

Revealing cross-disciplinary information through formal knowledge representation – a proposed Metadata for ancient Cypriot inscriptions

V. Vassallo, E. Christophorou, S. Hermon, F. Niccolucci
The Cyprus Institute – Science and Technology in Archaeology Research Center
CyI - STARC
Nicosia, Cyprus

Abstract—This paper presents recent developments in constructing a cross-disciplinary metadata for ancient Cypriot inscriptions, integrating information regarding the objects themselves and their digital “surrogates” (3D models, photographic documentation, digital texts, transliterations, etc.).

Keywords—ancient Cypriot inscriptions; digital epigraphy; Digital Humanities; cross-disciplinary metadata

I. INTRODUCTION: INSCRIPTIONS AND HISTORY OF RESEARCH

Ancient inscriptions are one of the most direct evidences of past activities, ideas, or events as narrated for posterity by past people. Understanding the intrinsic information embedded in ancient inscriptions may serve as a first source of gathering information about past societies, individuals, historic events or religious / mystical beliefs of cultures in the past. They are also a valuable link between history and archaeology, archaeology and social sciences, history and geology, gender studies and sociology, philosophy of religion or history of sciences. Ancient inscriptions may be investigated by many different, but ultimately complementary disciplines: archaeology (investigating, among others, the context of the finds themselves), geology (providing details on the material upon which inscriptions were carved), philology (analyzing for example the text, the writing style, the scripture, etc.) or studies revealing aspects of their carving, preservation, conservation, etc. Additional information may regard methods of their investigation: under microscope, 3D documentation, photographic campaign or musealization.

Recent developments in Internet technologies enhanced the access to information and triggered a tremendous effort in creating online accessible repositories of information regarding, among others, ancient inscriptions. In the past few years, several projects on digital epigraphy made inscriptions available to the wider community and have greatly contributed to sharing, understanding and studying ancient inscriptions. Scholars promoted initiatives and conventions to electronically transfer ancient texts (such as Greek and Latin inscriptions) and faced the peculiarities of presenting and editing inscriptions in digital repositories. The Text Encoding Initiative (TEI) for

Epigraphy (EpiDoc) has been a major starting point towards the standardization of electronic publication of Greek and Latin inscriptions, which has been spread and applied by various digital epigraphy projects [1]; [2]. At the same time, researchers in digital epigraphy developed various tools to advance the study and research of inscriptions online [3]. Consequently, such effort highlighted the importance of having explicit descriptions of terminologies, shared vocabulary and agreed upon thesauri, the ultimate goal being the development of domain ontologies, the basis of shared, distributed repositories of knowledge on Ancient Inscriptions [4]. These will help in their turn in creating cross-disciplinary digital libraries for ancient inscriptions, which, together with digital tools for interaction with such content, will compose the research infrastructures of tomorrow.

This contribution presents an ongoing research on the development of a cross-disciplinary metadata for Ancient Cypriot inscriptions, integrating information regarding the objects themselves and their digital “surrogates” (3D models, photographic documentation, digital texts, transliterations, etc.), within the frame of EAGLE, “Europeana network of Ancient Greek and Latin Epigraphy”, an EU funded project under the umbrella of the CIP-Best Practice Network.

EAGLE brings together some of the most prominent European institutions and cultural archives in the field of Classical Latin and Greek epigraphy, to provide Europeana with a comprehensive collection of unique historical and archaeological sources which constitute an important aspect of European culture. The Science and Technology in Archaeology Research Center (STARC) of the Cyprus Institute (CyI) is among the content providers that will contribute Ancient Cypriot literary inscriptions which will derive from the Archaia Kypriaki Grammateia Digital Corpus (AKGDC) project, funded by the A.G. Leventis Foundation [5]. The project contains a corpus of Ancient texts that includes a wide range of literary genres such as epic, lyric and dramatic poetry, epigrams inscribed on stone, prose, medical and philosophical texts and covers the ancient Cypriot literary production in a time span of c. thirteen centuries (from the 7th century BC to the 6th century AD). Consequently, the research presented here, on the cross-disciplinary metadata schema for Ancient Cypriot

inscriptions, is based on the set of inscriptions that will be published online within the AKGDC project.

