DELIVERABLE

Project Acronym:	EAGLE
Grant Agreement number:	325122
Project Title:	Europeana network of Ancient Greek and Latin Epigraphy

---

Content harmonisation guidelines, including GIS and terminologies - Second Release

D 2.2.2

version n°4.0

Authors:

Pietro Maria Liuzzo (UHEI)
Donato Fasolini (UAH)
Anita Rocco (UNIBA)

Collaborators:

Andrea Mannocci (CNR-ISTI)
Chiara Latronico (EUROPEANA)
Claudio Prandoni (Promoter)

Reviewers:

Franco Zoppi (CNR-ISTI)
Francesco Mambrini (DAI)

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	X
C	Confidential, only for members of the consortium and the Commission Services	

Revision History

	Date	Authors	Organization	Description
v.1	12 January 2014	Pietro Liuzzo	UHEI	Deliverable was put together from the work of the WG1 and 2, updating all relevant parts from release D2.2.1.
v. 2	26 January 2014	Claudio Prandoni	PROMOTER	Revision of the document and updates for submission to the Working Group
v. 3.1	26 January 2014	Pietro Liuzzo	UHEI	Updates with feedback from WG and latest updates to vocabularies
v. 3.2	11 February 2014	Franco Zoppi and Francesco Mambrini	CNR-ISTI and DAI	Review and integration of suggested corrections
v. 3.3	16 February 2014	Claudio Prandoni	PROMOTER	Further revision and changes to vocabularies section
v. 4	18 February 2014	Pietro Liuzzo	UHEI	Updated and fixed minor typos

More details on contributors of each section can be found where appropriate.

Statement of originality:
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Contents

List of Tables

List of Figures

List of abbreviations

Table 2: List of Abbreviations

Abbreviation	Meaning
AE	L'Année épigraphique
BP	Best Practice: a technique or methodology that, through experience and research, has proven to reliably lead to a desired result
BPN	Stands for Best Practice Network. Eagle is a BPN, which is *an instruments that aims to promote the adoption of standards and specifications for making European digital libraries more accessible and usable by combining the consensus building and awareness raising function of a thematic network with the large-scale implementation in real-life context of one or more concrete specifications or standards by its members.* This is the definition provided by the ICT *Policy Support Programme Competitiveness and Innovation Framework Programme 2010* of the European Commission
CIDOC-CRM	International Committee for Documentation - Conceptual Reference Model
CIL	Corpus Inscriptionum Latinarum
DOW	Description of Work
EDM	Europeana Data Model
EMW	EAGLE Media Wiki
LOD	Linked Open Data
MS	Milestones
M	Month
OA	Open Annotation
RDF	Resource Description Framework
SI	Success Indicators

Continued on next page

Table 2 – continued from previous page

Abbreviation	Meaning
SKOS	Simple Knowledge Organization System
HE	Harmonization Effort
HT	Harmonization Tools
URI	Unique Resource Identifier
WC	Wikimedia Commons
WP	Work Package is one of the main parts in which the Eagle Project is articulated • WP1: Project management • WP2: Networking and best practices • WP3: Metadata model, mapping and ingestion • WP4: EAGLE Aggregator and Image Management infrastructure • WP5: End-user dedicated services • WP6: Dissemination and exploitation
WPL	Work Package Leader. These are: • WP1: Promoter • WP2: UHEI • WP3: CYI • WP4: CNR-ISTI • WP5: DAI • WP6: UNIROMA1
WG	Working Groups are the core groups of activity in WP2. They are: • WG1: GIS and Terminologies • WG2: Translations and content curation • WG3: IPR and User Engagement

Summary of all Best Practices and Harmonization Actions

Best practices

Harmonization Efforts

Harmonization Tools and Resources

List of Contributors

The following list includes members of the Working Group on GIS and terminologies and all those members of the international community which provided support with feedback, initial suggestions, documentation, technical support, and all other forms of direct and indirect participation in the development of the contents of this deliverable, via personal meetings, conversations at conferences and events, online meetings, mail exchanges on mailing lists, etc. . We would like also to acknowledge here all those contributions which would not be possibly referenced otherwise. While contributions have been stated precisely where possible in the deliverable, it should be underlined that the responsibility for the contents stays with the project members.

Table 3: List of Contributors

Name	Affiliation
Bridget Almas	Tufts University
Elton Barker	Open University
Dávid Bartus	EÖTVÖS LORÁND TUDOMÁNYEGYETEM
Marie-Claire Beaulieu	Tufts University
Simon Bence	EÖTVÖS LORÁND TUDOMÁNYEGYETEM
Alice Bencivenni	Università di Bologna
Monica Berti	Universität Leipzig
Gabriel Bodard	King's College London
Alfredo Buonopane	Università degli Studi di Verona
Hugh Cayless	Duke University
Sven Ole Clemens	Deutsches Archäologisches Institut
Hannah Cornwell	AshLI project
James Cowey	University of Heidelberg
Pietro De Nicolao	Wikimedia Italia
Denitsa Dzhigova	Universität Heidelberg
Jacob Egger	Paris Lodron Universität Salzburg
Tom Elliott	New York University

Continued on next page

Table 3 – continued from previous page

Name	Affiliation
Silvia Evangelisti	Università di Foggia
Antonio Enrico Felle	Università di Bari
Donato Fasolini	Universidad de Alcalà
Francisca Feraudi-Gruénais	Heidelberg Akademie der Wissenschaften
Diego Ferreira	TemaTres
Joaquìn L. Gómez Pantoja	Universidad de Alcalà
Brigitte Gräf	Heidelberg Akademie der Wissenschaften
Frank Greishaber	University of Heidelberg
Ethan Gruber	American Numismatic Society
Corey A. Harper	New York University
Sebastian Heath	Institute for the Study of the Ancient World
Piso Ioan	Universitatea Babes Bolyai
Leif Isaksen	University of Southampton
Askold Ivantchik	Ausonius
Daniel Kinzler	Wikimedia Deutschland
Allison Kupietzky	The Israel Museum Jerusalem
Susan Hazan	The Israel Museum Jerusalem
MaximilianHeiden	Deutsches Archäologisches Institut
Michael Howard Kay	Saxonica - Wrox
Stephen Lambert	University of Cardiff
Lorenzo Losa	Wikimedia Italia
Francesco Mambrini	Deutsches Archäologisches Institut
Luca Martinelli	Wikimedia Italia
Marion Lamé	CNR – Istituto di Linguistica Computazionale
Richard Light	expert in Linked Open Data
Elly Milonas	Brown University
Milagros Navarro Caballero	Ausonius
Silvio Panciera	Università di Roma
Olga Pelcer	Historical Institute of Montenegro
Johnathan Prag	ISicily project
Lydia Pintscher	Wikimedia Deutschland
Tatjana Sandon	University of Edinburgh
Nathalie Prévot	Ausonius
Eydel Rivero Ruìz	Universidad de Alcalà
Anita Rocco	Università di Bari
Eleonora Santin	CNRS - UMR HiSoMA
Miguel Angel Sicilia	Universidad de Alcalà
Rainer Simon	Austrian Institute of Technology
Simona Stoyanova	Universität Leipzig

Continued on next page

Table 3 – continued from previous page

Name	Affiliation
Christian Uhlir	Paris Lodron Universität Salzburg
Rada Varga	Universitatea Babes Bolyai
Scott Vanderbilt	Roman Inscription of Britain online (Los Angeles)
Valentina Vassallo	The Cyprus Institute
Herbert Verreth	KU Leuven
Bryan Ward-Perkins	University of Oxford
Claudio Zaccaria	Università degli Studi di Trieste
Andrea Zanni	Wikimedia Italia

We would also like to acknowledge all contributors of the EAGLE conferences and workshops.[1] And finally we would also like to acknowledge for their continued contribution to the dissemination of epigraphic knowledge the ca. 500 volunteers in Wikimedia Commons[2] who have in time uploaded and shared photos of greek and latin inscriptions and user:Fae who enabled us to see them.

[1] See Orlandi et al. 2014, list of contributors.

[2] `https://commons.wikimedia.org/wiki/Category_talk:Media_contributed_by_EAGLE/reports/volunteers`

Executive Summary

This Deliverable on *Content harmonization guidelines, including GIS and terminologies* is the second release of a document defining best practices selected and used during the activities of the Working Group on Content Harmonization, GIS and Terminologies (WG 1) as well as the Working Group on Translations and Content Curation (WG 2, Liuzzo, Santucci, and Prandoni 2013, 15f.).

The vocabularies as well as all other tools adopted and developed which are here presented:

- aim to ensure the continuous addition of qualitative and quantitative content by the EAGLE BPN;
- are designed to be extensible and inclusive in order to facilitate enrichment and harmonization of metadata using controlled vocabularies and gazetteers (Description 2013, 25 part B);
- foster in a most practical way interoperability and integration of the content providers systems, in order to enrich the value of the aggregated epigraphic content (Description 2013, 11 part B);
- are developed to produce and support the creation of further Linked Open Data resources and annotations;
- take into account the latest developments in schemas, ontologies and systems which are involved and affect stakeholders;
- are aligned to the development of the Metadata schema and of the EAGLE portal.

Therefore the results here presented are, in the order in which they appear in the document

- the principles, methodology and results of the gathering in Trismegistos - Places of all GIS information provided by Content Providers;
- the Vocabularies developed by the WG on all shared fields of information and the methodology used to produce and maintain them (the print out has been omitted from this second release because a new html version of the vocabularies is available online;
- efforts in the harmonization of contents, including the EAGLE Mediawiki for translations and the EAGLE Zotero Group Bibliography;

- tools to facilitate harmonization XSL Transformation stylesheets developed to support the integration of vocabularies in the metadata model.
- partnership established in the framework of the Networking Task (Liuzzo, Santucci, and Prandoni 2013, 7f.) to enhance results at the best possible level for the international community, especially the collaboration with Wikimedia projects

At the end of each chapter is available a list of the key best practices identified, which are also listed in a summary at the beginning of the document.

All the tools developed are available and free license software and content, embrace the Linked Open Data principles and strive to be challenging new directions for the development of a top quality network of resources in the BPN (Description 2013, 53-4 part B).

Chapter 1

Introduction

Zeng and Chan 2006 in their work on Metadata interoperability say

> the results of efforts to improve interoperability can be observed at different levels:
>
> **Schema level** Efforts are focused on the elements of the schemas, being independent of any applications. The results usually appear as derived element sets or encoded schemas, crosswalks, application profiles, and element registries.
>
> **Record level** Efforts are intended to integrate the metadata records through the mapping of the elements according to the semantic meanings of these elements. Common results include converted records and new records resulting from combining values of existing records.
>
> **Repository level** With harvested or integrated records from varying sources, efforts at this level focus on mapping value strings associated with particular elements (e.g., terms associated with subject or format elements). The results enable cross-collection searching.

In this guidelines the second level is the main aim. The work has been carried out following the methodologies outlined in Liuzzo, Santucci, and Prandoni 2013 with the collaboration of all the members of the working group. Aim of the two working groups involved in this tasks (WG1 on GIS and Terminologies especially, but also WG2 on Translations and Content curation) is the harmonization of entries in the different databases under all possible aspects without flattening existing divergences. This tasks have been faced in the framework set by the work on the Eagle Metadata Schema and its foreseen developments (Rivero Ruiz, Andrea, and Vassallo 2013 and *addenda*) which dealt with the first of the levels mentioned above mainly. All guidelines for interoperability and harmonization here

suggested are also thought about with the third level in mind as well as with attention to the work being done on mobile applications(Martin 2014), interface (AA.VV 2013), portal requirements (Mambrini 2013) and portal (Prandoni, Alfarano, and Casarosa 2014).

The data aggregated by the Eagle BPN and its new affiliated members, collaborators and partner projects is very heterogenous in format, quantity, background concepts of data etc. Efforts towards record harmonization had in most cases therefore to be limited to a specific subset (metadata of texts of inscriptions), although all that could be done to harmonize the whole data involved has been done.

For the whole set of aggregated contents the Eagle BPN focused on mapping and harmonization of Geographic information and bibliographic information (and partially of controlled vocabularies), two major tasks eased by the existence among the Content Providers of contents which were already structured. This does not decrease the amount of work required to curate contents, eliminate duplicates and obtain tools of real interest and usability for a typically very skeptical community of users (classicists, epigraphists, ancient historians, etc.).

The subset of information related instead to the metadata of texts of inscriptions provided the occasion to work on the harmonization of a metadata format for aggregation according to the existing international standards; and secondly to undertake an effort towards a major *desideratum* of our user community: the harmonization of controlled vocabularies and terminologies.

Incremental development and publication

Now that also the website and API of the EAGLE project are online we have satisfied the requirement of the community of users to have direct access and visibility of resources, even if they are continuously in progress. The best practice which the developer word in the linked open data context sees as the best is in clear opposition to the habits of the epigraphic community where a perfect final edition, whatever the time it might require to make is the agreed best practice. This is also due to the kind of publication means used until now, printed publication, requiring a finished product not one that is subject to incremental development. The resources of the EAGLE Consortium have therefore been published as incrementally developed and dynamic resources and best practices have been selected also to make this possible, thanks to the cooperation and input of external advisors and Europeana professionals. A consistent SKOS version of the vocabularies was produced which reflected fully the characteristics recognized as best practices and consequently the abandonment of the previously used TEMATRES software was decided as the constraints it imposed limited the work to be done and its exploitation.[1] This change allowed also to meet look and feel requirements from the users, as the possibility to select a language from the multilingual vocabularies and the possibility to see a simple hierarchy of terms instead of the all relational articulation. All the source files, the transformations handling the SKOS (much simpler than many of those available around) and the content to be

[1] The TEMATRES instances remain as a back up but are not anymore updated.

published are all available in a long term sustainable GIT group which has been shared with members which will land stability to the whole of the development efforts under the framework of the EAGLE project.

Overall best practices

What we believe in fact to be key to useful guidelines is the responsibility towards the **users** and groups interested in the content provided for quality and commitment to the same research purposes. This is the reason of the close connection and interaction of this guidelines with the efforts of User Engagement tasks.

On the other side as second key best practice we have, and we continue to, activate as many useful partnerships with sister projects as possible, to enhance the possibilities to reach the needs of the first key best practice.

Interaction is key to real success both in interoperability for the production of useful linked data (ancient data in the case of EAGLE). To this respect this guidelines ought to the LAWD (Linked Ancient World Data) community, the Epidoc Community, the Europeana Network, the Pelagios Network and the Perseus project a wealth of opportunities to practice this key best practice which is interaction.

These guidelines therefore always keep the above two points in mind and reasons for the choices made in the points suggested are based on these two points.

Key Best Practices Summary for this section

BP 1: Responsibility towards users Seek best quality and usability. Understand needs of main groups of users and focus efforts towards task which bring direct benefit.

BP 2: Interaction and Networking Understand differences, involve stakeholders, build skills as well as tools to simplify harmonization in the interest of everyone and with an eye to the future.

BP 17: Incremental Development Availability of resource online with a incremental development approach so that frequent updates do not prevent the use of the existing resources at their present stage of development.

Chapter 2

Updates and main changes from D2.2.1

Main changes to this document included in this second release are:

1. new and updated best practices;
2. new and updated resources;
3. new and updated guidelines.

The structure of this document omits now the full text in the appendixes of the vocabularies because all these are now fully available and dynamically accessible, readable and constantly updated on the EAGLE portal together with the other resources made available by the EAGLE BPN. The code and source data is also available in a GIT group of the EAGLE BPN, and is now fully integrated with the Content checker, the EAGLE portal and Europeana.

2.1 New and Updated best practices

2.1.1 Incremental Development

All resources built and curated by the EAGLE BPN are now accessible online directly from the homepage of the EAGLE website. These include:

- vocabularies (`http://www.eagle-network.eu/resources/vocabularies/`);
- translations' wiki (`http://www.eagle-network.eu/resources/translations/`);
- Zotero Bibliography (`http://www.eagle-network.eu/resources/bibliography/`).

Each resource is introduced briefly, is open to feedback and meets requests of interaction and usability which could not be met before. The Vocabularies are now published directly into the EAGLE website and are browsable by language, hierarchy, full relational table and concept by concept. The user is also able to distinguish between preferred and translated terms.

2.1.2 Alignment of gazetteers

Trismegistos GEO has implemented the cooperation with Pleiades and Pelagios and performed an alignment of the gazetteer of toponyms between Trismegistos and Pleiades with very good results, although the automated part of the work did not turn out to be successful as much as expected.

2.1.3 Further steps towards Linked Open Data

Although content negotiation is not yet implemented several requirements from external advisors and Europeana have been implemented as best practices:

- extended use of URIs;
- replacement of examples;
- rematching and alignment with external vocabularies.

We have nevertheless not accepted forced hierarchy and full standardization of languages and relations as this is not a shared best practice with the scientific community of epigraphers. While we had been advised to reduce the languages of the vocabulary to one or the other we have decided to keep the consistency with the attested use of terms and give instead tools to better browse the vocabularies and full labeling of the information provided.

2.2 New and Updated Resources

2.2.1 The EAGLE Vocabularies

The Vocabularies (`http://www.eagle-network.eu/resources/vocabularies/`) have been updated with new terms, new relations, and new examples and definitions, new ways of browsing and searching, further searches options and are integrated with the portal look and feel.

2.2.2 The EAGLE Mediawiki of Translations

The EAGLE Mediawiki (`www.eagle-network.eu/wiki/index.php/Main_Page`) has been enriched with many new translations from school projects, workshops, conferences and individual volunteers. A board of editors has been set up and a full set of guidelines as been published.

2.2.3 The GIT group repository

Following external advice we have adopted a GIT repository (`https://github.com/EAGLE-BPN`) to control versioning and reuse of the source data and to share

the results with the international community. The GIT repository contains most of the reusable code developed to work with inscriptions by the EAGLE BPN.

2.2.4 The Zotero Group Bibliography

Structured bibliographies of the Content providers have been merged and made directly available in a Zotero Group (`https://www.zotero.org/groups/eagleepigraphicbibliography`), where they have been incremented by direct entry from external partners and update from non-structured data sources. The bibliography of the proceedings of the first EAGLE international conference has been developed from this resource and also integration with the epigraphic data is ongoing.

2.2.5 Images contributed to Wikimedia Commons

A large number of images have been edited in Wikimedia Commons (`https://commons.wikimedia.org/wiki/Category:Media_contributed_by_EAGLE`), activating a large number of volunteers in that community and using the GLAM Toolkit two initial subsets of photos have been uploaded to Wikimedia Commons for reuse and sustainability purposes. Work on refinement of metadata continues smoothly.

2.3 New and Updated Guidelines

2.3.1 Guidelines and Tutorials

The Zotero bibliography, the upload and elaboration of metadata in Wikimedia Commons, as well as the editing of the SKOS vocabulary in GIT have seen a need for new guidelines and tutorials. The users nevertheless preferred ad hoc introductions rather than documents. Only the EAGLE Mediawiki received instead a full re-elaboration of the guidelines which where integrated in the wiki so that new and returning users could benefit of a continuously available reference. This guidelines include the use of Wikibase, the access guidelines and most of all guidelines for the production of new contents. Guidelines for the translation of Greek and Latin inscription have been elaborated by the EAGLE BPN and have been discussed at the Paris International Conference.

2.3.2 Tailoring and specialization of up-conversion and harmonization code

The Code which performs the harmonization of editorial conventions among text, although not yet perfect, have been extensively restructured and improved to cover many more cases and needs:

- disambiguation by addition of all known ids (which included the production of a comparative table of ids for disambiguation);
- export of multiple text outputs to the EAGLE metadata model;
- a number of complex instances of transformation of abbreviations, meaningful gaps, signs and symbols, have also been implemented;
- connection of bibliography with the online Zotero bibliography.

