could form the basis for the creation of a more complex (but more complete) ontological instrument for the effective modelling of (ancient and modern) textual entities.

Among future activities, we aim to investigate the close correlation of graphemes with the linguistic units (such as phonemes) of which they are conceptual representations and the way in which, through phonemes, the thought of the speaker (and therefore of the writer) materializes in the form of linguistic expressions to become text. We will then extend CRMtex with the new entities suitable to describe such complex linguistic phenomena.

References

[1] Papyri.info Initiative, http://papyri.info.
[2] Trismegistos Portal, http://www.trismegistos.org.
[3] EAGLE Project, https://www.eagle-network.eu.
[4] Epigraphy.info Initiative, http://epigraphy.info.
[5] PARTHENOS Project, http://www.parthenos-project.eu.
[6] ARIADNEplus Infrastructure: https://ariadne-infrastructure.eu.
[7] A. Felicetti, F. Murano, P. Ronzino, F. Niccolucci, CIDOC CRM and Epigraphy: A Hermeneutic Challenge, in: Extending, Mapping and Focusing the CIDOC CRM. Proceedings of CRMEX 2015, P. Ronzino, ed., CEUR-WS.org, 2016, 55–68 (http://CEUR-WS.org/Vol-1656).
[8] A. Felicetti, F. Murano, *Scripta Manent*: A CIDOC CRM Semiotic Reading of Ancient Texts, International Journal on Digital Libraries 18/4 (2017), 263–270.
[9] T. Kollatz, EPIDAT - Research Platform for Jewish Epigraphy, in Crossing Experiences in Digital Epigraphy. From Practice to Discipline, A. De Santis, I. Rossi, eds., Warsaw – Berlin, De Gruyter, 2018, 231–239.
[10] Epigraphic Database Heidelberg Ontology, https://edh-www.adw.uni-heidelberg.de/edh/ontology.
[11] D. Calvanese, P. Liuzzo, A. Mosca, J. Remesal, M. Rezk, G. Rull, Ontology-Based Data Integration in EPNet: Production and Distribution of Food During the Roman Empire, Engineering Applications of Artificial Intelligence 51 (2016), 212–229.
[12] EpOnt (Epigraphic Ontology Working Group): https://groups.google.com/forum/#!forum/epont.
[13] EpiDoc: Epigraphic Documents in TEI XML, http://sourceforge.net/p/epidoc/wiki/Home/.
[14] Nomisma.org Collaborative Project, http://nomisma.org.
[15] CIDOC CRM v.6.2.3, http://www.cidoc-crm.org/Version/version-6.2.3.
[16] J. Lyons, Human Language, in: Non-Verbal Communication, R. A. Hinde, ed., Cambridge University Press, New York, 1972, 49– 85.
[17] F. de Saussure, Course in General Linguistics, Lausanne – Paris, Payot, 1916, translated and annotated by Roy Harris, London, Duckworth, 1983.
[18] E. Pulgram, The Typologies of Writing-Systems, in: Writing Without Letters, W. Haas, ed., Manchester, Manchester University Press, 1976, 1–28.
[19] G. Sampson, Writing Systems: A Linguistic Introduction, London, Hutchinson, 1985 [2nd ed. Sheffield, Equinox, 2015).
[20] H. Rogers H., Writing Systems: A Linguistic Approach, Oxford, Blackwell, 2005.
[21] CIL VI 1139.
[22] R. Lanciani, Pagan and Christian Rome, Boston – New York, Houghton, Mifflin and Company, 1892.
[23] CRMinf Argumentation Model, http://www.cidoc-crm.org/crminf.
[24] CRMtex Model for the study of ancient texts, http://www.cidoc-crm.org/crmtex/.
[25] CIDOC CRM Compatible Models and Collaborations, http://www.cidoc-crm.org/collaborations.