II. STATE OF THE ART

The definition of the metadata schema for Ancient Cypriot inscriptions derives from a first assessment and comparisons of what is available in the field of digital epigraphy and in the community of Digital Humanities [6]. The formal descriptions of inscriptions used in various digital epigraphy repositories has been thoroughly examined and analysed¹: first result of this survey highlighted that only few digital archives use standardized domain related metadata schema (e.g. TEI EpiDoc). Others describe their data using, simplified and generic cross-domain standardized metadata (e.g. Dublin Core). Very few use CIDOC-CRM (or any other reference model), while the rest publish their data online through proprietary schemas that are only constituted of limited fields.

Following the analysis of the formal description of the metadata schema or structures used by different institutions, we carried out comparisons of the different schemas, trying to make a first mapping of the metadata fields and identify the common ones within the various digital inscription initiatives. As a general rule, it has been noted that in all examined cases the description of the corresponding digital surrogates (of the original inscriptions) is missing. Therefore, the metadata containing information related to the epigraphic inscription solely, is characterized as “a written monument” [8].

III. CROSS-DISCIPLINARY METADATA FOR CYPRIOT INSCRIPTIONS

The assessment of available metadata schema has brought to consider the elements used in other collections: in particular EpiDoc, being the schema mostly used by the epigraphist community, presents a core that is largely used for the part relative to the inscription itself. Among the digital epigraphy projects evaluated, most are based on EpiDoc specifications. An EpiDoc file is a representation of the edition of one inscription or a group of inscriptions in XML. At a minimum level of description, the file may contain a text in Greek or Latin just with editorial siglae. It may also contain apparatus criticus, translation, commentary, place of finding, description and dating of the text or object, and other information elements

¹ Over the last few years, many epigraphic corpora have been digitized and different projects are available online. Starting from the first pilot project ‘Inscriptions of Aphrodisias’ in 2002 we arrive to the federative experience of Eagle “Electronic Archive of Greek and Latin Epigraphy”, that in 2003 started to collect existing projects dedicated to the electronic archiving of the ancient epigraphy documentation and in 2005 and 2009 received the consistent contribution of 4 important projects already online: Epigraphische Datenbank Heidelberg (EDH), Epigraphic Database Roma (EDR), Epigraphic Database Bari (EDB) and Hispania Epigraphica (HEp) [7]. Currently, the EAGLE “European network of Ancient Greek and Latin Epigraphy” project converged with the previous consortium and was expanded with more digital projects; Some of the projects under the EAGLE umbrella are: the Last Statues of Antiquity of the University of Oxford, Ausonius of the University of Bordeaux, Ubi erat lupa of the University of Salzburg, etc. For the assessment we evaluated all these digital projects and the metadata schemas that they employ.

that are usually published in scholarly editions² [9].

The metadata schema for ancient Cypriot inscriptions presented below and currently under development has various goals and aims to answer several challenges: it describes in detail the “digital surrogate” and its provenance, it provides related information about the context (ancient and modern) of the inscription and mostly it is organized in such a way that a harvesting to larger initiatives, for instance Europeana, or future research infrastructures, will be enabled. The metadata is based on previous research in STARC [12], taking into consideration LIDO³ and CARARE⁴ metadata schemas and having at its base CIDOC-CRM as a reference model.

Our proposed metadata schema is organized in five groups, which correspond to five different research domains (see below). The model is organized in wrappers (and sub-wrappers) that refer to all possible information that an inscription may contain. The schema has been conceptualized as a multidisciplinary tool and includes the necessity to engage different disciplines and the appropriate representation of them.

This new metadata schema focuses on different elements that compose the knowledge about the ‘inscription’. Ancient inscriptions can, in fact, be investigated and studied as a multidisciplinary subject where complementary disciplines converge: archaeology, philology, chemistry, geology, conservation, visualization and museology are some of the areas to be taken into consideration for a fuller description of these artefacts. For this reason, the need for a more structured and detailed way to present metadata regarding Ancient Greek inscriptions of Cyprus arose. This effort will help experienced and non-experienced users to gain a better understanding not only about the inscription itself but to place it in its wider context.