Chapter 3

Geographical Informations in Trismegistos

3.1 Definitions

> The most fundamental advantage of the emerging networked knowledge environment is that it provides a much-improved technological basis for sharing resources of all sorts from all sources. This situation increases the importance of effective access to information. Place, along with time, topic, and creator, is one of the fundamental components in how we define things and search for them. (Buckland and Lancaster 2014)

Due to this consideration and its implication it is important to better define the terms which come into play first of all: *place* and *toponym*.

3.1.1 Place

Among many possible definition available, the working definition adopted here for ”Place” is the one used at Trismegistos.

> The term ’place’ is used in the broadest sense, referring not only to towns and villages, but also to regions, districts and to all kinds of micro-toponyms (e.g. town quarters and streets, kleroi and other plots of land, rivers, sanctuaries.[1]

[1] Contents in this section are taken from Trismegistos.

3.1.2 Toponym

Toponym

A place can have more than one name (*Toponym*), depending on the language, the period or the type of text; e.g. the names Apollonopolis Megale (Magna) / Ano (Superior), Apollonias, Behdet, Bachthis, Polis Phoibou and Mesen all refer to modern Edfu.

All toponyms referring to one place are listed on a single card, which has a unique TM Geo_ID number. This number contains no information, but is just created in a numerical order.[2] With regard to ancient places it is not always clear what is a real toponym and what is a common noun referring to a geographical item (also called appellatives in linguistic studies). Any toponym listed in the geographical index of a publication is also listed in the geographical database.[3]

3.1.3 Workflow

Variants

For every ancient place the different names are listed in separate fields, one for every ancient language used: 'Greek', 'Latin', 'Egyptian' (sc. hieroglyphic, hieratic, demotic), 'Coptic'. The names in the fields 'Greek', 'Latin' and 'Coptic' are written in the script of the sources (with a Unicode alphabet); in the field 'Egyptian' a Unicode transcription is given, no facsimile of the original hieroglyphs or script. In the field 'Other' Latin alphabet transcriptions are given of toponyms in other languages such as Arabic, Hebrew, Nabataean, Meroitic, Old Persian and Assyrian. In the field 'Modern' the name is written (in Latin alphabet transcription) of the present-day toponym that corresponds with the ancient name. The names in all these fields are 'standardized' to a certain extent, the way one expects in the geographical index of a book. In the field 'Variants' all these names are brought together in Latin alphabet transcriptions, and other variant spellings or translations of the names can be added. To facilitate searching by non-specialists, the transcriptions in this field contain as little diacritical signs as possible. For Egyptian names the scientific transcription is therefore replaced by a more phonetic one (e.g. Ḫftḥ-n-Bḥ > Chefeteh-n-Bouchis). Greek names are transliterated into the Latin alphabet, not translated into Latin, although a common Latin form can be added as one of the variants (e.g. Ὀξύρυγχος > Oxyrynchos, variant Oxyrhynchus). For some very common names (e.g. Egypt, Alexandria, Thebes) the English spelling is used. From all these variants a standard name is chosen both for the ancient and the modern name. The ancient standard name is displayed in the field 'Standard', the modern name in the field 'Modern' already mentioned. If a place has no ancient name, the modern name is displayed in the field 'Standard'.

[2] If two places are identified and their cards joined, the Geo_ID number of the old card is preserved but henceforward contains only a reference to the new card.

[3] Talbert 2000, Vf. gives complete guidelines on this theme as a supplement to the maps and Gazetteer of the Barrington Atlas.

3.1.4 Modern Toponyms

It is not always easy to determine which modern name should be used to refer to an ancient site. The name of the modern hamlet, the village, the town or the major city in the neighbourhood, all these modern names can be used to refer to the ancient site. If this is the only site on the territory of these places, one can list all four names next to each other in the field 'Variants' and pick as standard the name most commonly used in scholarly literature to refer to the site, without any risk of confusion. If there are, however, several sites on the territory of the town, it is best to choose the name of the hamlet or village as standard name and to refer to the town in the field 'Location'.

Another problem arises when a major city covers a lot of modern toponyms and when it is important not to loose the information about the different actual sites where documents have been found. These sites are put separately in the field 'Variants', preceded by the phrase 'sites including ...'; e.g. L00 Alexandria - sites including Kom el-Dikka; L01 Memphis - sites including Saqqara. The Provenance file (see below) contains a field for the information about these specific sites.

A geo-card has been created for every toponym in the full lists, referring to the place where an inscription has been found or written.

The process of data control has therefore been dialogical: the update of information in Trismegistos has been a way to clarify possible inconsistencies in a bi-directionally valuable work of curation of contents.

All these toponyms have been processed into the Trismegistos Geo file, which now counts 42.550 toponyms, from all over the *Imperium Romanum* and beyond.

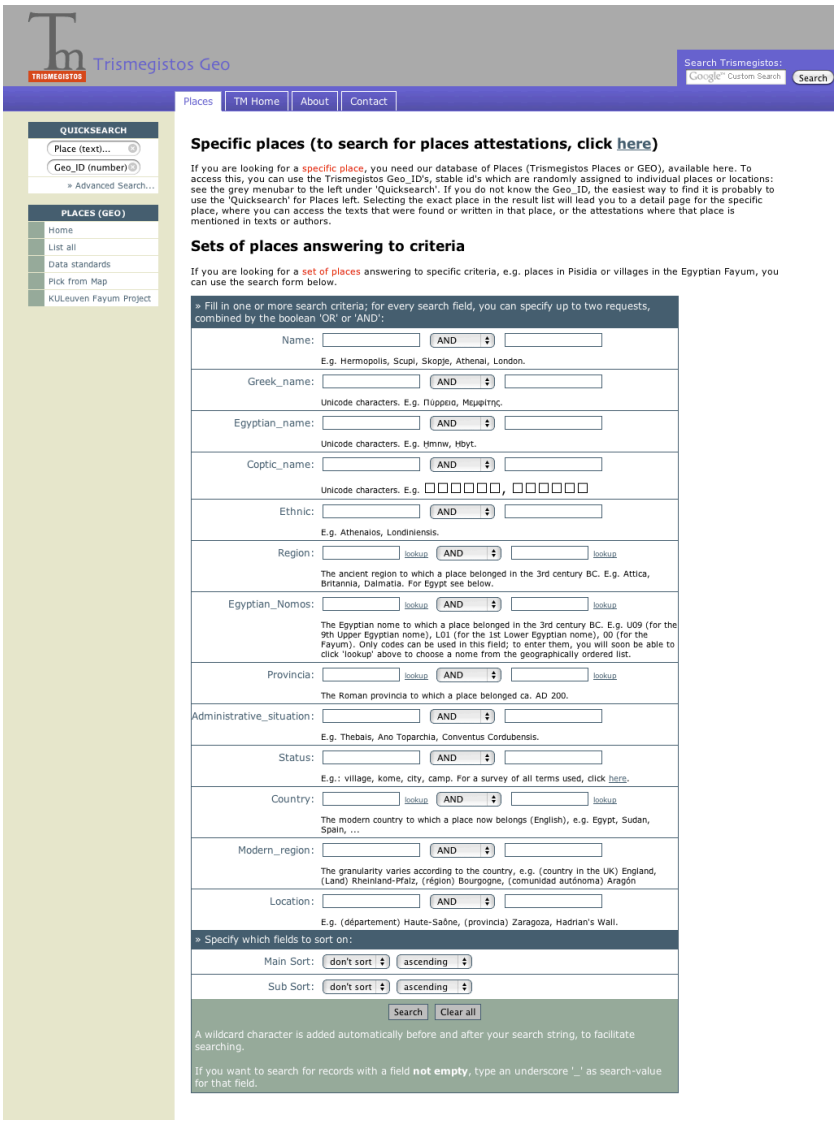


Figure 3.1: Trismegistos Geo Search

Every toponym in TM Geo is identified by a unique number. E.g. Roma = TM Geo 2058.

Figure 3.2: Rome Card in Trismegistos Geo

This number is linked to a set of geographical data: the ancient name of the place, the modern name, the modern country it belongs to, the ancient region (in the 3rd century BC), the Roman *Provincia* (in ca. 200 AD), the modern region (3.2) and a more precise location within a modern district or its distance towards a nearby town. e.g. TM Geo 20422:

- Latin name: Narbo
- modern name: Narbonne
- modern country: France
- ancient region: Gallia
- Roman provincia: Narbonensis
- modern region: Languedoc-Roussillon
- location: in the département Aude

Since not every set of data contains all the necessary information, work is still going on in order to complete the contents of the TM Geo cards.

It is important for searches that the spelling of the ancient toponyms is harmonized. Toponyms from the Greek east are transliterated into latin alphabet, but usually also a Latin spelling is added. E.g. standard name: Nikopolis, Latin variant: Nicopolis. Also other variant spellings have been be added.

3.2 General structure of Geographical information

The regional definitions belong to a long established use in the studies of antiquities. Due to the swift and fast changes in the administration and structure of the Roman Empire, not to speak of Greek cities and hellenistic Kingdoms, for the purpose of cataloguing, reference is often made to specific periods in which the administrative structure was particularly clear from the sources we have, and lasted long enough. This might be made much better and far more precise using LOD practices but would require a separate project due to the complexity of the scope, so the one presented here would be necessarily temporary best practice and a good ground to improve data in the future. Nevertheless in each bibliographical reference specific information on each item can be found.

Initially the following structure was shared and discussed in the working group as the regional organizational framework.

3.2.1 Ancient regions (3rd century BC)

- Acarnania
- Achaia
- Achaia Phthiotis
- Aemilia
- Aeolis
- Aethiopia
- Aetolia
- Africa
- Apulia
- Arabia
- Arcadia
- Argolis
- Armenia
- Arsinoites
- Attica
- Bactria
- Bithynia
- Boeotia
- Bosporus Cimmerius
- Britannia
- Bruttium
- Campania
- Cappadocia
- Caria
- Cataonia
- Cilicia
- Corsica
- Creta
- Cyclades
- Cyrenaica
- Dacia
- Eastern desert
- Elam
- Elis
- Epirus
- Etruria
- Euboea
- Galatia
- Gallia
- Germania (sc. east/north of the Rhine)
- Hibernia
- Hispania
- Illyricum
- Ionia
- Isauria
- Isthmos
- Italia
- Lower Egypt
- Laconia
- Latium
- Lemnos
- Lesbos
- Liguria
- Locris
- Lucania
- Lycaonia
- Lycia
- Lydia
- Macedonia
- Malis
- Marmarica
- Mauretania
- Media
- Mesopotamia
- Messenia
- Moesia
- Mysia
- Noricum
- Northern Sinai
- Numidia
- Paeonia
- Palestina
- Pamphylia
- Pannonia
- Paphlagonia
- Phocis
- Phoenicia
- Phrygia
- Picenum
- Pisidia
- Pontus
- Raetia
- Rhodos
- Samnium
- Sardinia
- Sarmatia
- Scythia
- Sicilia
- Sinai
- Sogdiana
- Sporades
- Syria
- Thasos
- Thessalia
- Thracia
- Transpadana
- Troas
- Upper Egypt
- Umbria
- Venetia
- Western coast
- Western desert

3.2.2 Roman provinciae (ca. AD 200)

- Achaia
- Aegyptus
- Africa Proconsularis
- Alpes Cottiae
- Alpes Graiae
- Alpes Maritimae
- Alpes Poeninae
- Aquitania

- Arabia
- Armenia
- Asia
- Baetica
- Belgica
- Bithynia et Pontus
- Britannia
- Cappadocia
- Cilicia
- Corsica
- Creta
- Cyprus
- Cyrenaica
- Dacia
- Dalmatia
- Epirus
- Galatia
- Germania Inferior
- Germania Superior
- Hispania Citerior
- Iudaea
- Lugdunensis
- Lusitania
- Lycia et Pamphylia
- Macedonia
- Mauretania Caesariensis
- Mauretania Tingitana
- Mesopotamia
- Moesia Inferior
- Moesia Superior
- Narbonensis
- Noricum
- Numidia
- Pannonia Inferior
- Pannonia Superior
- Raetia
- Sardinia
- Sicilia, Melita
- Syria
- Thracia

3

3.2.3 Italian regiones (from 41 BC)

- Aemilia (Regio VIII)
- Apulia et Calabria (Regio II)
- Bruttium et Lucania (Regio III)
- Etruria (Regio VII)
- Latium et Campania (Regio I)
- Liguria (Regio IX)
- Picenum (Regio V)
- Roma
- Samnium (Regio IV)
- Transpadana (Regio XI)
- Umbria (Regio VI)
- Venetia et Histria (Regio X)

3.2.4 Roma

- Regio I Latium et Campania
- Regio II Apulia et Calabria
- Regio III Bruttium et Lucania
- Regio IV Samnium
- Regio V Picenum
- Regio VI Umbria
- Regio VII Etruria
- Regio VIII Aemilia
- Regio IX Liguria

- Regio X Venetia et Histria
- Regio XI Transpadana

These are accepted denomination for all Content providers. The chronological limits implied in this are a choice due to the complexity which would need to be reflected for different earlier and later periods. For example, while the archaic world would need mostly a city by city organization with little and very localized broader territorial organization, on the other end of the chronological spectrum of antiquity, there are provinces and administrative divisions of the late antiquity which are subject to very frequent changes. The above articulation allows a good localization and a fair information.

3.3 Trismegistos Geo and Pelagios

In the Framework of a memorandum of understanding between the EAGLE BPN and the PELAGIOS 3 project,[4] the gazetteer of places in Trismegistos Places will be exported as OA[5] to join the Pelagios Network. This will involve an alignment between Pleiades and Trismegistos Geo, which will be probably done via Wikidata and will produce multiple identifiers for places. This result is a potential drawback in terms of our intended goal of producing unique identifiers; however we have differently defined identities and to keep them separated but linked is the path towards to richest possible linked ancient data network.[6]

Trismegistos GEO gathered some 5300 identifications of TM Geo IDs with Pleiades IDs and some 1000 identifications with Geonames: these numbers provide a concrete starting point for aligning with Pelagios the EAGLE data. Compared to the 42.550 Trismegistos GEO ID this might seem not much and we did hope that more identifications could be made automatically by matching TM Geo with Pleiades. Nevertheless the number of exact matches given from the automated process was not so great: the small differences in spelling in both databases and the large number of homonyms prevented a broader harvest of automatic matches. On the other hand, checking manually one by one the 35.000 remaining TM Geo's without a Pleiades number in http://ryanfb.github.io/pleiades-static-search/, is an enormous task… TM Geo I finished the identification of the 3434 toponyms in the *Itinerarium provinciarum Antonini* and of the 3287 toponyms in the *Tabula Peutingeriana*.[7] In Pleiades on the other hand, sometimes specific references are given to these sources. If it would we possible to get a list of all

The *Itinerarium Antonini* and the *Tabula Peutingeriana*

[4] Many extremely interesting application related to this project can be seen in the blog http://pelagios-project.blogspot.de/

[5] http://www.openannotation.org/

[6] A similar case can be found in Stevenson 2012.

[7] For which 1190 toponyms are already identified - the website `http://francia.ahlfeldt.se/index.php` is a great help for both texts.

Pleiades numbers with a reference to these two sources, we can link quite easily from Pleiades to TM Georef (which contains the attestations of the two texts) and so to the Pleiades-number field in TM Geo. For the *Itinerarium* the situation is quite straightforward: we entered every toponym as it occurs in the edition of Cuntz (1929). Every toponym has been placed in its modern country, its ancient region (3rd century BC) and its Roman provincia (ca. 200 AD). For 2784 instances out of 3434 we have a Pleiades number (and for 650 instances we don't). For the *Tabula Peutingeriana* we entered every toponym as it occurs in the online edition of `http://www.cambridge.org/us/talbert/talbertdatabase/prm.html`, IF it contains an actual name. The Tabula also lists some entries which are unnamed, and those we only incorporated if I happened to find an identification with a known place (but I did not check out every last one of these items). This new information is already available in Trismegistos online and is updated at the end of every week.

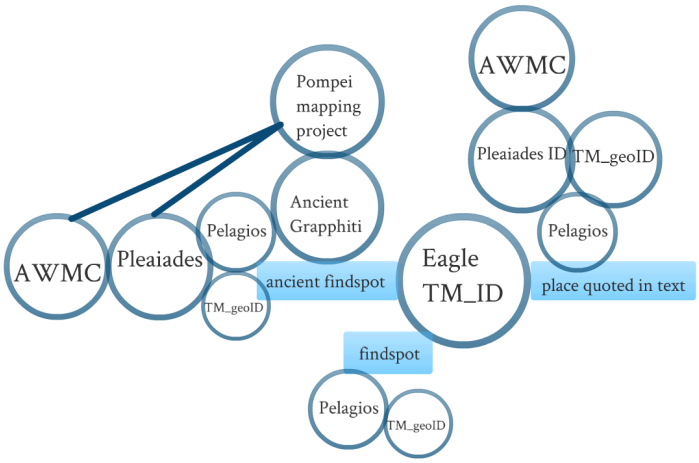


Figure 3.3: Linked Geo-IDs

Trismegistos and Pelagios

This is an affirmed and successful best practice, already implemented by some of EAGLE members and among world wide digital ancient world project. This again is a step towards making data available for Linked Open Data project and contributes towards the creation of the Web of Resources.

EAGLE/Trismegistos will host annotations relating to EAGLE partner's content. Pelagios 3 will makes annotations available through its Demonstrator API, so that also each EAGLE BPN partner and the EAGLE portal will be able to exploit with Pelagios 3 API and Widgets this data to enrich contents. Eagle portal will have Pelagios 3 widgets and the data will also be used for EAGLE mobile applications for image recognition and storytelling.

To exploit the harmonization effort taken on this regard, we have prepared also tools that allow to enrich metadata directly before ingestion with URIs pointing to the Trismegistos cards. This harmonization tool is a XSL stylesheet which matches the contents which have been sent to Trismegistos with the data in the metadata record and is integrated with the other tools described below (see p.61).

As a last step in this field for harmonization we look forward to and endorse the practice of Europeana for Semantic enrichment.[8] Data about places will be enriched in EDM with links to four types of reference resources, including geoNames as envisaged also in the EAGLE Metadata Model (Rivero Ruiz, Andrea, and Vassallo 2013) for places.

[8] Isaac, Clayphan, and Haslhofer 2012, p. 37

Key Best Practices Summary for this chapter

BP 3: Toponyms and identifiers Give each place a single card, which has a unique ID number (Pleaides and Trismegistos Geo are the two existing gazetteers for ancient places). This number should contain no information but be a stable reference. No reuse of IDs.

BP 3.1: Alignment of gazeteers Pleaiades and Trismegisto Gazeteers have been aligned with an automated process and a lot of hand work.

BP 4: Variant names Names shall be also in Latin alphabet transcription. Variant spellings or translations of the names shall always be preserved, for example in a specific field with some diacritics to facilitate searching by non-specialists.

BP 5: Spelling Harmonization Latin transliteration and spelling can be added in variant names to enhance the possibilities to find relevant results for users.

HE 1: Networking and related contents Open Annotation will be made for all toponyms to be part of larger places network, as the Pelagios Network. This allows connection with all other related resources via toponyms.

HT 1: XSLT for harmonization of places id and URI allows to insert TM Geo ID. When HE1 will be done with the ID it will be possible to see all possible related content from the numerous partners of Pleaides and Pelagios network.

Chapter 4

Vocabularies

4.1 Linked Open Data Vocabularies for Epigraphy

As Di Stefano Manzella [1] clearly explained classification is no easy issue in any field: epigraphy is no exception to this rule. Traditionally the CIL VI (Rome) classification has been used as a reference, as this typology has served as a model for all epigraphic production in the Roman Empire. There are nevertheless new glossaries and classification curated by CIL, which retain the limits of a formal classification, together with the benefits of this.

Problems are various, and include also the use of terms across vocabularies and the doubts which might be generated by archaeological chance. Piso 2001, pp. XI-XII for example noted:[2]

> Au vrai, l'ancienne école épigraphique s'est peu souciée du support archéologique des inscriptions et du contexte archéologique dans lequel elles étaient découvertes. Pour le support archéologique j'ai tenté chaque fois qu'il m'était possible d'en identifier la catégorie et la fonction. Puisqu'on tend, à juste titre, vers une histoire totale, il est important, par exemple, de savoir si telle plaque inscrite était encastrée dans la base d'une statue honorifique, dans un monument funéraire ou dans le mur d'un édifice quelonque. Il y a ensuite toute une série de monuments qui présentent les éléments d'un autel, notement le socle, le dé médien et le couronnement, mais sans en remplir le fonction. Dans ce cas c'est la fonction qui compte, et non pas l'aspect extérieur. Ce qui caractérise un autel, c'est le focus, tandis

[1] Manzella 1987, p. 109
[2] Following a current of studies which has its main in G. Susini and J.-N. Bonneville.

> que dans le couronnemente d'une base (piédestal) sont souvent pratiqués des orifices qui servent à fixer la statue. Pourtant, les deux éléments manquent souvent, car on peut toujour recourir à un focus mobile et, d'autre part, une statue peut se soutenir par son propre poids. En pareil cas, j'ai noté "autel ou base de statue".

Epigraphy is one of the disciplines for the study of History, for some it is not a discipline on its own. But for this same reason there is nothing that can be neglected and maximum precision is required. Now that also an entirely new discipline with its own criteria, methods and definitions is becoming truth, this is even more necessary and needs to be done in the most up to date and efficient way for the benefit for both Epigraphy, Digital Epigraphy and all cognate disciplines as Ancient History, Classics, Archaeology.

In the description of the object which bares an inscription, for example, at least four kind of information should be looked at:

- date;
- function;
- material;
- state of preservation.

Sometimes it is impossible to distinguish because:

- objects with different function might be identical when observed;
- an object type might depend on its function and viceversa.

For example we can take two object which might be identified as one or the other or might not. Altars and statue base may appear identical to the observers, and the same object may be interpreted by different scholars as either one or the other. And yet they represent two distinct object types when one comes to classifications.

An altar is a monument which is self defined and it is used to make sacrifices and libations. The base of a statue instead is the support for a statue.

It will certainly be an altar if there is a *focus* for sacrifices or traces of it, but that can be missing: it could have also been made of metal or ceramic and therefore be movable. An altar would also be recognizable from a decoration with *pulvini* or a *patera* or a *urceus*. The text of the inscription can also be a way to define the monument type. On the other end we are certain that the monument is a Statue base if there are signs of the statue and where it was fixed to its base. But also in this case a statue could just stand because of its weight. This is also the case when we deal with techniques of execution of an inscription. Those can also be multiple and an inscription can be for example both a graffito and painted. For this reason in the metadata model these information are allowed for multiple times, but vocabularies do not define multiple terms. This is one best practice in the definition of vocabularies which is very important for the nature of contents at

play. This is also the reason why the EAGLE BPN believes that it is best practice to provide not just definitions of all main terms but also examples from different areas and in different state of preservation.

The Europeana LOD practice for metadata recommends the adoption of machine readable vocabularies. The WG has adopted Linked Open Data practices and approaches[3] to address the existing problems in classification and in publication of a machine readable vocabulary of values (controlled vocabulary).[4] LOD

Linked Open Data principles as stated by Berners-Lee (2006) are as follows:

1. use URIs as names for things;[5]
2. use HTTP URIs so that people can look up those names;
3. when someone looks up a URI, provide useful information, using the standards (RDF, SPARQL);[6]
4. include links to other URIs, so that they can discover more things.

The Working Group therefore initially looked for a ready tool which would allow direct RDF publication of the data.

TemaTres[7] was initially chosen from a series of available options open or commercial

VocBench a free software for publishing thesauri: http://aims.fao.org/tools ;
GINCO a thesaurus management software: http://data.culture.fr/ ;
PoolParty a commercial software: http://www.poolparty.biz/test-demo/ ;
TopBraid a commercial software: http://www.topquadrant.com/solutions/ ;
iQvoc a commercial software: http://iqvoc.net/ .

This free Vocabulary Management system compared to the others offered a very high usability, export facilities which allow for SKOS, DC and several other standards, constant and active support from developers.

Compared against the above principles we have done the following to create a useful tool for networking and harmonization which could be based on currently used best practices but also allow for further improvement:

1. each item of EAGLE vocabularies in TemaTres has a URI, also non preferred and target terms.
2. all EAGLE vocabularies uris are HTTP.

[3] Bizer, Heath, and Berners-Lee 2009

[4] Harper et al. 2012, pp. 4-5 for the definition. See also Isaac, Clayphan, and Haslhofer 2012

[5] Defined as *generic means to identify any entity that exists in the world.* Bizer, Heath, and Berners-Lee 2009

[6] *RDF provides a generic, graph-based data model with which to structure and link data that describes things in the world.* Bizer, Heath, and Berners-Lee 2009

[7] http://r020.com.ar/tematres/

3. upon acknowledgment of the requirements of existing users, we provide for each main term:

    - definition;
    - bibliography;
    - alternative terms;
    - translations of terms.

4. In the EAGLE vocabularies there are URIs linking to several other vocabularies and more will be constantly added by users.

Still to be achieved is the possibility of content negotiation for data users.

TemaTres has been a very useful tool to develop the initial version of the EAGLE vocabularies but has problems and takes assumptions which did not make it any more fit for purpose for the EAGLE project as both external advisers and Europeana professional told us during social media consultations and the aggregation process respectively. This was in part linked to the different handling of languages which the EAGLE BPN adopted, which is closer to scientific practice rather than to the defined clarity required at top level from users of the data. A consistent SKOS version of the vocabularies was produced which reflected fully the characteristics recognized as best practices and consequently the abandonment of the previously used TemaTres software was decided as the constraints it imposed limited the work to be done and its exploitation.[8]

In facts the many problems and complexities which Manzella 1987 underlined and exposed can find a solution using the LOD approach.[9] Both the limits of choices and hierarchical organization can be bypassed using unique identifiers and relations among those. Poli-hierarchical structures as well as the presence of many possible denomination with specific and self standing definition allows for both precision, consistency, sustainability over time of this approach.[10]

For example, an inscription can easily be classified both as *magistrati populi romani* and as *decuriones municipales*; a document can be classed both under the fasti and under the individual. These major advantages can be exploited when the effort of publishing open machine readable material is undertaken and a document can be then classified chronologically and alphabetically without the need of an authorial choice.

Nevertheless, the major possible advantage is that there is no patent need, from a LOD approach, to distinguish among Greek and Latin inscriptions: they can find their place together in a Linked Data edition and benefit of other efforts in this direction (See section 7 page 71.).

[8] The TemaTres instances remain as a back up but are not anymore updated.

[9] World Wide Web Consortium 2012; Bizer, Heath, and Berners-Lee 2009

[10] Harpring 2010

4.1.1 Methodology

As for Metadata Vocabulary, also in the establishment of controlled vocabularies of terms the best practice is to minimize the creation and proliferation of vocabularies by harmonizing the existing ones and matching whenever possible.[11] The alignment had to take place first among the vocabularies in use, in the form of simple lists. These will be then matched to existing vocabularies. This mapping task requires a very broad expertise in the field and indeed some terms turned out to be problematic and are still under discussion. There is no problem in this and indeed it is an enrichment process, when tools that allow for this discussion and consequent changes are in place. Therefore we believe there is no effective and real vocabulary without a dynamic community behind it discussing its terms and maintaining it. Openness and availability alone are not enough. The WG supported this method as a best practice which further stakeholder could use for the alignment of medium sized heterogenous terminologies. A first draft has been then proposed to several mailing lists in the Digital Humanities field for comments and met a very positive welcome as a first extensive attempt to address a centuries long problem. This fully supports and exceeds what envisaged in the DOW(Description 2013, 10 part A) already at this first release:[12]

> Aim of this task is to provide guidelines and recommendations to ensure the continuous addition of qualitative and quantitative attractive content by the EAGLE BPN, including the definition of vocabularies and terminologies to be used to enrich and harmonize the content and the analysis of the tools and practices to adequately geo-reference it. Best practices will be identified on how to include in EAGLE metadata model the most suitable vocabularies and terminologies and how to harmonize the geo-referencing of the epigraphic content.

The same approach that the Pleiades and Pelagios project took for geographical data was adapted here to seven other possible ways of linking data for the international community.

The following are the principles used to chose a main term among the aligned ones, its related terms and terms in target vocabularies. They partially follow and mix both the preferred practice in establishing vocabularies of the Linked Data community, and those of the epigraphers in the EAGLE WG on GIS and Terminologies. No Vocabulary' language has priority and the vocabulary is fully multilingual but with specific and consistent declarations of the language for each occurrence:

[11] Harper et al. 2012, p. 7

[12] The all process of development of these vocabularies has been carried out together with the effort of definition of the metadata model (Rivero Ruiz, Andrea, and Vassallo 2013) at all stages.

1. on occurrence of aligned terms, the language more attested was used, or, preferably, the one term which *had already a definition*;
2. where conflicting options could be found the English was chosen and where no English was available the Latin;
3. in cases when one term was only attested in one vocabulary, the original language of that vocabulary was maintained;
4. terms in the main language of a vocabulary which were not chosen as a main term were stored as a *Target Vocabulary* term of the main chosen one rather then as translated terms;
5. any translation of terms, once for each language, is inserted as a *Non Preferred Term*.

4.1.2 Content Enrichment with Tematres

Content Enrichment

Content enrichment has been initially and conveniently done with the use of the aligned vocabulary only in TemaTres.[13] Among the advantages of TemaTres there was the possibility to expand the search from the buttons in the bottom right corner which perform searches of the term being viewed in the main Google search engines.

We entered only one term in one language for each of the aligned terms. Due to a limit in language encoding of the chosen software, we also decide not to input terms written in the same way in different languages, if not in the case of Target Vocabularies.

Most entries have been enriched with bibliography and examples also in the new html representation of the vocabularies. They retained all functionality of the TemaTres instances and have also the possibility to further query on external databases and on the EAGLE portal directly from the vocabulary. Instead of inserting images in the TemaTres vocabulary the links for further searches allow the user to perform further searches.

TemaTres did not allow the declaration of independent languages for alternative terms, we have then added these consistently in the final export, which became the base for the further developments suggested by our network. and declared declared In a second data curation phase we have then added language specifications for each term, directly in the SKOS file together with all specific terms.

```
<skos:Concept rdf:about="http://www.eagle-network.eu/voc/material/lod/30">
      <skos:prefLabel xml:lang="de">(vulkanischer) Tuff</skos:prefLabel>
      <skos:altLabel xml:lang="de">CIL: Tuff, vulkanischer vor allem Peperin</skos:altLabel>
      <skos:altLabel xml:lang="en">(volcanic) tuff</skos:altLabel>
<skos:altLabel xml:lang="de">Tuff</skos:altLabel>
      <skos:altLabel xml:lang="es">Toba</skos:altLabel>
      <skos:altLabel xml:lang="fr">Tuf</skos:altLabel>
      <skos:altLabel xml:lang="fr">tuf volcanique</skos:altLabel>
```

[13] See Liuzzo, Santucci, and Prandoni 2013, 26f.

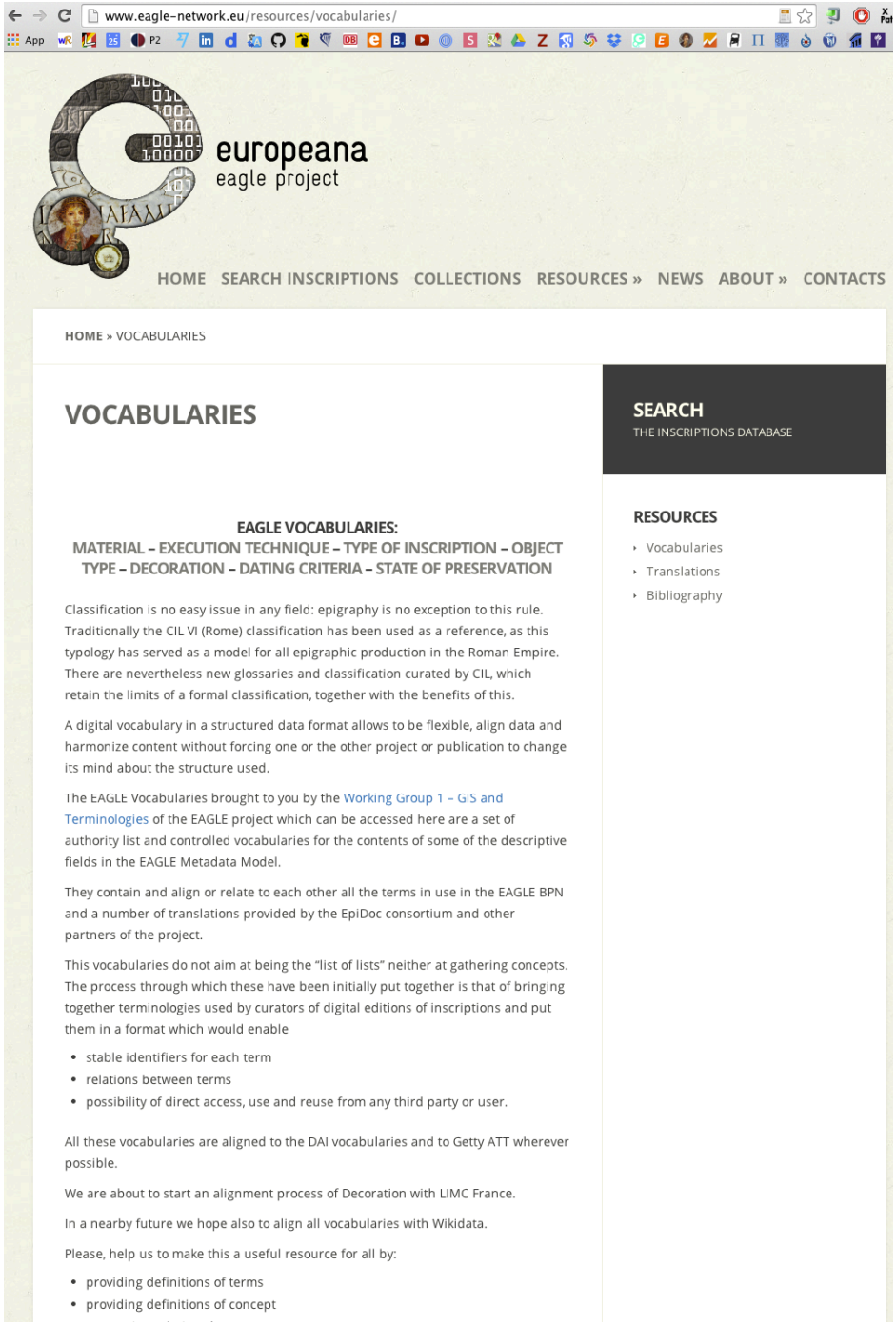


Figure 4.1: The access page to the EAGLE Vocabularies

FASTI, LEGES, ACTA

EAGLE Vocabulary - Type of Inscription

preferred label	relation	term	language
Fasti, Leges, Acta			la
	Translated term	Calendario, elenco di magistrati, legge, senatoconsulto, costituzione imperiale, editto, decreto, atti di collegio	it
	Translated term	Fasten / Verzeichnisse,	de
	Translated term	Fastes, lois, actes	fr
	Translated term	Gesetze, Akten	de
	Translated term	Official text	en
	Translated term	επίσημο έγγραφο	el
	Translated term	النص الرسمي	ar
	Definition	calendars, edicts of magistrates, every type of legislative document (laws, decrees of the senate, imperial constitutions, edicts, decrees, acts of collegia)	en
	Definition	it. calendari, elenchi di magistrati, ogni tipo di documento legislativo (leggi, senatoconsulti, costituzioni imperiali, editti, decreti, atti di collegi, ecc.)	it
	Examples	HD000674	de
	Examples	EDR003059	la

```
    <skos:altLabel xml:lang="hu">tufa</skos:altLabel>
    <skos:inScheme rdf:resource="http://www.eagle-network.eu/voc/material/"/>
    <skos:broader rdf:resource="http://www.eagle-network.eu/voc/material/lod/28"/>
    <skos:exactMatch>
      <skos:Concept rdf:about="http://archwort.dainst.org/thesaurus/de/vocab/?tema=6303">
        <skos:prefLabel xml:lang="de">Tuff</skos:prefLabel>
      </skos:Concept>
    </skos:exactMatch>
    <skos:closeMatch>
      <skos:Concept rdf:about="http://chc.sbg.ac.at/sri/thesaurus/35"/>
    </skos:closeMatch>
    <dct:created>2013-08-26 16:54:50</dct:created>
  </skos:Concept>
```

Each term acquired in TemaTres a URI in this form
`http://www.eagle-network.eu/voc/typeins/lod/113`

where the part after "voc" is the name of the vocabulary and the number is a unique reference into the specific vocabulary. This is not the consecutive number in each letter alphabetic list. The decision to build uris with the vocabulary name as well, is dictated by the presence of the same term in different vocabularies.

The above link will point to the HTML representation of a corrsponding skos concept, and the skos version will be reached by selecting the uri
`http://www.eagle-network.eu/voc/typeins/skos/113`

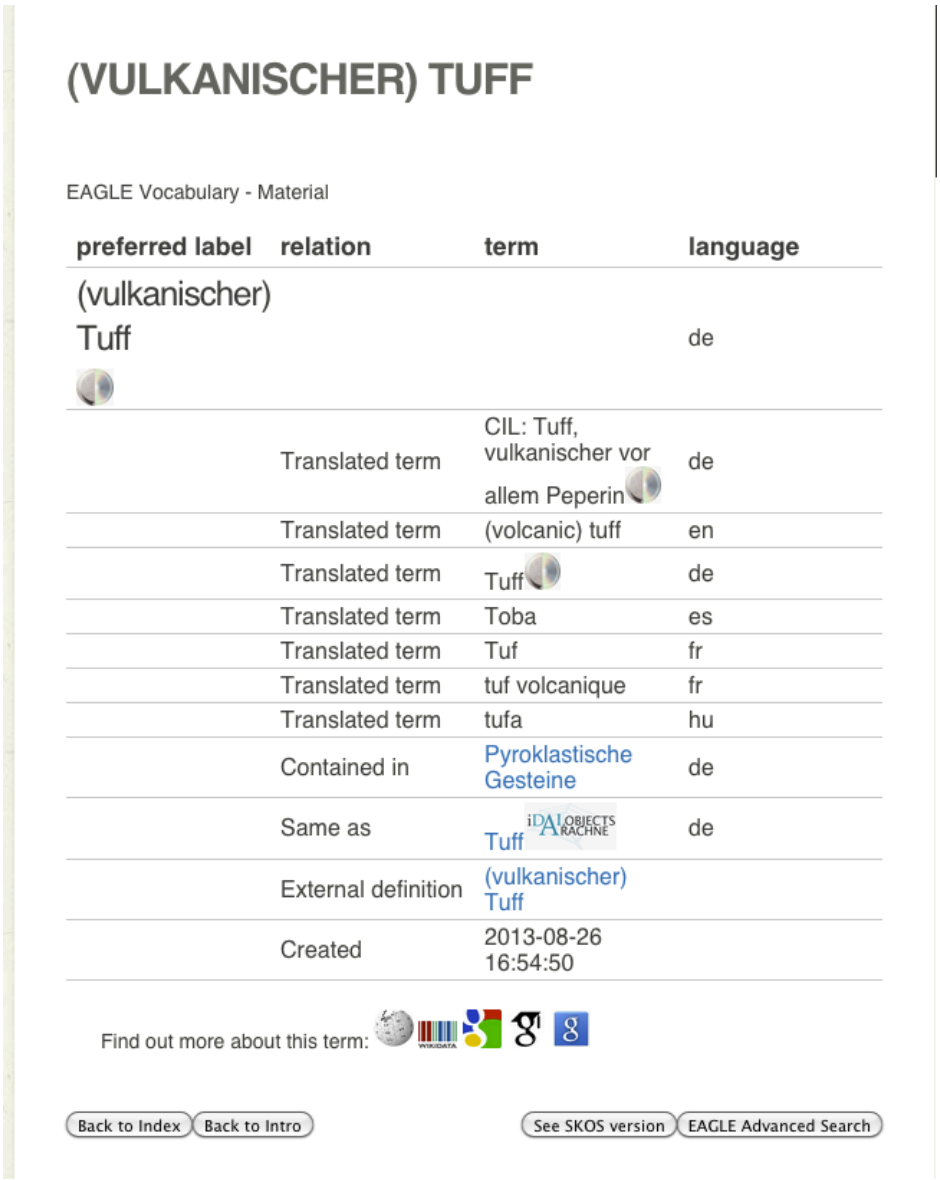
Content negotiation is now possible. When a request is sent for an item, e.g. `http://www.eagle-network.eu/voc/material/304` a code 303 is returned and the user is redirected to the requested resource. Therefore is a browser sends the request the system points to `http://www.eagle-network.eu/voc/material/lod/304.html`, if anther client request RDF `http://www.eagle-network.eu/voc/material/skos/304.rdf` will be returned

The SKOS/RDF files exported from TemaTres and curated with language and some further useful data can be found in the EAGLE open repository (see p.66):
`https://github.com/EAGLE-BPN`

All these source files, the transformations handling the SKOS (much simpler than many of those available around) and the content released are all available in this long term sustainable GIT group which has been shared with a number of members which will contribute to land stability to the whole of the development efforts under the framework of the EAGLE project.