For this reason, another feature to visualize this cross-disciplinary aspect is foreseen: sub-wrappers will be identified by flags that will help researchers to edit or find the sections that they are interested in. For example, ‘Archaeo’ is the flag that corresponds to the sub-wrapper that contains metadata

² EpiDoc is the schema which is most commonly used in the epigraphy community but there are different and discordant opinion on its use. In the last years it has been underlined that EpiDoc is not providing a way to encode computational semantics, a feature that provides the basis for sharing epigraphical information in the perspective of the new researches in open linked data [10]. Other scholars instead heavily promote the use of EpiDoc [11], since they underline that it serves not just as a data interchange format but also supports the creation of sophisticated digital editions and corpora of inscriptions that are “fully queryable and manipulable”(as cited by [11]). They also suggest that EpiDoc could be potentially used as a way of storing, preserving and distributing epigraphic data in a digital format [2]. Currently, within the European project EAGLE “European network of Ancient Greek and Latin Epigraphy”, a common metadata schema based on EpiDoc core and mapped to CIDOC-CRM ontology has been developed and probably could offer a solution to this debate.

³ Its definition in an XML schema, together with the specification document, can be found at www.lido-schema.org. LIDO is the result of a collaborative effort of international stakeholders in the museum sector, starting in 2008, to create a common solution for contributing cultural heritage content to portals and other repositories of aggregated resources. Being an application of the CIDOC Conceptual Reference Model (CRM) it provides an explicit format to deliver museum’s object information in a standardized way.

⁴ CARARE metadata schema specification can be found at: <http://www.carare.eu/eng/Resources/CARARE-metadata-schema-outline-v1.0>

elements related to ‘Archaeology’ (the archaeological context, the place of discovering, etc.); ‘Phil’ is the flag that will stand for the sub-wrapper that contains the metadata elements related to ‘Philology’ (genre of the text, metre, etc.); ‘Geol’ is the flag that will stand for the sub-wrapper that contains the metadata elements related to ‘Geology’ (e.g. material of the support), and so forth. Table I describes how flags related to wrappers.

TABLE I.

‘Archaeo’	Place of discovery, Ancient name, Region / District, Regio antique, Country of discovery, Date of discovery, Discoverer, Condition of discovery, Event of discovery [...]
‘Phil’	Genre, Metre, Scripture, Writing direction, Number of signs, Style of writing, Spelling mistakes [...]
‘Geol’	Material of the support, Geographical provenance of the material [...]

Going into further details regarding the inscription itself, particular attention is paid to the metadata fields that describe its content, which will allow users to have a holistic view about it. This metadata schema takes of course into consideration the description of the Real World Object and results into an elaborated source of information whereby one gets knowledge about the item itself. This information ‘narrates’ the ‘life’ of the object from the ancient times until the present era. For example, there is a focus on the people and places that are mentioned in the text of the inscription and their connection, as well as on the relations between the event in which it has been created within a particular time frame. This information reveals the interconnection between time, space and people that will help users to interpret and understand an inscription in full. From the philological perspective, the metadata schema provides vital information for philologists who can access data like: critical apparatus, genre of the text of the inscription, metre, philological debates and spelling errors. Attention is given also to the Bibliography and to the Commentary, as wrappers organized in their turn in sub-wrappers. In Table II some examples represent the metadata fields and the description of the elements (in red, when possible, the specific information taken from the Epigram E21 of the AKG Digital Corpus).

TABLE II.

INSCRIPTION	
Title	The title of inscription that names the inscription, e.g. in AKG, E1, E72 or in other corpora: The Inscription of Gaius Maximus etc. - E21
Inventory number	The ID number of the real object - 1959/VII-23/1, IG 208
Date	Exact date of the inscription
Time span/	The period in which the inscription is dated (e.g.