The following are the uri of the *EAGLE vocabularies, published in the new Resource section of the EAGLE portal*:

- `http://www.eagle-network.eu/voc/typeins.html` ;
- `http://www.eagle-network.eu/voc/objtyp.html` ;
- `http://www.eagle-network.eu/voc/material.html` ;



(VULKANISCHER) TUFF

EAGLE Vocabulary - Material

preferred label	relation	term	language
(vulkanischer) Tuff			de
	Translated term	CIL: Tuff, vulkanischer vor allem Peperin	de
	Translated term	(volcanic) tuff	en
	Translated term	Tuff	de
	Translated term	Toba	es
	Translated term	Tuf	fr
	Translated term	tuf volcanique	fr
	Translated term	tufa	hu
	Contained in	Pyroklastische Gesteine	de
	Same as	Tuff	de
	External definition	(vulkanischer) Tuff	
	Created	2013-08-26 16:54:50	

Find out more about this term:

Back to Index | Back to Intro | See SKOS version | EAGLE Advanced Search

Figure 4.3: HTML representation of the concept Tuff

- `http://www.eagle-network.eu/voc/writing.html` ;
- `http://www.eagle-network.eu/voc/decor.html` ;
- `http://www.eagle-network.eu/voc/statepreserv.html` ;
- `http://www.eagle-network.eu/voc/dates.html` ;

Among the latest improvement to the vocabularies Hungarian has just been added and Portuguese translations are on their way for most terms.

ADNUNTIATIO

EAGLE Vocabulary - Type of Inscription

preferred label	relation	term	language
Adnuntiatio			la
	Definition	Bekanntmachung; Ankdigung (z. B. von munera) Nichtrechtliche Verfungen(s.Rechtliche Verfung, fentlich / privat)	de
	Examples	HD053407	de
	Created	2013-08-01 12:27:53	
	Modified	2013-08-15 16:56:04	

Find out more about this term:

Back to Index | Back to Intro | See SKOS version | EAGLE Advanced Search

Figure 4.4: HTML representation of the concept Adnuntiatio

Careful and selected alignment to Getty Art and Archaeology Thesaurus has been done as well as with the DAI Vocabulary, also developed using TemaTres, and the CIL glossaries.[14] Alignment with Wikidata in the most efficient way is being studied with the Wikimedia Italia and Wikimedia Deutschland teams and will be performed as soon as possible following on one of the most widely recognized best practices for this kind of data, to be linked to major datasources. Further alignments might also be done in the future.

4.1.3 The Newly published vocabularies online

The change from TemaTres to a stand alone one allowed also to meet look and feel requirements from the users:

- the possibility to select a language from the multilingual vocabularies;

[14] http://cil.bbaw.de/cil_en/dateien/glossar.php

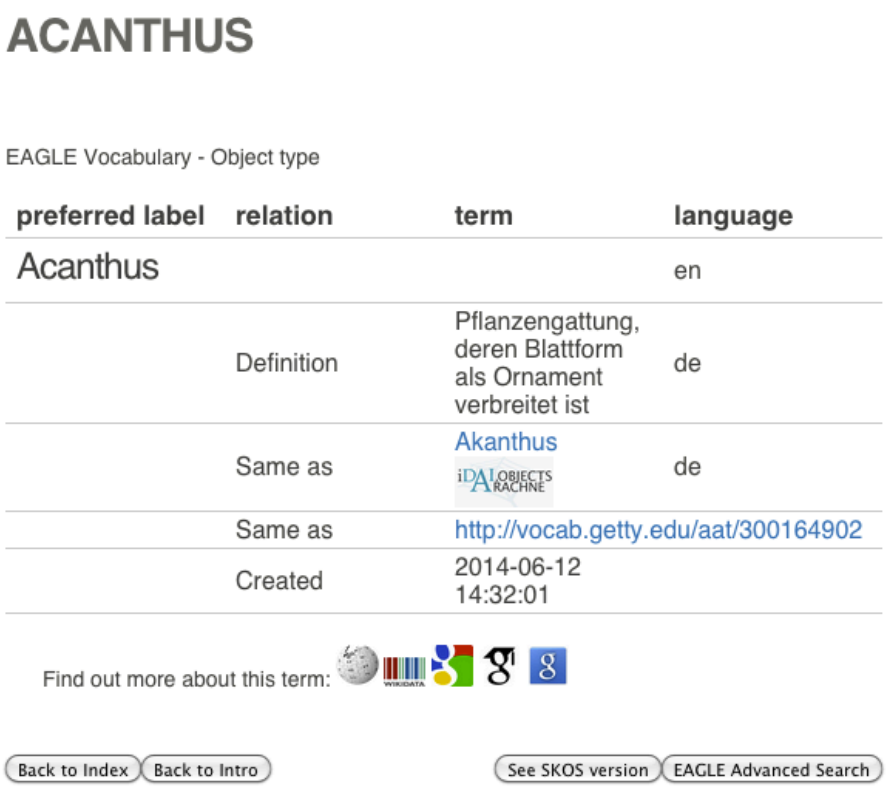
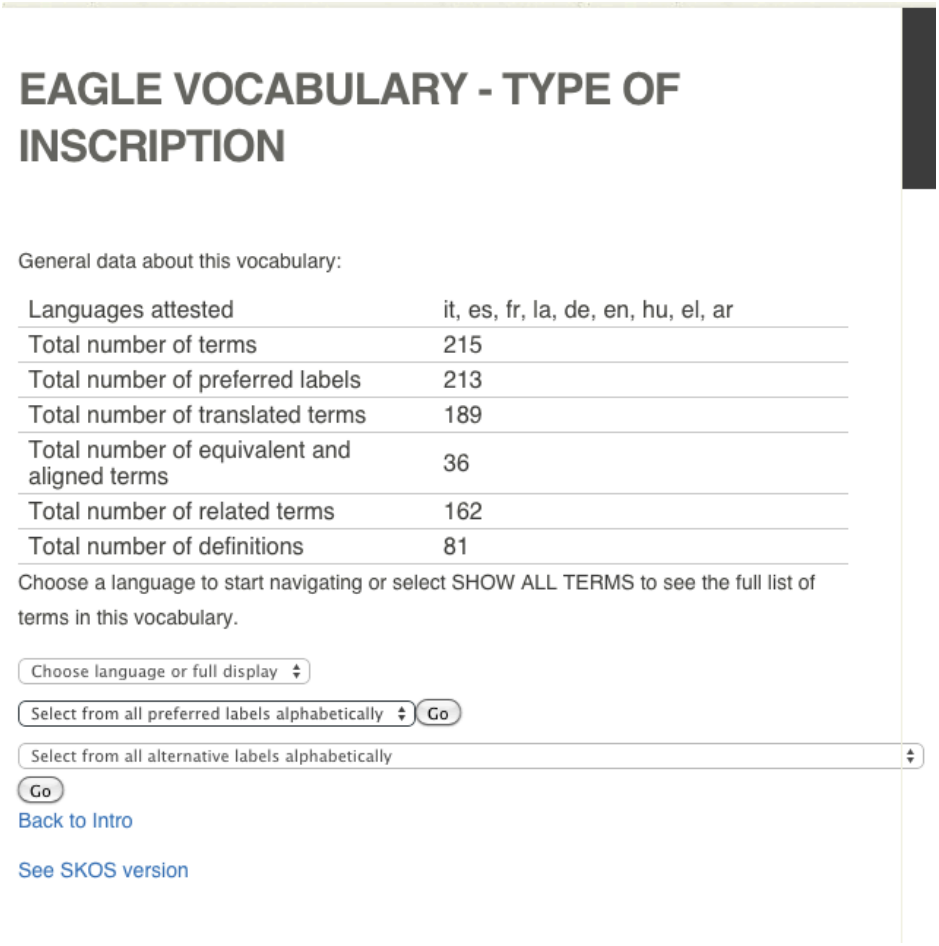


Figure 4.5: HTML representation of the concept Acanthus, with further searches, definitions and two aligned terms

- the possibility to see a simple hierarchy of terms instead of the all relational articulation;
- the possibility to prioritize the language selection and filter out all values in other languages;
- the possibility to further searches back on partner vocabularies and on the EAGLE portal;
- a google custom search engine to search the vocabularies;
- continuity of look and feel with the EAGLE portal.

While the choice of a flat kind of vocabulary is the best and most flexible in most cases, in two of the EAGLE Vocabularies we have opted for a hierarchical structure: Material (4.3) and Dating Criteria (4.7).[15] For these two vocabulary a visualization of the hierarchy is also possible on selection by the user.

[15]The Dating criteria vocabulary has been vastly improved during an editing spring organized *ad hoc* in January.



EAGLE VOCABULARY - TYPE OF INSCRIPTION

General data about this vocabulary:

Languages attested	it, es, fr, la, de, en, hu, el, ar
Total number of terms	215
Total number of preferred labels	213
Total number of translated terms	189
Total number of equivalent and aligned terms	36
Total number of related terms	162
Total number of definitions	81

Choose a language to start navigating or select SHOW ALL TERMS to see the full list of terms in this vocabulary.

Choose language or full display

Select from all preferred labels alphabetically Go

Select from all alternative labels alphabetically

Go

Back to Intro

See SKOS version

Figure 4.6: Start page for the Vocabulary Type of Inscription

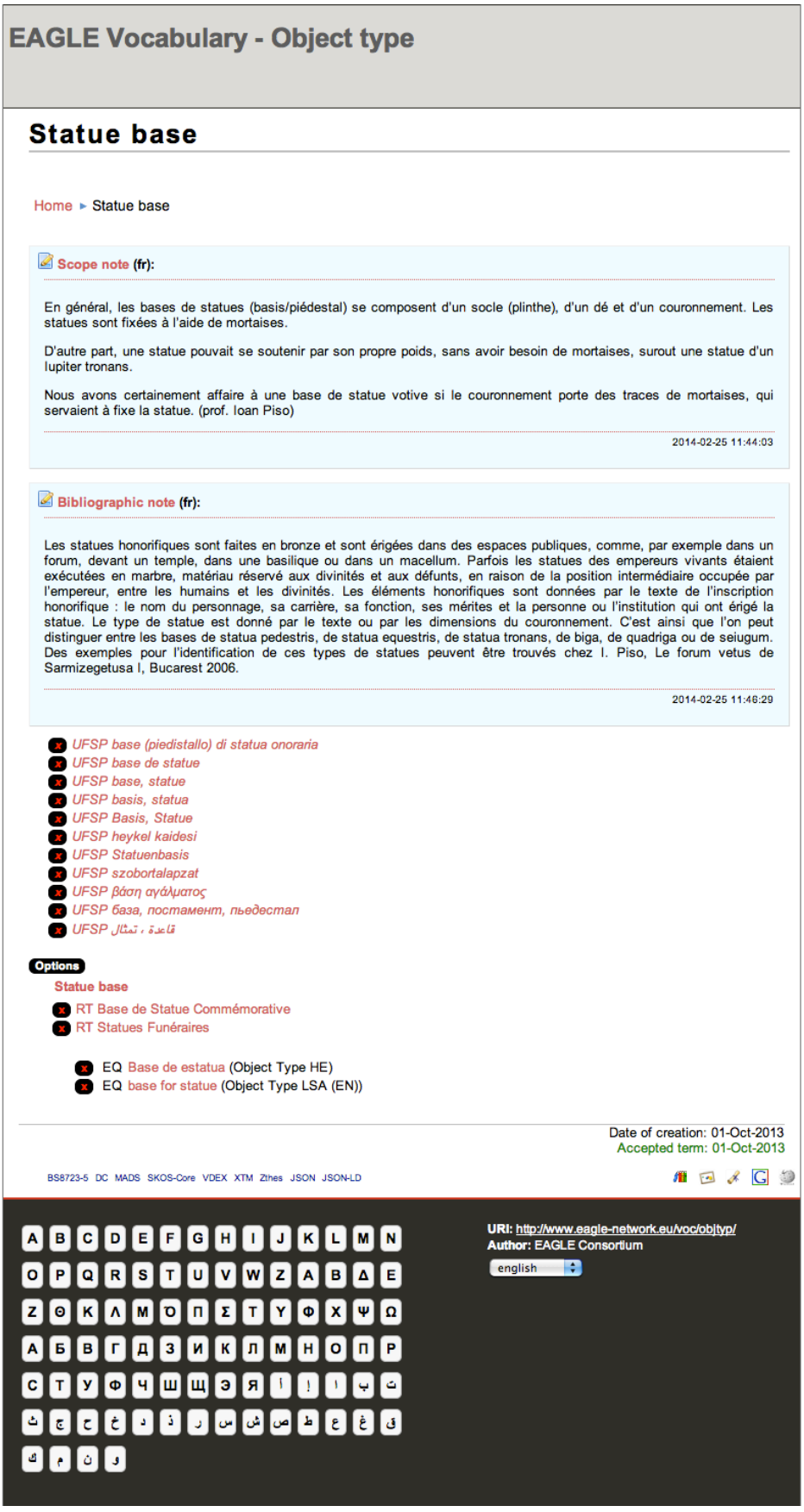
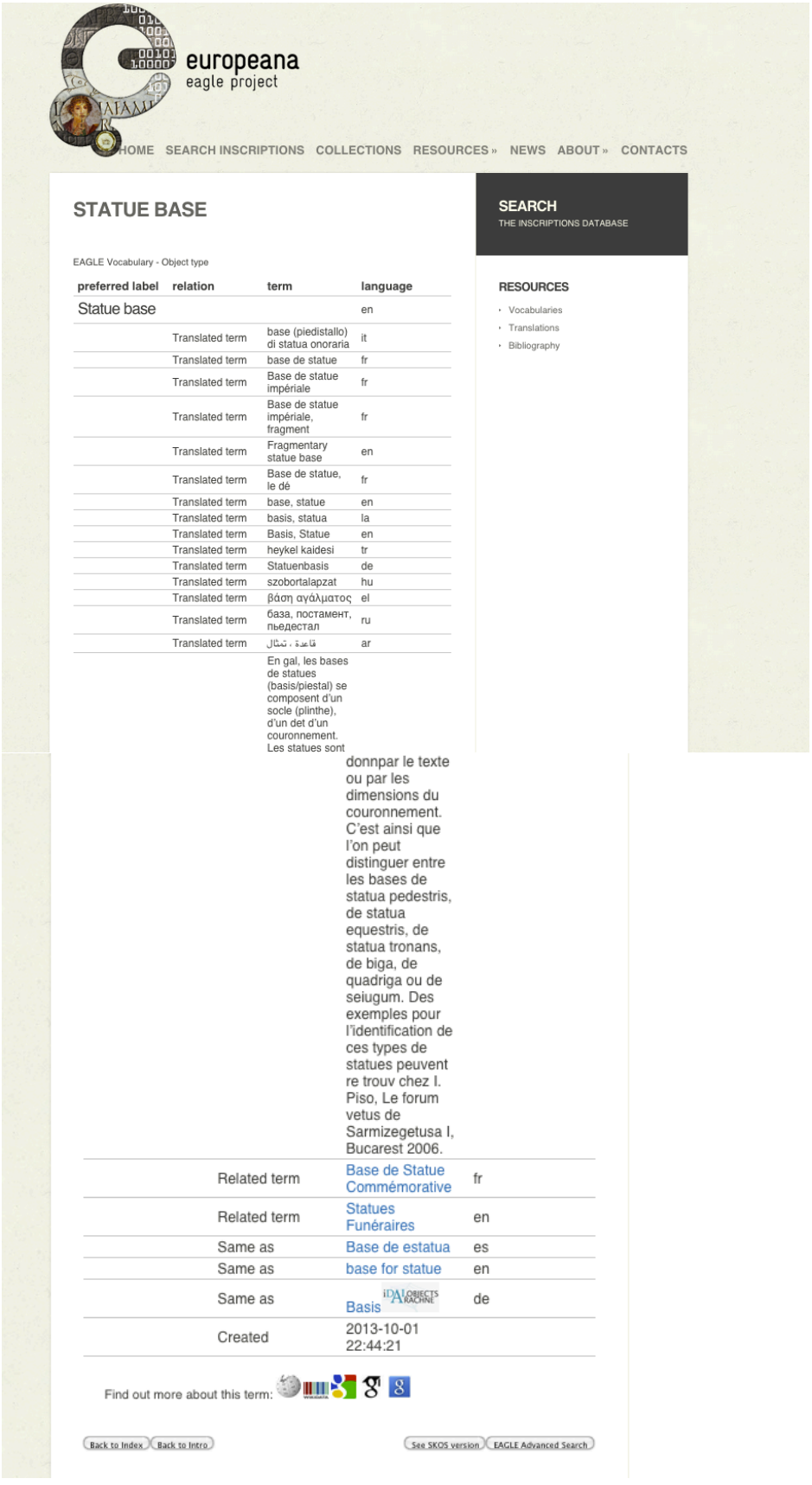


Figure 4.7: Statue Base in TemaTres



europeana
eagle project

HOME SEARCH INSCRIPTIONS COLLECTIONS RESOURCES » NEWS ABOUT » CONTACTS

STATUE BASE

SEARCH
THE INSCRIPTIONS DATABASE

RESOURCES
- Vocabularies
- Translations
- Bibliography

EAGLE Vocabulary - Object type

preferred label	relation	term	language
Statue base			en
	Translated term	base (piedistallo) di statua onoraria	it
	Translated term	base de statue	fr
	Translated term	Base de statue impériale	fr
	Translated term	Base de statue impériale, fragment	fr
	Translated term	Fragmentary statue base	en
	Translated term	Base de statue, le dé	fr
	Translated term	base, statue	en
	Translated term	basis, statua	la
	Translated term	Basis, Statue	en
	Translated term	heykel kaidesi	tr
	Translated term	Statuenbasis	de
	Translated term	szobortalapzat	hu
	Translated term	βάση αγάλματος	el
	Translated term	база, постамент, пьедестал	ru
	Translated term	قاعدة ، تمثال	ar
		En gal, les bases de statues (basis/piestal) se composent d'un socle (plinthe), d'un det d'un couronnement. Les statues sont donnpar le texte ou par les dimensions du couronnement. C'est ainsi que l'on peut distinguer entre les bases de statua pedestris, de statua equestris, de statua tronans, de biga, de quadriga ou de seiugum. Des exemples pour l'identification de ces types de statues peuvent re trouv chez I. Piso, Le forum vetus de Sarmizegetusa I, Bucarest 2006.	
	Related term	Base de Statue Commémorative	fr
	Related term	Statues Funéraires	en
	Same as	Base de estatua	es
	Same as	base for statue	en
	Same as	Basis	de
	Created	2013-10-01 22:44:21	

Find out more about this term:

Back to Index | Back to Intro | See SKOS version | EAGLE Advanced Search

Figure 4.8: Statue Base in in the new release

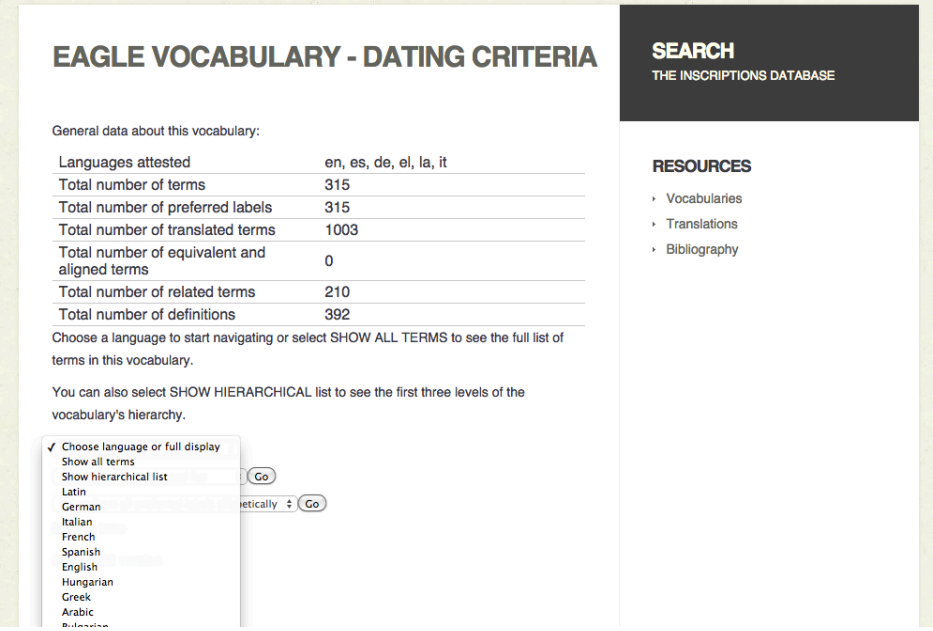


Figure 4.9: New visualization options for the dating criteria vocabulary

Figure 4.10: Hierarchy visualization for the dating criteria vocabulary

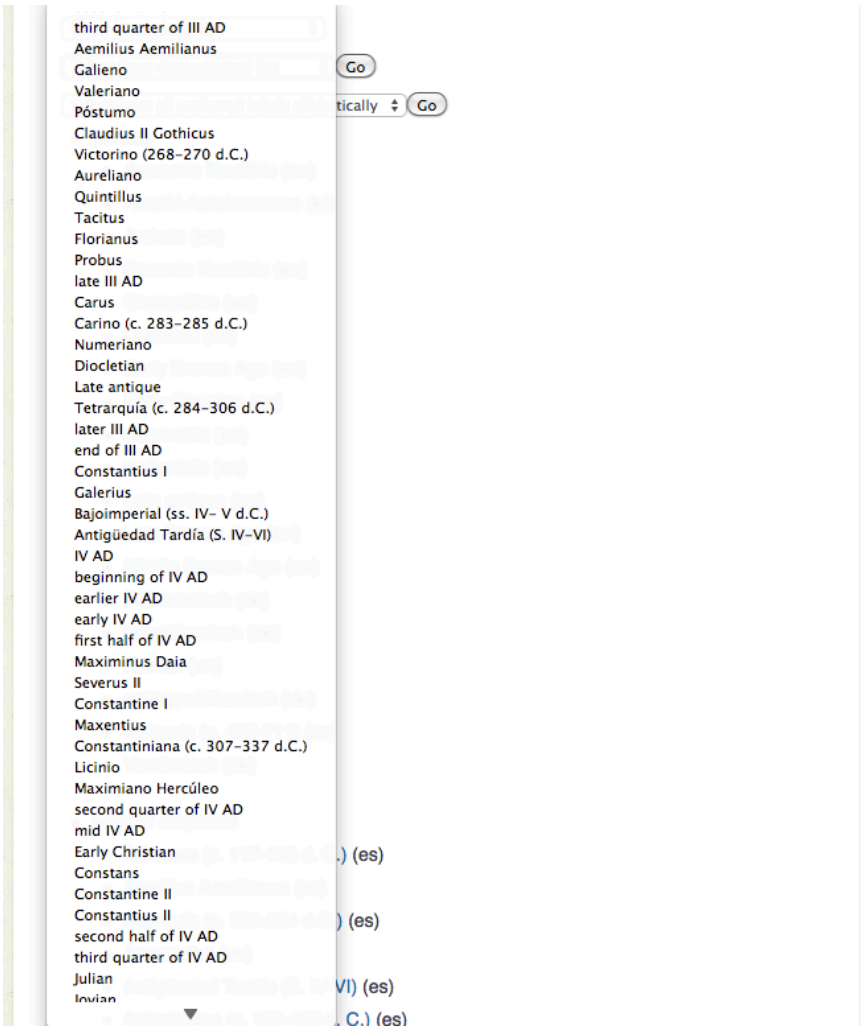


Figure 4.11: Chronological list of terms in the dating criteria vocabulary

Key Best Practices Summary for this chapter

BP 6: LOD Each vocabulary entry shall:

- **a** be related to existing resources
- **b** be a dereferenceable resource with an identifier in the form of an http url.
- **c** have links to with other entries
- **d** have links to external vocabularies
- **e** have links to external resources using the term consistently
- **f** be exposed and accessible for access, export, editing, reuse

BP 6.1: content negotiation Following LOD best practice we support content negotiation for the vocabularies

BP 7: Harmonization Minimize the creation and proliferation of vocabularies by harmonizing the existing ones.

BP 7.1: avoiding gerarchy and standardization Where possible we have chosen not to try and force a particular strucuture and to keep the list of term as simple as possible.

BP 8: Multilingualism Attested terms in any language shall be preferred values compared to one to one translations of terminologies as such.

BP 9: Open Software Use open software to allow iterability, high possibilities of interaction as well as flexibility to adapt to different needs.

HE 2: Vocabularies alignment Comparison of terms and definition is a task requiring constant activity, continuous interaction and networking.

HE 3: Definitions each vocabulary term should have definition(s), bibliography and examples.

HE 4: Relations each vocabulary term should be related in meaningful ways to its subordinate terms, broader terms, parallel terms and external representations.

HT 2: The EAGLE Vocabularies the TemaTres vocabularies developed provide a tool for harmonizing data without impact on local practices. Stakeholders have already demonstrated direct interest in this effort which has perspective comparable to those of geographical data.

HT 2.1: Further searches For each term a series of further searches is possible which enables the user to find out more about that term and related items. Also a google custom search dedicated to the vocabularies has been deployed.

HT 3: XSLT for harmonization of vocabularies id and URI allows to insert TemaTres URI in XML files, when the value in use is present in those, seamlessly.

4.2 Type of Inscription

4.2.1 Key Data

In this vocabulary are described all terms which concern the type of text written on a stone according to its function and aim as a text. (Manzella 1987, pp. 110-111)

URI	`http://www.eagle-network.eu/voc/typeins.html`
Language	it, es, fr, la, de, en, hu, el, ar
Preferred Terms	213
Non preferred terms	189
Terms with definition	81
Aligned terms	36
Terms	215

Table 4.1: EAGLE Vocabulary - Type of Inscription / EAGLE Consortium

4.3 Object Type

4.3.1 Key Data

The classification of objects on which an inscription can be written is extremely complex in many cases. For example, within the scope of the inscription which would fall under Statue Base, there are few cases in which what survives is an 'Inscribed plaque' – an inscription once affixed to a base (made of a different material) – and in other cases our 'statue base' in a wider context would more naturally be defined in another way: in particular 'Triumphal arch' (say, for the Arch of Constantine, that once supported statues), 'City gate' (like those of Rome, with statues of the emperors), or 'Columnar monument' (e.g. Trajan's Column – which of course supported a statue).

URI	`http://www.eagle-network.eu/voc/objtyp.html`
Language	en, de, la, fr, es, hu, el, ru, ar, it, tr, he
Date of creation	10/12/2013
Scope note	28
Terms	1209
Translated terms	1196
Aligned terms	610
Definitions	57

Table 4.2: EAGLE Vocabulary - Object type / EAGLE Consortium

4.4 Material

4.4.1 Key Data

This hierarchic Vocabulary relies on the definitions of the Simplified Petrography in Salzburg. Another classification can be found inCIL Material Glossar. No Definitions can therefore be found here. Most of the Alternative Terms in Bulgarian, Turkish, Croatian, Suomi, Arabic and Modern Greek come from the Epidoc Consortium efforts, as described in the general introduction (4).

The vocabulary is hierarchically designed to allow for both precise and generic definition, as well as for several types of uncertainty. In many cases, for example for precious/exotic materials (onyx, gold, etc.) precision is possible and desirable but for most banal materials only basic distinctions will generally be possible. In many cases scholars in epigraphy do not seek to decided whether a base was of Proconnesian or Thasian marble, nor to decide exactly what kind of stone the local material was (generally a form of limestone), in such cases, the best practice is to allow for different levels and choices.

A result of this approach has also been that this vocabulary should also be a complete list of possible material on which an inscription from antiquity can be found.

Many problems are involved with this kind of classification as well. ”Arcilla” for example is not in the same ”class” as ”terracotta, inscribed brick, etc”, because it pertains to the material in which the inscription was written or carved in, that is, like ”stone, metal or wood”. There are also specific terms which could be assimilated but are not, like ”Lumaquela”, the local name for the stone in which most of the inscriptions from Tarraco were carved in.

URI	`http://www.eagle-network.eu/voc/material.html`
Language	en, tr, fi, hu, la, es, fr, it, de, el, bg, ar, hr, he
Date of creation	11/10/2013
Terms	381
Translated terms	605
Aligned terms	101

Table 4.3: EAGLE Vocabulary - Material / EAGLE Consortium

4.5 Writing

4.5.1 Key Data

URI	http://www.eagle-network.eu/voc/writing.html
Language	la, it, fr, en, de, hr, el, ar, es, hu, sp, ru, bg
Date of creation	01/08/2013
Terms	67
Definitions	44
Non preferred Terms	147
Aligned Terms	22

Table 4.4: EAGLE Vocabulary - Writing / EAGLE Consortium

This vocabulary collects all writing techniques and types as attested in the EAGLE BPN. A major classification problem is faced on this respect by epigraphists. In same cases the distinction between the method used to produced a text is used as a principle for a definition, in some other the name of the writing technique is based on the result obtained, no matter of the method that was used. In particular *Punctim*, *litt. Scariphatis* and *litt. Eminentibus* are defintions of this second type. A first distinction that should be made when defining a writing technique should be whether

- the text has been obtained by adding material.
- the text has been obtained subtracting material.
- the text has been obtained modifying the material (as in cases of impression with a matrix which does not always involve subtraction of material)

Subtractive Techniques include texts which have been engraved, chiselled or graffiti. In these cases part of the surface of the material is removed to obtain the inscription. The following techniques among the one in the vocabulary would fit this definition:

- scalpro
- stylo
- digito
- terebro
- caelo
- a stecca

Additional Techniques include all those instances in which a different material is added on a base surface of the inscription. The following techniques among the one in the vocabulary would fit this definition:

- penicillo
- calamo
- carbone
- crustis
- tessellis
- litt. applicitis ex lamina vel ex brattea

There are indeed techniques which mix this two macro-types:

- litt. alveolatis
- litt. Ageminatis

Impression Techniques involve soft or still soft materials and do not fit in any of the previous two categories. material might be removed but this is not the rule and does not by itself determine the inscription. The same type can be used for those inscriptions which are obtained by fusion or liquefaction of a material in a mold, as this is a particular type of modification of a material. The following techniques among the one in the vocabulary would fit this definition:

- impression techniques
  - signaculo
  - typo
  - signaculo candente (impression with heath)
- techniques with a matrix
  - ex forma

There are writing techniques which do not fit any of the previous categories suggested.

Classification should also take into consideration the material on which they are applied which can for this classification purpose be hard or soft. In this cases in fact the technique is different not because of the method used but because of the particular status of the support in the moment in which the inscription was produced.

There are also inscriptions on ceramic which use the technique ”scalpro” with capital letters imitating those of lapidary inscriptions. Scalpro would tehrefore be fine for these rare examples as well.

hard	**soft**
atramento	Crustis
caelo	litteris ex forma
carbone	Signaculo
litteris ageminatis	Stilo
litteris alveolatis	Tessellis
litteris applicitis	
litteris eminentis	
litteris ex cartha aurea	
litteris scariphatis	
pictura	
punctim	
scalpro	
terebro	
typo	

4.6 Decoration

4.6.1 Key Data

Terms in this vocabulary describe any decoration surrounding an inscription. *Hederae*, *crosses*, and *chi-rho* monograms are probably among the most frequent decorations, but there are a number of classifiable features of decoration which support and are indeed vital in some cases, to the understanding of the inscribed text. Many artistic calssification could be suggested, but for the purposes of epigraphic research we offer a flat list of features which can be found surrounding an inscription.

URI	`http://www.eagle-network.eu/voc/decor.html`
Language	it, en, de, es, la, fr, hu
Date of creation	01/08/2013
Terms	626
Translated Terms	260
Aligned terms	279

Table 4.5: EAGLE Vocabulary - Decoration / EAGLE Consortium

4.7 State of Preservation

4.7.1 Key Data

State of preservation (which could probably be abbreviated to 'Condition') is a small vocabulary describing the general condition in which the inscription is at the moment of the latest inspection. In many cases we won't know if an inscription still exists.

URI	`http://www.eagle-network.eu/voc/statepreserv.html`
Language	en, it, de, la, es, fr
Date of creation	01/08/2013
Non preferred labels	6
Definitions	8
Terms	19

Table 4.6: EAGLE Vocabulary - State of preservation / EAGLE Consortium

4.8 Dating Criteria

4.8.1 Key Data

The common use among Members of the BPN is to insert, when this information is available two dates. A ‘Not before’ and ‘Not after’ dates. this is still a best practice and should definitely be preserved in any encoding. There are nevertheless divergences on historical periods and indeed in the way computations are done. In many cases only a generic date is possible. In this vocabulary we suggest a series of standardizations which equate some locution with a specific date interval, e.g. ”shortly after” with an interval of ca. 10 years. The WG has chosen to always explicitly suggest the corresponding TEI origDate element as a definition in this vocabulary. Within the TEI element it is also possible to express uncertainty and probability of dates.

URI	`http://www.eagle-network.eu/voc/dates.html`
Language	de, en, el, es, he, la
Date of creation	01/08/2013
Terms	330
Translated Terms	1036
Definitions	403

Table 4.7: EAGLE Vocabulary - Dating Criteria / EAGLE Consortium

Chapter 5

Harmonization of contents

5.1 Translations

The situation of these type of data was and still is that of a large series of very small, but high quality corpora. There are high numbers of translations of inscriptions out there, but very scattered and this has been the first real attempt to gather them all in one place. The EAGLE MediaWiki is, to our knowledge, the biggest existing database of translations of inscriptions as well as the first ever attempt to use Wikibase outside of Wikidata.

History and development

EAGLE BPN and Wikimedia Italia, subcontractor in the project have set up a Mediawiki (*The EAGLE Mediawiki*) which collects translations of inscriptions.

With Wikimedia Italia we have mapped some data from Content Providers of the EAGLE BPN, both original members and newly affiliated members. Some also still in the process of signing an affiliation agreement.

We uploaded automatically a subset of metadata including all available URI and text of the translations. We have added other known URIs to increase the ability to browse the Mediawiki.

Wikibase

After briefly considering *Semantic MediaWiki* as a way forward we have opted for *Wikibase*, although there where no precedents, as this would allow direct LOD export in RDF automatically, in line with other decisions within the EAGLE BPN and guidelines supported by Europeana.

Wikibase enabled us to import and export as well as to directly edit the Eagle Media Wiki as it is an open database online software and is also wiki-synthax free. To support individual addition to the database a tutorial has been developed.[1]

The following images are a sample record from the inscriptions of Brigetio and the main page of the wikibase.

Wikimedia Commons and EAGLE

Please, note the property "*Image(s)*": in this property a direct link to WC can be found. The photo from Wikimedia Commons is inserted and linked to the transla-

[1] See p.**??**. The EAGLE BPN choice has been supported by Wikidata with inclusion in `wikiba.se` website and mention in several wiki pages.

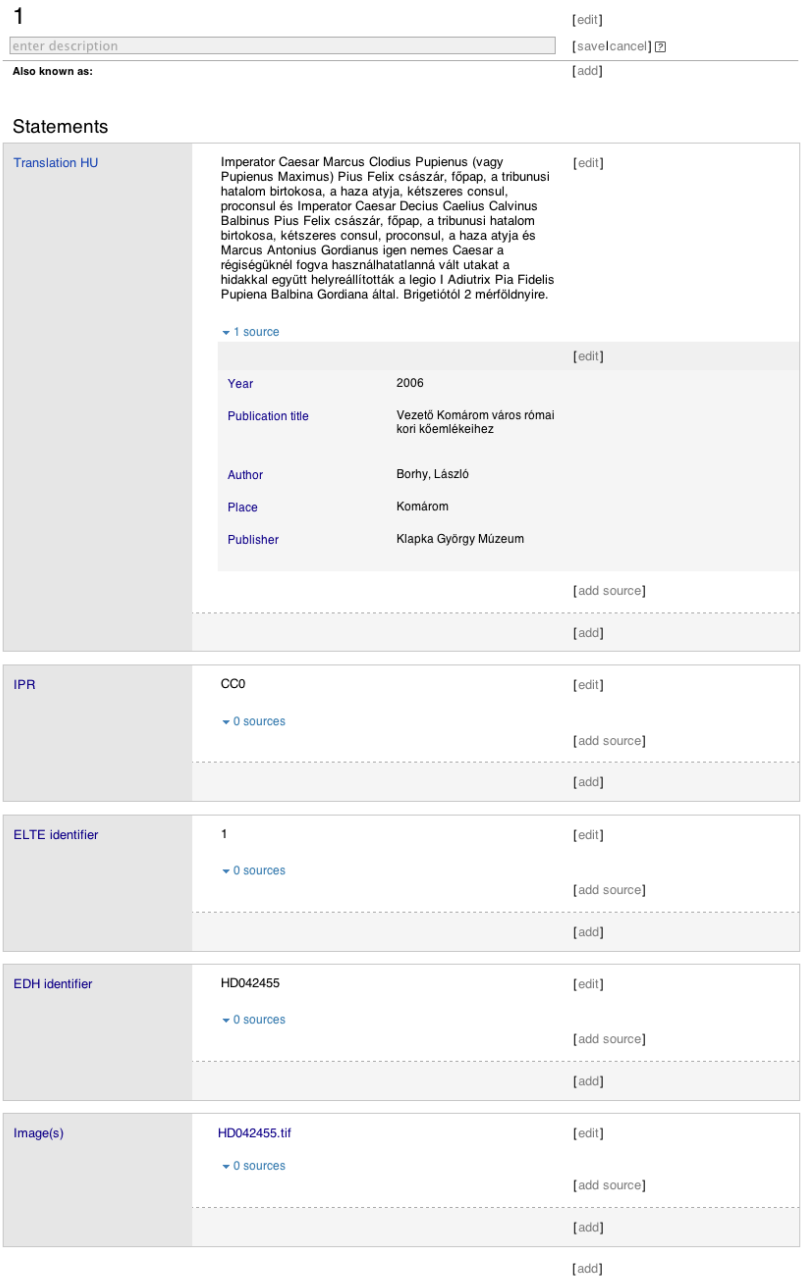
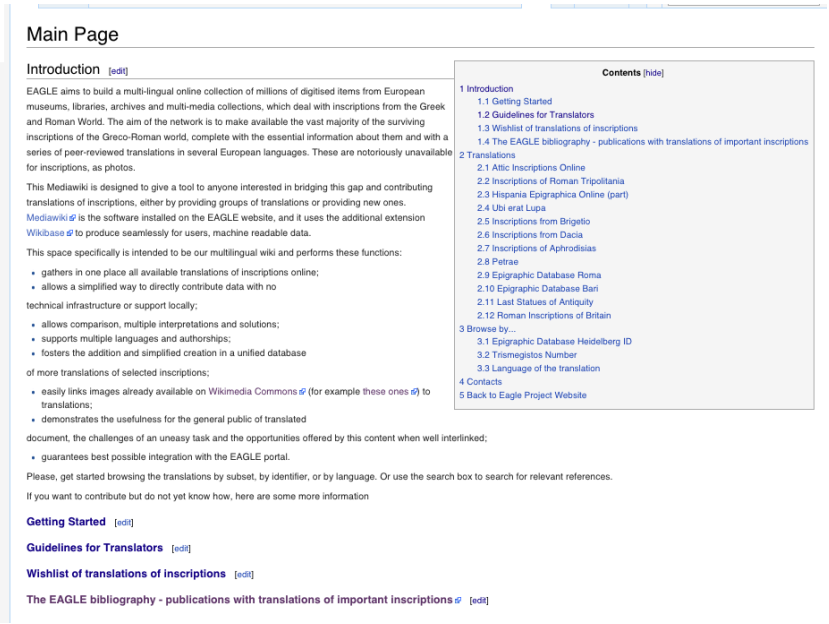


Figure 5.1: Brigetio 1

tions of the inscription as easily as by entering the name of the file. This has been a major benefit for us in the use of Wikibase, which really allow us to easily link


Main Page

Introduction [edit]

EAGLE aims to build a multi-lingual online collection of millions of digitised items from European museums, libraries, archives and multi-media collections, which deal with inscriptions from the Greek and Roman World. The aim of the network is to make available the vast majority of the surviving inscriptions of the Greco-Roman world, complete with the essential information about them and with a series of peer-reviewed translations in several European languages. These are notoriously unavailable for inscriptions, as photos.