Period	Hellenistic period etc.) - Cypro-Classical
Terminus ante quem	The earliest date that the inscription is dated - late 4th century BC
Terminus post quem	The latest date that the inscription is dated - early 3rd century BC
Author of inscription	The author of the inscription, e.g. the creator of the inscription itself, e.g. the name of the artist or the name of the artistic/artisan school
Honorand	To whom the inscription is devoted: e.g. the girl Caecilia who died young - Young (?) Myrto
Awarded by	The person that devotes the inscription, e.g. Caesar devotes this inscription to (...) - Unknown
Social role	e.g. The social role of the mentioned people (...)
Event	The event related to the purpose of the inscription, e.g. death of Augustus - Myrto’s death

Besides the inscriptions themselves, since the Cypriot collection consists of their digital “surrogates” (pictures, 3D models of inscriptions, movies etc.), the new metadata schema takes into consideration their descriptive features, their digital provenance and other related information. For example, through this metadata it is possible to describe the digital provenance of the 3D model of an inscription. The information is, therefore, organized in wrappers and sub-wrappers for the description of the acquisition (the technique, the tool, the specification of the tool, the specification of the output, etc.) the post-processing (the operative info, the file specifications, etc.) and the digital output (data format, software, according to the kind of digital resource obtained).

Table III shows some of the field’s elements regarding for example the description of a 3D model of an inscription.

TABLE III.

DIGITAL RESOURCE PROVENANCE	
Acquisition	
3D scanning	Type; Device specifications (Model, Software, Points_second, Time acquisition, Number of scans, Acquisition range, Accuracy, Target model, Number of targets); File specification (Number of points max, Data input format, Data output format, Data decimation, Data weight, Data compression)
Processing	
3D modelling	Operative Info (Operator, Collaborators, Paradata, Date); Technical info (Software); File specifications (Geometry type, Data of points max, Data input format, Data output format, Number of vertices, Number of faces, Data

	decimation, Data weight, Data compression)
Digital outcome	
3D model	Descriptive data (subject, type, description, link, Digital object, IsShownAt, rights); Technical info (software, Data format, Data weight), 3D model (Dimensions, number of vertices, number of faces, textures, rendering time)

IV. IMPLICATIONS FOR RESEARCH INFRASTRUCTURES

A repository containing a rich, multi-disciplinary metadata descriptive content, coupled with data organised according to a conceptual reference model (e.g. CIDOC-CRM) has several immediate benefits: it better and more accurately reflects the nature of past artefacts and better represents their multi-layered nature, it helps gaining new insights embedded in such objects and enhances the transparency of their interpretation. Moreover, such an approach enhances the integration of the repository into a wider perspective and enables its incorporation in global research infrastructures to be developed in the future. ARIADNE⁵, a newly established EU funded project aimed at developing a digital infrastructure for archaeological research, is one of the best candidates for providing such an integrative platform. As such, organising data according to a rich, cross-domain metadata and to a standard conceptual reference model (CIDOC-CRM) is the first step towards achieving a multi-disciplinary research infrastructure.

V. IMPLICATIONS FOR EUROPEANA

CyI, as Content Provider, will afford Europeana the metadata of the Cypriot inscription corpus, enabling a wider visibility and increased awareness to the Archaia Kypriaki Grammateia Digital Corpus Project (AKGDC). Within the framework of the EAGLE, the most prominent European institutions and archives in the field of Classical Latin and Greek epigraphy, will provide Europeana with a collection of unique historical sources. In particular, it will supply inscriptions coming from 25 countries, providing more than 1.5 M of images and related metadata, including translations of selected texts for the benefit of the general public. This aggregation process will be possible through the mapping from the metadata schema used by the Content Providers participating in the project to a common inscription-specific metadata model (developed within the project community) based on standards and recommendations (developed through related Europeana projects) and a service platform for epigraphy to allow for multi-format ingestion and multi-device. The publication into Europeana, since it is a web portal "public users-driven", guarantees at the one hand more visibility for the Cultural Institution within a wider community [13], presenting to the general public a type of subject usually addressed to small communities and often not published online. On the other hand, Europeana will host the contribution of specific content that it was not contemplated previously as a stand-alone aggregation, and maybe it was just published as an item inside museum collections that of course describe it as a museum

⁵ <http://www.ariadne-infrastructure.eu>

artefact and not as an inscription. In this way, Europeana will be enriched with a critical mass of quality-oriented content.