This Mediawiki is designed to give a tool to anyone interested in bridging this gap and contributing translations of inscriptions, either by providing groups of translations or providing new ones. Mediawiki is the software installed on the EAGLE website, and it uses the additional extension Wikibase to produce seamlessly for users, machine readable data.

This space specifically is intended to be our multilingual wiki and performs these functions:

- gathers in one place all available translations of inscriptions online;
- allows a simplified way to directly contribute data with no

technical infrastructure or support locally;

- allows comparison, multiple interpretations and solutions;
- supports multiple languages and authorships;
- fosters the addition and simplified creation in a unified database

of more translations of selected inscriptions;

- easily links images already available on Wikimedia Commons (for example these ones) to translations;
- demonstrates the usefulness for the general public of translated

document, the challenges of an uneasy task and the opportunities offered by this content when well interlinked;

- guarantees best possible integration with the EAGLE portal.

Please, get started browsing the translations by subset, by identifier, or by language. Or use the search box to search for relevant references.

If you want to contribute but do not yet know how, here are some more information

Getting Started [edit]

Guidelines for Translators [edit]

Wishlist of translations of inscriptions [edit]

The EAGLE bibliography - publications with translations of important inscriptions [edit]

Contents [hide]

1 Introduction
- 1.1 Getting Started
- 1.2 Guidelines for Translators
- 1.3 Wishlist of translations of inscriptions
- 1.4 The EAGLE bibliography - publications with translations of important inscriptions

2 Translations
- 2.1 Attic Inscriptions Online
- 2.2 Inscriptions of Roman Tripolitania
- 2.3 Hispania Epigraphica Online (part)
- 2.4 Ubi erat Lupa
- 2.5 Inscriptions from Brigetio
- 2.6 Inscriptions from Dacia
- 2.7 Inscriptions of Aphrodisias
- 2.8 Petrae
- 2.9 Epigraphic Database Roma
- 2.10 Epigraphic Database Bari
- 2.11 Last Statues of Antiquity
- 2.12 Roman Inscriptions of Britain

3 Browse by...
- 3.1 Epigraphic Database Heidelberg ID
- 3.2 Trismegistos Number
- 3.3 Language of the translation

4 Contacts

5 Back to Eagle Project Website

Figure 5.2: The Eagle Media Wiki

resources in WC, in the EAGLE BPN databases and in other projects. The last but not less important benefit of this is the ability to enter several translations (in the same or in different languages).

We did face some problems, related to the nature of the contents: not always the same inscription is considered as such by all editors, neither there is always agreement on integration of the text and consequently on how the translation should be.

The flexibility of the Wikibase software has been vital to all the stages of the process. The enterprise has been highly successful till now. We hope it will carry on at an healthy rate with more and more data being linked and entered. [2]

5.2 Bibliography with Zotero

Structured bibliographies following the best practices set from librabrians for share and reuse of metadata on publication, not only promote correct behaviours for data reuse and citation but can help in deduplication process and alignment as well as to further research. Among Content Providers of the EAGLE BPN several have a structured bibliography and a decision was made to make it publicly available in

[2] Further details can be found in the proceedings of the EAGLE international Conference (D2.4) and will be given in D2.5, on the multilingual mediwiki .

the most reusable way.

The Working Group suggested as a best practice one that would meet the needs of the user community: this group of users needs quick and precise reference to publications for direct use in articles and other publications developed with popular WYSIWYG word processors. Moreover the bibliography needs to be a research tool reliable and updated.

The Working Group suggested therefore the use of a group in Zotero as best practice, compared to other bibliography management systems online which would be to pay for or not enough collaborative. Moreover Zotero is the one bibliography management use most used by practitioners.

Main bibliographies where imported from EDB, DAI, EDH, and HEpOnline using a bibtex format, and a de-duplication process for each bibliographic references of a total of more than 27.000 records began. The Group has been created and is here:[3]

EAGLE Epigraphic Bibliography

Some important benefits of this software for the users can be:

- free and open-source license;
- easy import and export;
- one-click metadata and full text imported from major sites storing bibliographic information such as (JSTOR, Google Scholar, AMAZON etc.);
- facilitated editing and online and offline usability and syncing;
- integration with the most popular word processors. Drag and drop functions;
- export and print in many different bibliographic styles both citations and bibliography.

On the other side also the bibliographic databases benefit from this effort:

- more exposition of the bibliographic database;
- integration and enrichment with other databases;
- multiplication of export formats (bibtex, bookmarks, mods, rdf, xml, etc);
- easier organization of contents of the bibliography with tags and folders;
- multiple input and quick additions of items.

Especially on import, references can be moved from the database in many formats (see:How to import in Zotero), whereas some tools as EndNote need a multiple steps procedure (see: EndNote in Zotero).

This task will also allow for large communities of users in academic institutions worldwide to be engaged in the effort of bibliography harmonization in the domain of ancient epigraphy. There has been a long felt behind need for a structured

[3]Another group in Zotero which provides an example of best practice is the EpiDig Group. there are also an existing Epigraphy Group and Linked Open Data Zotero Bibliography.

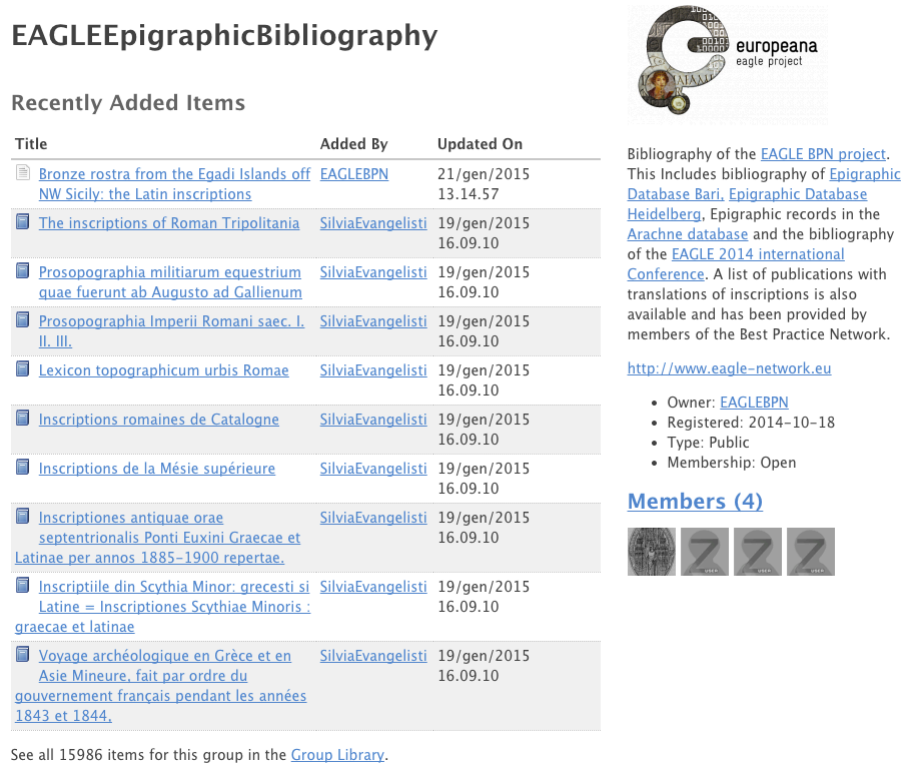


Figure 5.3: Eagle Epigraphic Bibliography (Zotero Group)

common reference bibliography available online. With the use of the resources offered by Zotero in future projects and developments it might also be possible to archive relations between bibliographic entries and therefore highly enrich the quality of research possibilities.

Zotero pros and cons

Nevertheless, some unsolved issues must also be noted:

stable keys/handle/uri each item receives a URL when it is imported into Zotero; based on our reading of the documentation in the user forum, this url is supposed to be stable.[4] However, the url is not intended to be a unique and universal key for each publication (e.g. for a specific book), but of course a unique reference in that personal or group bibliography. So for example this is the same record in a personal bibliography and in the general group is different

- itemKey/FXJQAB3Z
- itemKey/7CUVAQ9J

morover, these urls are recreated if you happen to update the same record

[4] https://forums.zotero.org/discussion/13090/record-id/

with a batch upload (see Zotero - record id).
The workaround mentioned in the forums is the same other bibliography management systems use: insert your own disambiguated and unique handle/key in an ”extra” field. In this way the info contained in the record is preserved (by users, not by Zotero).

cardinality of fields with urls to inscriptions This is not possible in one record. On one side this would be a ”citation VS reference” problem (working definitions can be found below). On the other there is also a question of granularity. Zotero allows for a relation option, which would also be bibliographically meaningful (see: Relations in Zotero). This means, that if we make one record for each inscription as an online publication we can then have one record for the publication it is in and as many relations as we want. Here again there is a problem although: these relations are not visible online neither exported in any export format, but they can be visualized if one has downloaded the plugin or standalone and in reports (which can be customized, se reports). A relation cannot be created in the web service. Faolan C.P., Zotero Administrator, in a recent support forum post wrote:

> I don’t believe this was possible at the time of the first post. It now is, and is generally planned but I’m not certain when it will land.[5]

Eagle status

The situation of the EAGLE consortium is such that some databases have a citation only system (e.g.EDR), some have both citations and a reference bibliography, but not connected among them (e.g. EDH), some only have a reference bibliography (e.g. DAI). Moreover, on the citation level there is on one side the greatest heterogeneity, on the other there is an almost complete effort of standardization done by Trismegistos in the task for de-duplication of items.

This complex situation is indeed of some benefit when trying to find a best practice to harmonize this kind of data, because the cases at play represent most of the possible situations and a solution to be found for this will be with more probability useful for future stakeholders.

Work-definitions

The following are initial guidelines for this sub-task in content curation and harmonization.

By **Citations** we mean abbreviated forms of bibliographical information to give direct reference, identification and basic information on an item. They are not therefore to be confused with identifiers which might be taken from a bibliographic repertoire as l’Année épigraphique.

By **References** we mean the list of full bibliographic records to which usually citation point, which can be identified by a unique number from a bibliographic repertoire.

[5] https://forums.zotero.org/discussion/22534/zotero-web-interface-showing-related/

Best practice for a commons Epigraphic bibliography therefore include sets of suggestions for both citation and references and efforts of harmonization from the existing situations which can also allow for further and consistent expansion and long term sustainability.

- On the side of citations the aim is to have a stable reference, possibly a URI, to be used as a pointer and identifier, which is different or parallel to the reference to a bibliographic entry in a repertoire which might be less or more inclusive.
- On the side of references, completeness is the key guideline, together with structure and relations according to the Working Group, as the cases in which bibliography is entered as string data turn out not to be accessible for reuse
- bibliographic repertoires identifiers and other inscriptions identifiers should be used as tags to allow recognition of contents and searches based on the inscription contained in a specific publication rather than on its title. This is possible because EDH has done for years a systematic work about this.
- inscriptions should be linked to the structured online bibliography rather than being a separate resource

These are points which perfectly fit a LOD approach which can use RDF in XML for example.

The Zotero Group Epigraphic Bibliography with URI for each item and each tag in proper and evident relation to other items, is what provider should use, seamlessly to the local format of citations.

When importing the data a large amount of work has been done prior to the final import to start work on definition and enlargement to harmonize contents of the bibtex input files. The large amount of data did not allow for perfection but we expect large user's contribution. Each item as a stable URI embedded in the metadata to avoid the potential ambiguity of the Zotero assigned URIs and all reference to bibliographic repertoires are also given in the form of retrievable tags. the only distinction kept is that the corresponding identifier of an inscription is given as such also. Further standardization of formats of tags had to be carried out but could not be thoroughly done due to the large amount of data.

currently the group bibliography contains

- Epigraphic Bibliography Heidelberg;
- Epigraphic Bibliography Bari;
- DAI bibliographic entries for inscriptions;
- References to Corpora according to use in EDR;
- Hispania Epigraphica bibliography;
- The Proceedings of the first EAGLE international Conference' bibliography;
- A list of publications containing translations of inscriptions.

Also an initial tool integrated with the upconversion tool as been developed to link this bibliography online to the records of inscriptions exported to the aggregator. This assignes all relevant URIs to a give string entry into a record on the basis of string matching on a structured unique bibliographic entry. These include the corresponding record and a direct link to all possible tags uris provided from Zotero so directly enabling the enlargement of a bibliographic search via this tool.[6]

<bibl>
AE 2006, 0995.
<ptr type=”zoterotaglink” target=”https://www.zotero.org/groups/
eagleepigraphicbibliography/items/tag/AE 2006 0995”/>
<ptr type=”edh” target=”http://edh-www.adw.uni-heidelberg.de/edh/
bibliographie/B013026”/>
<ptr type=”zotero” target=”http://zotero.org/users/2138134/items/
CMPG9F3B http://zotero.org/users/2138134/items/UWPQSVC4”/>
</bibl>

[6]See the GIT repository.

Key Best Practices Summary for this chapter

BP2: Interaction and Networking Joining GLAM and other major efforts in the Wikimedia context allows for internal and external interaction as well as for a very high impact on users and facilitation to engage them.

BP 6: LOD Also for translations and bibliography ”LOD ready” tools should be used not only for external possible connections to happen but also for more specific internal tasks as linking structured bibliographical data and building a common full citation-reference system.

BP 8: Multilingualism Dealing with translations the more languages available, the better. Also to have more translations in one language, makes them comparable and consequently the quality of data much higher.

BP 10: Usability Wikibase offers the easiest available way to write structured RDF triples with no need to know anything about RDF or triples.

BP 11: Shared online bibliography Although several published and curated bibliography exist, best results for interaction and engagement of users in data enrichment can be reached only with online tools.

HE 5: collecting translations scattered data need to be first of all collected from sources:

- **a** published (in a Zotero library)
- **b** unpublished
- **c** already online

HE 5.1: Adding new translations It is now possible to add new translations with a variety of methods which guarantee control but also usability.

HT 4: Eagle Wikibase this tool allows for perfect integration with Wikimedia Foundation projects: Wikimedia Commons, Wikidata, Wikipedia as the main examples.

HT 5: Open Zotero Group bibliography A Zotero open group bibliography allows the comunity to edit and contribute to a common highly curated resource and to use this huge reference for their own research work.

Chapter 6

Standards and tools for data harmonization

6.1 Epidoc as Harmonization format for inscriptions

The main harmonization guideline adopted is to align the XML format of data provided for aggregation and ingestion, to the TEI specification EpiDoc. This well established standard, broadly used in many high quality projects,[1] allows for a very easy alignment, for the production of an XML file compliant with international standards and for high flexibility for integration of the vocabularies and places gazetteer in use.[2]

> EpiDoc is an international, collaborative effort that provides guidelines and tools for encoding scholarly and educational editions of ancient documents. It uses a subset of the Text Encoding Initiative's standard for the representation of texts in digital form and was developed initially for the publication of digital editions of ancient inscriptions (e.g. Inscriptions of Aphrodisias, Vindolanda Tablets). Its domain has expanded to include the publication of papyri and manuscripts (e.g. Papyri.info). It addresses not only the transcription and editorial treatment of texts themselves, but also the history and materiality of

[1] See: https://sourceforge.net/p/epidoc/wiki/Home/ and http://wiki.digitalclassicist.org/Category:Projects. Among these especially King's College projects as the Inscriptions of Aphrodisias and the Inscriptions of Roman Tripolitania.

[2] Please, note that this is one of two standards adopted in the Eagle Metadata Model. Rivero Ruiz, Andrea, and Vassallo 2013, Part II, 2.1

the objects on which the texts appear (i.e., manuscripts, monuments, tablets, papyri, and other text-bearing objects).[3]

In the Gentle Introduction all steps taken towards this are explained and clarify the historical reasons for which this standard among others is the one preferable for epigraphers.

> Over the last century, epigraphers have wrestled with the issues involved in representing non-verbal information within their written texts. Until the end of the 19th century publishers could be expected to produce a facsimile of the text, but this became decreasingly common, and publishers did not demonstrate a parallel willingness to provide a full photographic record of every text. The conventions which have been painfully developed to indicate missing text, abbreviations, etc. have been more or less generally agreed since the 1930s and overlap, to some extent, with those used in papyrology and palaeography. All epigraphers have had to deal with the issues involved in moving this to an electronic environment - for example, finding a font which will permit underdotting; but most of us have now adjusted to these new constraints. [...] The need for agreed standards is not limited to epigraphy. Since 1987 an international consortium of scholars principally in the humanities has been working together to develop and refine a set of guidelines for describing the structure and content of documents. The results of this endeavour have produced an encoding language, realized in XML and described by the name of the group - TEI, the Text Encoding Initiatve.[4]

6.2 Upconversion to Epidoc

In cooperation with UHEI tools have been developed to support the alignment and harmonization of data from content providers to international standards for what concerns digital editions of inscriptions.

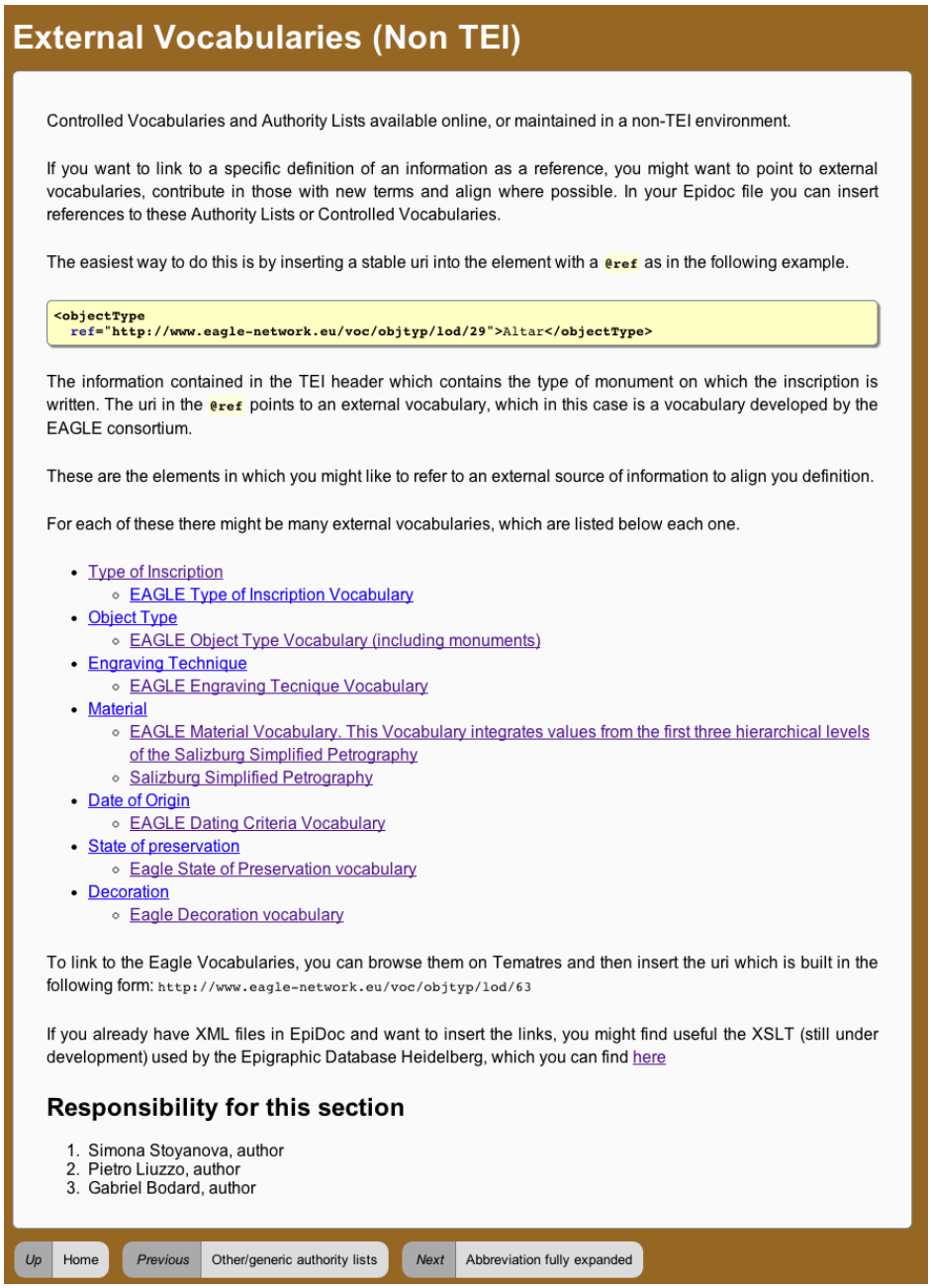
This is all freely available also for potential new partners in the Git repository.

All references in the .xsl are absolute, in order to guarantee reusability.

XSLT for EPIDOC

Given the template described in Part III, and ANNEX II of the EAGLE Metadata Schema (Rivero Ruiz, Andrea, and Vassallo 2013), the following XSLT convert from string epigraphic texts in marked up TEI-EPIDOC XML, following the EpiDoc guidelines (Elliott et al. 2007).

[3] https://sourceforge.net/p/epidoc/wiki/Home/
[4] http://www.stoa.org/epidoc/gl/dev/intro-eps.html



External Vocabularies (Non TEI)

Controlled Vocabularies and Authority Lists available online, or maintained in a non-TEI environment.

If you want to link to a specific definition of an information as a reference, you might want to point to external vocabularies, contribute in those with new terms and align where possible. In your Epidoc file you can insert references to these Authority Lists or Controlled Vocabularies.

The easiest way to do this is by inserting a stable uri into the element with a `@ref` as in the following example.

```
<objectType
  ref="http://www.eagle-network.eu/voc/objtyp/lod/29">Altar</objectType>
```

The information contained in the TEI header which contains the type of monument on which the inscription is written. The uri in the `@ref` points to an external vocabulary, which in this case is a vocabulary developed by the EAGLE consortium.

These are the elements in which you might like to refer to an external source of information to align you definition.

For each of these there might be many external vocabularies, which are listed below each one.

- Type of Inscription
  - EAGLE Type of Inscription Vocabulary
- Object Type
  - EAGLE Object Type Vocabulary (including monuments)
- Engraving Technique
  - EAGLE Engraving Tecnique Vocabulary
- Material
  - EAGLE Material Vocabulary. This Vocabulary integrates values from the first three hierarchical levels of the Salizburg Simplified Petrography
  - Salizburg Simplified Petrography
- Date of Origin
  - EAGLE Dating Criteria Vocabulary
- State of preservation
  - Eagle State of Preservation vocabulary
- Decoration
  - Eagle Decoration vocabulary

To link to the Eagle Vocabularies, you can browse them on Tematres and then insert the uri which is built in the following form: `http://www.eagle-network.eu/voc/objtyp/lod/63`

If you already have XML files in EpiDoc and want to insert the links, you might find useful the XSLT (still under development) used by the Epigraphic Database Heidelberg, which you can find here

Responsibility for this section

1. Simona Stoyanova, author
2. Pietro Liuzzo, author
3. Gabriel Bodard, author

Up Home | Previous Other/generic authority lists | Next Abbreviation fully expanded

Figure 6.1: Epidoc Guidelines

These are a series of XSLT based on Chetc.txt (by Hugh Cayless, Elli Mylonas, Gabriel Bodard and Tom Elliott) and further support from the Epidoc Collaborative

(especially from Gabriel Bodard) which:[5]

1. allows to convert epigraphic texts with various encoding and conventions from string to Epidoc markup and valid against the The EpiDoc RelaxNG schema.
2. Populate appropriate elements with available common URI from the vocabularies:
   - `http://www.eagle-network.eu/voc/typeins.html` (4.2)
   - `http://www.eagle-network.eu/voc/objtyp.html` (4.3)
   - `http://www.eagle-network.eu/voc/material.html` (4.4)
   - `http://www.eagle-network.eu/voc/writing.html` (4.5)
   - `http://www.eagle-network.eu/voc/decor.html` (4.6)
   - `http://www.eagle-network.eu/voc/statepreserv.html` (4.7)
   - `http://www.eagle-network.eu/voc/dates.html` (4.8)
3. build a title for contents which do not have it, using descriptive data.

6.2.1 Step 1

One set of XSLT to serve different needs

Each project uses different conventions and therefore the regular expressions used to match particular situations are different. The two currently developed initial XSL are `edh-epidoc.xsl` and `he-epidoc.xsl`.

This are the same, unless for the fact that call appropriated `.xsl`

Each XSLT is applicable to any file is in the main directory, if it is a valid xml utf-8 file. The string text needs to be

- escaped if it is not
- normalized if it needs so

The process of mark up of the string text in div[@type=”edition”] is accomplished in several steps to guarantee consistency and precision.

The `textstructure.xsl` looks for marker of different sections and tokenize them to apply the same .xsl to each section of the text which needs to be contained by an <ab> element. If there is only one part it applies following instructions to that only.

Each section of text is then processed by the `brackets.xsl`. Normalizing brackets is important for the following steps and splits individual semantic values.

[5]The following xsl will be maintained aligned with the Epidoc guidelines at all stages guaranteeing an effort free alignment to these international conventions for partners who can continue to apply local conventions for editing.

The notation [ort 3], which would mean that a supplied text is followed by a gap of three letters, is divided into [ort][3].

The normalized string which results from this process is then passed to the `upconversion.xsl` which works using the a specific operation to search for regular expressions patterns (`xsl:analyze-string`) and substitute them with correct xml elements.

The following is an example in which the pattern <E=F> is matched and substituted

```
<xsl:analyze-string select="." regex="(&lt;)(.)=(.)(&gt;)">
  <xsl:matching-substring>
   <choice>
     <corr>
       <xsl:value-of select="regex-group(2)"/>
     </corr>
     <sic>
       <xsl:value-of select="regex-group(3)"/>
     </sic>
   </choice>
</xsl:matching-substring>
```

The result will be

```
<choice><corr>E</corr><sic>F</sic></choice>
```

The result of this template is then passed to a further template which gives consistent numbers (`insertnumbers.xsl`). Empty lines do not need to have numbers, so Xpath is used to evaluate where to put a 0 as value of the @n in the <lb> element.

Starting from this

------] / e[t?] Q(---) Bl(a)e[sus?] / contub/ernalis / eius / d(e) s(uo) l(ibens) l(aetus) d(edit)

The result of this processes is then the following

```
<ab>
<lb n="0"/><gap reason="lost" extent="unknown" unit="line"/>
<lb n="1"/>e<supplied reason="lost" cert="low">t</supplied>
<abbr>Q</abbr>
Bl<expan><ex>a</ex><abbr>e</abbr></expan><supplied
    reason="lost" cert="low">sus</supplied>
<lb n="2"/>contub
<lb break="no" n="3"/>ernalis
<lb n="4"/>eius
<lb n="5"/><expan><abbr>d</abbr><ex>e</ex></expan>
<expan><abbr>s</abbr><ex>uo</ex></expan>
<expan><abbr>l</abbr><ex>ibens</ex></expan>
<expan><abbr>l</abbr><ex>aetus</ex></expan>
<expan><abbr>d</abbr><ex>edit</ex></expan>
</ab>
```

6.2.2 Step 2

On the following elements

```
<objectType>Stele</objectType>

<material>Marmor</material>

<term>Weihinschrift</term>

<origDate
  notBefore-custom="0001"
  notAfter-custom="0100"
  datingMethod="http://en.wikipedia.org/
  wiki/Julian_calendar">I cent</origDate>

<rs type="execution">engraviert</rs>

<rs type="statpreserv">incomplete</rs>

<rs type="decoration">hedera</rs>
```

populate the metadata with the vocabulary

A series of xsl, one for each vocabulary related to that element is called to match the content of the element with the vocabulary entry into an export of the *TemaTres* vocabulary stored on git and regularly updated.

Each entry with non preferred and target vocabulary entries looks like the following

```
<skos:Concept rdf:about="http://www.eagle-network.eu/voc/typeins/lod/73">
      <skos:prefLabel xml:lang="en">Akklamation</skos:prefLabel>
      <skos:altLabel xml:lang="en">Acclamation</skos:altLabel>
      <skos:altLabel xml:lang="en">Aclamación</skos:altLabel>
      <skos:scopeNote xml:lang="de"> Ausrufe; Anrufungen.
      Wahlslogans, Glkwsche, Sinnsprhe. Nicht
         bei Defixio oder Anrufungen an Gott (s.Gebet). </skos:scopeNote>
      <skos:historyNote xml:lang="en">Examples: vivas;
      spes in Deo salvo episcopo Marciano
         HD008112 HD018228</skos:historyNote>
      <skos:inScheme rdf:resource="http://www.eagle-network.eu/voc/typeins/"/>
      <skos:related rdf:resource="http://www.eagle-network.eu/voc/typeins/lod/75"/>
      <skos:related rdf:resource="http://www.eagle-network.eu/voc/typeins/lod/74"/>
      <skos:exactMatch>
         <skos:Concept rdf:about="http://www.eagle-network.eu/voc/typeins/lod/243">
            <skos:prefLabel xml:lang="it">Acclamatio</skos:prefLabel>
         </skos:Concept>
      </skos:exactMatch>
      <dct:created>2013-08-01 12:27:53</dct:created>
      <dct:modified>2013-08-20 14:17:18</dct:modified>
   </skos:Concept>
```

so the xsl looks for matchings in `skos:prefLabel` and in `skos:altLabel` and then reports in a @ref the URI of the parent `skos:Concept`, to give the following result:

<objectType ref="http://www.eagle-network.eu/voc/objtyp/lod/250">Stele</objectType>
<material ref="http://www.eagle-network.eu/voc/material/lod/48">Marmor</material>
<term ref="http://www.eagle-network.eu/voc/typeins/lod/80">Weihinschrift</term>
<origDate
notBefore-custom="0001"
notAfter-custom="0100"
datingMethod="http://en.wikipedia.org/wiki/Julian_calendar"
period="http://www.eagle-network.eu/voc/dates/lod/92">I cent</origDate>
<rs type="execution" ref="http://www.eagle-network.eu/voc/writing/lod/3">engraviert</rs>
<rs type="statpreserv" ref="http://www.eagle-network.eu/voc/statepreserv/lod/5">incomplete</rs>
<rs type="decoration" ref="http://www.eagle-network.eu/voc/decor/lod/155">hedera</rs>

6.2.3 Step 3

To populate the element <title> with meaningful information the `maketitle.xsl` checks the content of the file and looks for Type of Inscription and Object type. If both information are present the form of the title will be “inscription on object” if only one of the two is present that either of them will be used. It might happen that none is know. In that case the xsl looks for the CIL or AE bibliographic reference. Only in the unlikely event that there is none of this information the title “Inscription” is given.