VI. CONCLUSIONS

This research is anticipated to contribute to the study of digital inscriptions in particular and to give a new take in the Digital Humanities community in general. At a later stage it will be developed further and enhanced. The new CyI cross-disciplinary metadata schema designed for Ancient Cypriot Inscriptions can be promoted as a model for other digital inscription repositories which present similar content in order to describe it and place it in its wider context.

ACKNOWLEDGMENTS

The results presented in this paper were partially supported by a grant from the EU funded project EAGLE.

REFERENCES

- [1] Mahoney A. Electronic Textual Editing: Epigraphy, in TEI - Text Encoding Initiatives. http://www.tei-c.org/About/Archive_new/ETE/Preview/mahoney.xml
- [2] Cayless H., Roueché C., Elliott T., Bodard G., "Epigraphy in 2017." *Digital Humanities Quarterly* 3.1 (2009). https://www.forth.gr/ics/isl/publications/paperlink/doerr3_caa2004.pdf
- [3] Cayless H., Tools for Digital Epigraphy. In *Proceedings of the Association for Computing in the Humanities/Association for Literary and Linguistic Computing*, Athens GA, 2003.
- [4] Doerr M., Schaller K., Theodoridou M., Integration of complementary archaeological sources (2004) http://www.ics.forth.gr/ics/publications/paperlink/doerr3_caa2004.pdf
- [5] Pitzalis D., Christophorou E., Kyriacou N., Georgiadou A., Niccolucci F., Building Scholar e-communities Using a Semantically Aware Framework: Archaia Kypriaki Grammateia Digital Corpus. VAST: International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage 2012, pp. 89-95.
- [6] Ronzino P., Amico N., Niccolucci F., Assessment and comparison of metadata schemas for architectural heritage, *Proceedings of CIPA 2011*.
- [7] Felle A. E., Esperienze diverse e complementari nel trattamento digitale delle fonti epigrafiche: il caso di EAGLE ed EpiDoc, in *Diritto romano e scienze antichistiche nell'era digitale. Convegno di studio* (Firenze, 12-13 settembre 2011), Torino 2012 [Collectanea Graeco-Romana. Studi e strumenti per la ricerca storico-giuridica, 10], 117-130.
- [8] Manacorda D., Archeologia ed epigrafia: problemi di metodo a proposito di CIL, VI 8960, in *Archeologia w teorii i w praktyce*, A. Buko-P. Urbańczyk (red.), Warszawa 2000.
- [9] Alvarez F. L., Garcia-Barriocanal E., Gomez Pantoja J. L., Sharing Epigraphic Information as Linked Data, in *Metadata and Semantic Research: 4th International Conference, MTSR 2009* (Salvador Sanchez-Alonso, Ioannis N. Athanasiadis eds.)
- [10] Gómez-Pantoja J. L., Álvarez F. L., García-Barriocanal E., From relational databases to Linked Data in Epigraphy: *Hispania Epigraphica Online*, Elena García-Barriocanal, Zeynel Cebeci, Mehmet C. Okur and Aydın Öztürk (eds.), *Metadata and Semantic Research (Proceedings of the 5th International Conference, MTSR 2011, Izmir, Turkey, October 12-14, 2011)*, Communications in Computer and Information Science Volume 240, Springer, Berlin Heidelberg, 2011
- [11] Babeu A., Classics, "Digital Classics" and Issues for Data Curation. <http://guide.dhcurator.org/research-practices/classics/>
- [12] Ronzino P., Niccolucci F., Hermon S., "A metadata schema for Cultural Heritage documentation", in *EVA 2012 Florence*, Firenze University Press.
- [13] Vassallo V., Piccininno M., Aggregating Content for Europeana: A Workflow to Support Content Providers, *Lecture Notes in Computer Science*, 2012, Volume 7489, Theory and Practice of Digital Libraries, Rasmussen and F. Loizides (eds.), Springer.