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?xml-model
   href="http://www.stoa.org/epidoc/schema/latest/tei-epidoc.rng" schematypens="http://relaxng.org/ns/structure/1.0"
 ?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:space="preserve" xml:lang="de" xml:base="ex-epidoctemplate.xml">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Weihinschrift auf Stele</title>
      </titleStmt>
      <publicationStmt>
        <authority>Epigraphische Datenbank Heidelberg</authority>
        <idno type="URI">
          http://edh-www.adw.uni-heidelberg.de/edh/inschrift/HD067013
        </idno>
        <idno type="URI"/>
        <availability>
          <licence target="http://creativecommons.org/licenses/by-nc-sa/3.0/">
            This file is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.
          </licence>
        </availability>
      </publicationStmt>
      <sourceDesc>
        <msDesc>
          <msIdentifier>
            <country>
              <placeName type="modern" ref="http://www.geonames.org/countries/si/">Slowenien</placeName>
            </country>
            <region>
              <placeName type="modern" ref="http://www.geonames.org/">Podravska statistična regija</placeName>
            </region>
            <settlement>
              <placeName ref="http://www.geonames.org/3190249">Hajdina</placeName>
            </settlement>
            <repository ref="www.trismegistos.org/">
              Graz, Landesmus. Joanneum, Lapidarium Schloß Eggenberg
            </repository>
            <collection/>
            <idno/>
          </msIdentifier>
          <physDesc>
            <objectDesc>
              <supportDesc>
                <support>
                  <objectType ref="http://www.eagle-network.eu/voc/objtyp/lod/250">Stele</objectType>
                  <material ref="http://www.eagle-network.eu/voc/material/lod/48">Marmor</material>
                  <p/>
                  <dimensions unit="cm">
                    <height>-62</height>
                    <width>-30</width>
                    <depth>12.5</depth>
                  </dimensions>
                  <rs type="decoration"/>
                </support>
              </supportDesc>
              <layoutDesc>
                <layout>
```


Figure 6.2: Eagle Epidoc export with harmonization XSLs

6.2.4 Simplified Workflow

A simplified workflow for data curation is a necessary best practice to guarantee that technical knowledge is not an obstacle to the production of high data quality.

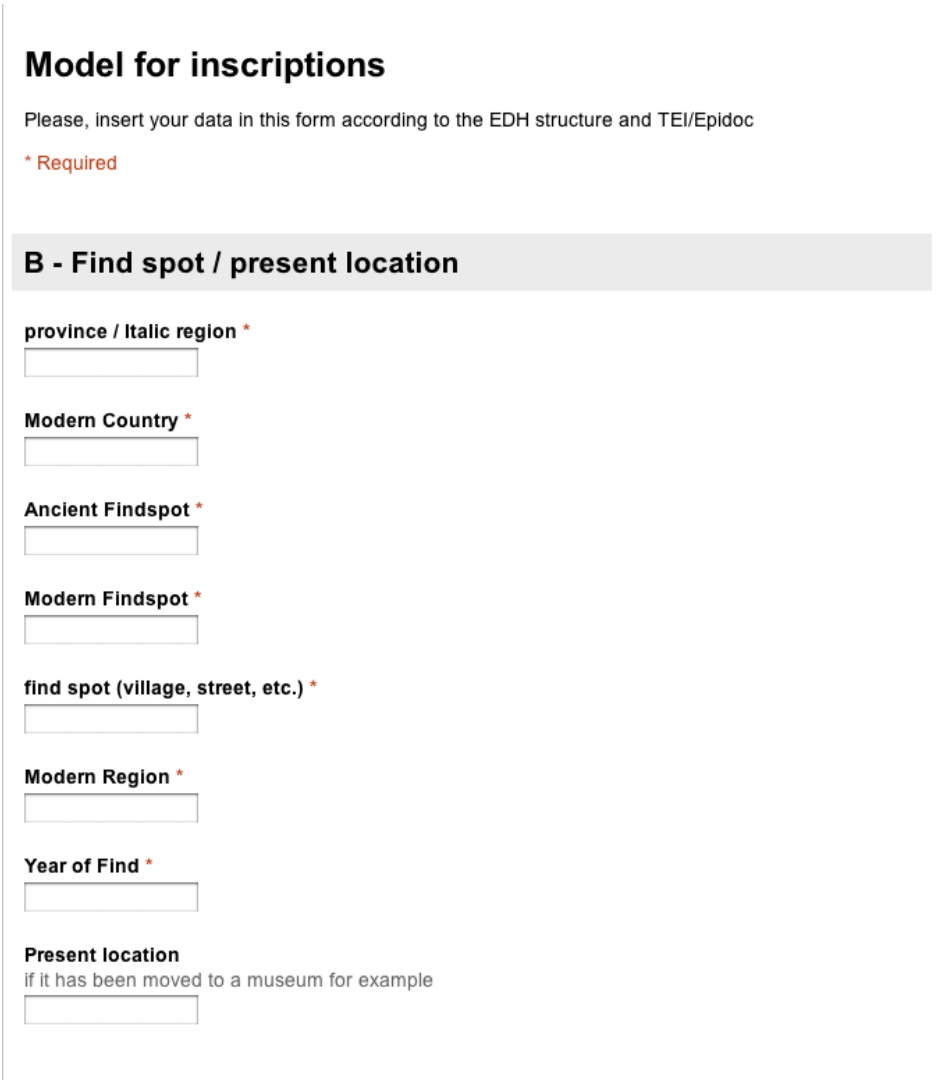


Figure 6.3: Epidoc data entry from google Drive

We have tested few options and decided within the WG to use well known software and develop transformation support. Data is entered in a google form that returns results in HTML. An xsl transformation which takes into account all the previous ones is set in place to transform those data (inserted according to strict principles) into the same EpiDoc format as from the exports and according to the Eagle Metadata Model Specification.

Key Best Practices Summary for this chapter

BP 2: Interaction and Networking Constant and useful cooperation has to be carried out with communities both digital and non digital.

BP 12: Raw data online All code and reference data should be available as raw data or via API.

BP 12.1: Stability and sustainability Making data standard and accesible for reuse makes it really sustainable on the long term

BP 13: Automatization and curation No export should be delivered unchecked: exporting in different format should be an occasion for further consistency check towards cleanness of data.

BP 14: Semantically meaningful data No random flat XML should be produced if it is possible to produce tools that elaborate results according to meaningful standards.

BP 15: Openly accessible XSL XSL and generally useful code should always be openly accessible for reuse, for example in GitHub, together with updated versions of content related code, as vocabularies for reuse and data exploitation.

HE 6: Conversion and Crosswalking tools should be elaborated so that bidirectional transformations are possible.

HE 6.1: Specialization and Tailoring While part of the code are common each content provider has an adapted verion of it. While it would be advisable to parametrize the code rather than having several copises, it is better for use with different engines and systems to have separate code.

HE 7: Simplification of workflow for external stakeholders it is essential that workflow is simplified, both by using well known tools and to guarantee that the user experience is satisfactory and not frustrating.

Chapter 7

Harmonization and interaction with international projects

7.1 Parnerships and interactions with international projects

As stated one of the key best practices for content harmonization is to seek maximum interaction with partner projects.[1]

7.1.1 Epidoc Community

The interaction with the Epidoc Community has been continuous for the development of the metadata model which is based on Epidoc as initial content to the establishment of vocabularies which are now contained also in the Epidoc Guidelines[2].

7.1.2 Perseus

The partnership with the Perseus project involves the development of interaction tools between the Eagle Media wiki of translation and Perseids.

The Perseids project is opening up the Perseus texts for user corrections, contributions and annotations, both in the classroom and

[1]The list of affiliated and cooperating partner attracted by EAGLE can be seen here: http://www.eagle-network.eu/about/get-involved/

[2]http://www.stoa.org/epidoc/gl/dev/idx-nontei.html

out. Work submitted through Perseids will make its way back into the Perseus library (slowly now, more quickly later).[3]

In fact the work done in Perseids project answers perfectly to the needs of the consortium who stated concern about the quality of content inserted by online unknown users and which in this way could be controlled while producing metadata already in XML.

7.1.3 Pelagios

Pelagios 3 is a Linked Open Data initiative, in which independent online resources are linked together by annotating common references to places (currently ancient Greek and Rome, but soon to be expanded) as defined in a URI-based gazetteer. These annotations can then be aggregated either by member projects or third parties for a variety of purposes, including search, contextualisation and visualisation.[4]

Working with the data provided by all the consortium to KULeuven for aggregation in Trismegistos Geo, to produce annotations in OA Data Model will make possible the tasks described in section 3.3.

To give an example of this annotation, things would look similar to this:

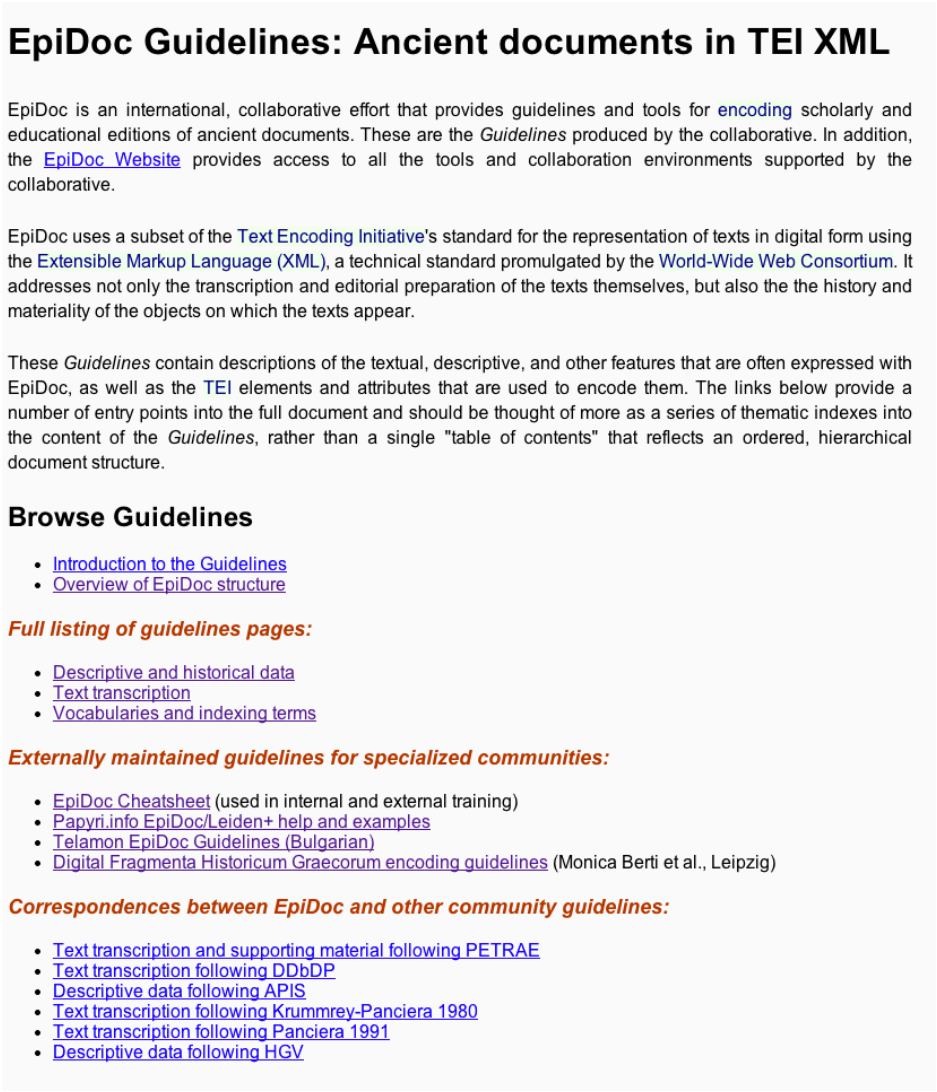
```
<http://edh-www.adw.uni-heidelberg.de/edh/inschrift/HD004306> a pelagios:AnnotatedThing ;
 dcterms:title ”Inschrift HD004306 (or any other suitable title for this item” ;
 <www.trismegistos.org/place/1163> a oa:Annotation ;
 oac:hasTarget <http://edh-www.adw.uni-heidelberg.de/edh/inschrift/HD004306> ;
 oa:annotatedAt ”2013-11-12”;
 oa:annotatedBy ”Trismegistos”;
 oa:hasBody [ cnt:chars ”Korinthos” ; a pelagios:Toponym ] ;
 oa:hasBody <http://pleiades.stoa.org/places/570182#this> ;
 oa:hasBody <http://www.trismegistos.org/place/1163> ;
```

7.1.4 Europeana API

In order to achieve the best possible harmonization we have been in touch with the team working at Europeana API to get as many inscriptions already in Europeana as possible to enrich and harmonize them all also at stages before the aggregation.

[3] See also http://sites.tufts.edu/perseusupdates/, quote from the Memorandum of Understanding between EAGLE and Perseids.

[4] http://pelagios-project.blogspot.de/p/about-pelagios.html



EpiDoc Guidelines: Ancient documents in TEI XML

EpiDoc is an international, collaborative effort that provides guidelines and tools for encoding scholarly and educational editions of ancient documents. These are the *Guidelines* produced by the collaborative. In addition, the EpiDoc Website provides access to all the tools and collaboration environments supported by the collaborative.

EpiDoc uses a subset of the Text Encoding Initiative's standard for the representation of texts in digital form using the Extensible Markup Language (XML), a technical standard promulgated by the World-Wide Web Consortium. It addresses not only the transcription and editorial preparation of the texts themselves, but also the the history and materiality of the objects on which the texts appear.

These *Guidelines* contain descriptions of the textual, descriptive, and other features that are often expressed with EpiDoc, as well as the TEI elements and attributes that are used to encode them. The links below provide a number of entry points into the full document and should be thought of more as a series of thematic indexes into the content of the *Guidelines*, rather than a single "table of contents" that reflects an ordered, hierarchical document structure.

Browse Guidelines

- Introduction to the Guidelines
- Overview of EpiDoc structure

Full listing of guidelines pages:

- Descriptive and historical data
- Text transcription
- Vocabularies and indexing terms

Externally maintained guidelines for specialized communities:

- EpiDoc Cheatsheet (used in internal and external training)
- Papyri.info EpiDoc/Leiden+ help and examples
- Telamon EpiDoc Guidelines (Bulgarian)
- Digital Fragmenta Historicum Graecorum encoding guidelines (Monica Berti et al., Leipzig)

Correspondences between EpiDoc and other community guidelines:

- Text transcription and supporting material following PETRAE
- Text transcription following DDbDP
- Descriptive data following APIS
- Text transcription following Krummrey-Panciera 1980
- Text transcription following Panciera 1991
- Descriptive data following HGV

Figure 7.1: Eagle Vocabularies in the Epidoc Guidelines

7.2 Content Curation and Harmonization in Wikimedia Commons

In order to give a complete effort of harmonization, already existing and editable sources of information as Wikimedia Commons can be curated. Adding relevant information and links to metadata of images is a best practice to perform content curation and encourage quality content, enriching those data with the ones coming from local experts which can give scientific input to the available content.

There are in facts many categorized and organized photographs in Wikimedia Commons with no link to a scholarly online edition of the text. It is a best practice according to the WG to exploit existing effort and provide contributions in order to harmonize it and correct it were possible to meet research quality standards.

There is in Wikimedia Commons this picture for example:
`https://commons.wikimedia.org/wiki/File:Mausoleu_de_Favara_2.jpg`
it corresponds to nr. 12143 in HEpOl

What we want to do is edit the metadata to make them like the one in this example: with a link to a database and correct categories.

https://commons.wikimedia.org/wiki/File:HD011512_1.tif

The summary section, might look like this:

```
{{Information
|Description=
{{ca|Mausoleu de Favara. Favara de Matarranya, Saragossa}}
{{en|Favara mausoleum. Favara, Saragossa}}
{{es|1=Mauseloe romano de Fábara (Zaragoza), correspondiente con las inscripciones romanas:
* Frontón: CIL II, 5851 = HEp. 8, 566a = AE 2000, 777a: D(is) M(anibus) / L(uci) Aemili Lupi
* Friso (puntos de anclaje de letras de bronce): HEp 08, 566 = AE 2000, 777b}}
|Source={{Own}}
|Date=2003-04
|Author= [[User:Baldiri|Baldiri]]
}}
```

In this case we simply add the url of the inscription in HEpOl, for example like this:

```
{{Information
|Description=
{{ca|Mausoleu de Favara. Favara de Matarranya, Saragossa}}
{{en|Favara mausoleum. Favara, Saragossa}}
{{es|1=Mauseloe romano de Fábara (Zaragoza), correspondiente con las inscripciones romanas:
* [http://eda-bea.es/pub/record... ...12143 Hispania Epigraphica Online]
* Frontón: CIL II, 5851 = HEp. 8, 566a = AE 2000, 777a: D(is) M(anibus) / L(uci) Aemili Lupi
* Friso (puntos de anclaje de letras de bronce): HEp 08, 566 = AE 2000, 777b}}
|Source={{Own}}
|Date=2003-04
|Author= [[User:Baldiri|Baldiri]]
}}
```

The EAGLE BPN will also take the effort of checking for obvious mistakes, but unless they are patent, they will not be corrected; the wrong information will simply

Size of this preview: 531 × 599 pixels. Other resolutions: 213 × 240 pixels | 425 × 480 pixels | 532 × 600 pixels | 680 × 768 pixels | 907 × 1,024 pixels | 2,435 × 2,748 pixels.

Language select: Show all

Original file (2,435 × 2,748 pixels, file size: 7.73 MB, MIME type: image/jpeg); ZoomViewer: flash/no flash

Summary [edit]

Description	**Deutsch:** Grabstein aus Brigetio, ca. 173 n. Chr. Heute im Ungarischen Nationalmuseum Budapest. Epigraphische Datenbank Heidelberg HD016792 **Inschrift:** Ae(lio) Septimo opt(ioni) leg(ionis) I / [Ad]i(utricis) desideratus est / [bello 3]aris qui vix(it) [; **Übersetzt:** „Dem Aelius Septimus, optio der legio I Adiutrix, vermisst im Krieg gegen die Naristen, der lebte..." Literatur: CIL 03, 04310 = CIL 03, 10969 = RIU 02, 00509 = AE 1962, 00221 = AE 1991, 01305
Date	30 September 2012, 15:40:03
Source	Own work
Author	Hartmann Linge

Figure 7.2: Metadata curation in Wikimedia Commons

be commented out. The provided link will indeed by itself provide those who will want to know further precise information.

The EAGLE BPN decided to take some effort to organize also categories as a best practice for content harmonization.

In the above example there are all the ones we want.

```
[[Category:CIL II 005851]]
[[Category:AE 2000, 0777a]]
[[Category:AE 2000, 0777b]]
```

There are many photos in Wikimedia Commons that have no such category, as in this example.

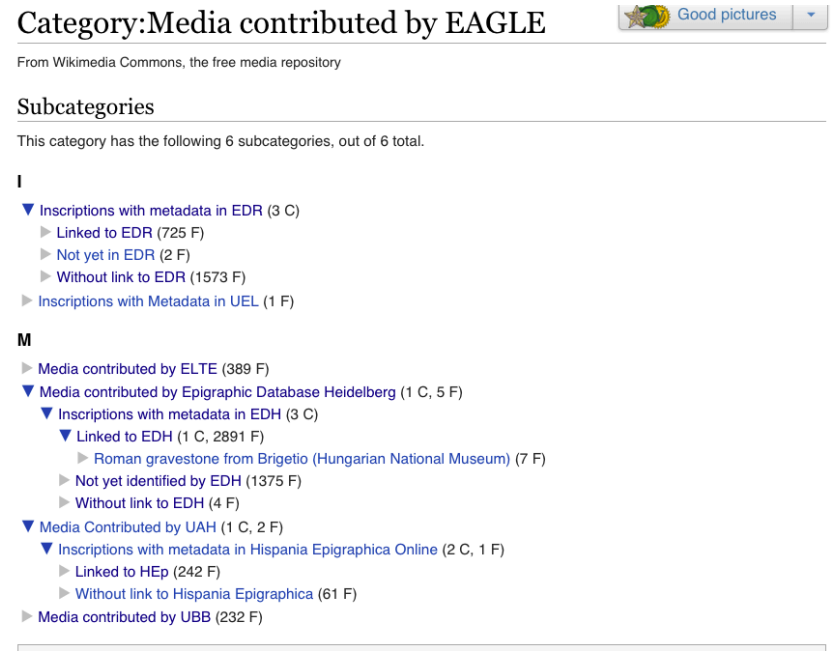


Figure 7.3: Media contributed by EAGLE - Categories

Figure 7.4: Categories for publication identifiers

We will add consistently categories which refer to CIL, AE and the EAGLE BPN member who contributed the file in case it is a file originally uploaded by one of the EAGLE Members.

If a required Category does not exist yet it will be created and properly placed under an existing Category.

For all newly uploaded images we apply also a Template specifit to the EAGLE project.

To upload a large amount of selected photos we have used the GLAMwiki Toolset, which is a tool specifically developed by Europeana for uploads of online images from databases and archives into Wikimedia Commons. Among the available tools this one has several advantages which relate to the specific characteristics of Galleries, Libraries, Archives and Museum. For example templates and categories which are meaningful to the institutions can be directly translated in the System Wikimedia Common uses without loss of information of any sort.

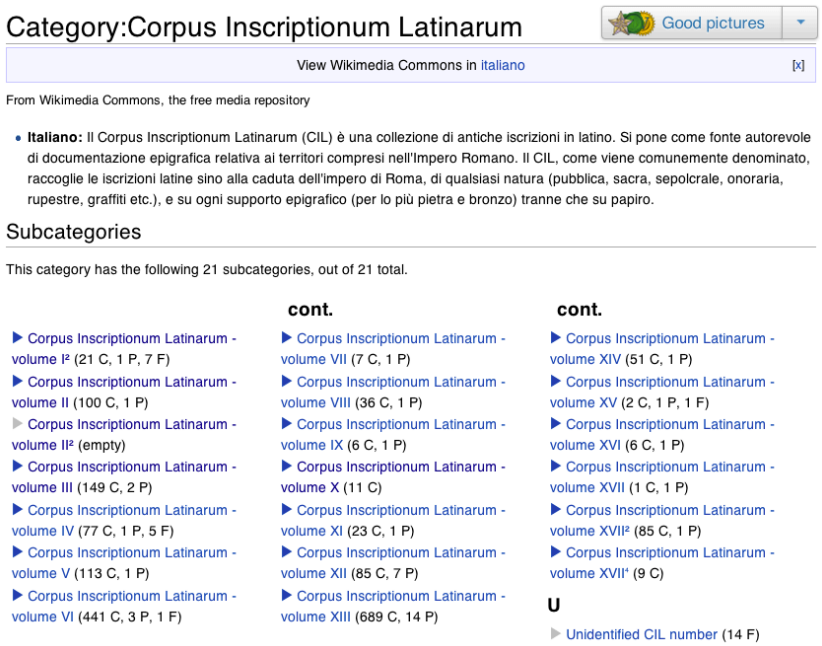


Figure 7.5: CIL Categories in Wikimedia Commons

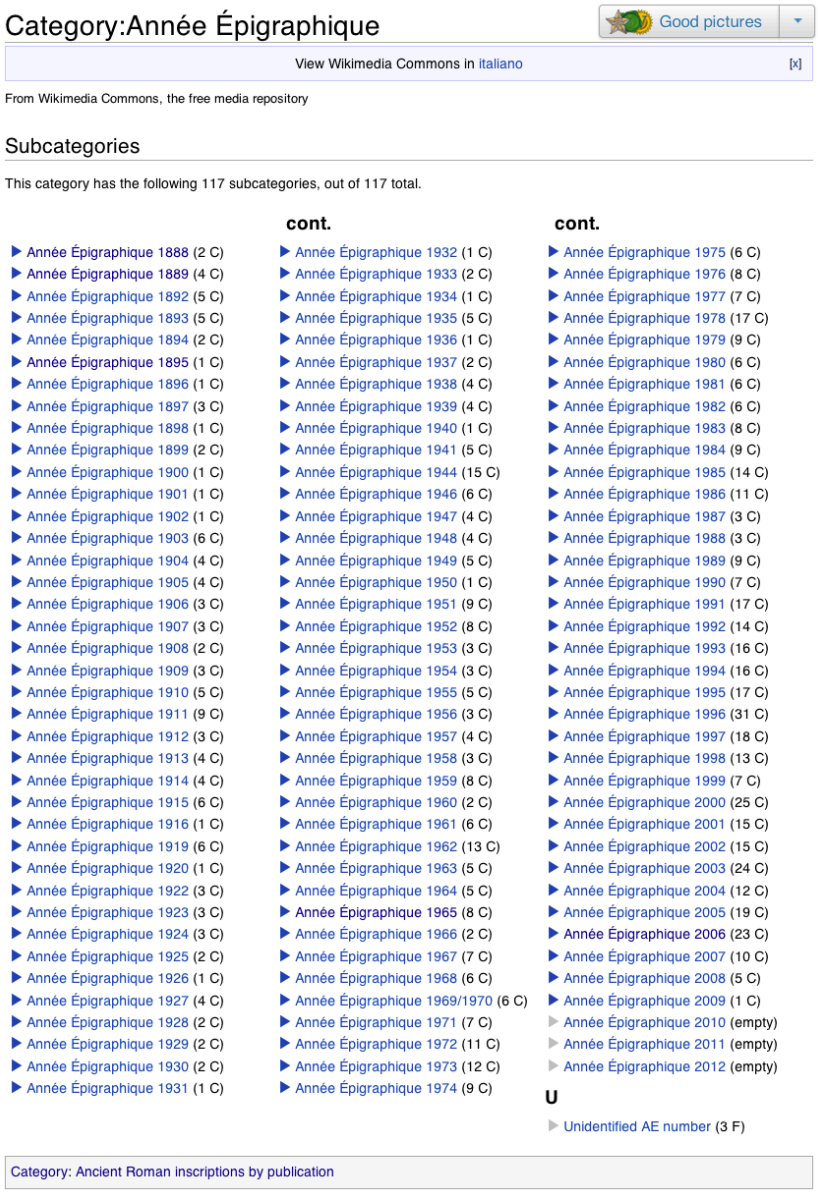
The process asks for a useful testing phase in which the user gets the assistance and guidance of more expert volunteers online and is then allowed to proceed in Commons directly.

For EAGLE we had to work on a XML dump to transform it into flat XML and rework all contents to feet them to the model.

Curation of bulk data Affected titles, format of potential categories, contents of metadata which needed to adhere to Wikimedia Commons standards and conventions, but the possibility to do this ahead of the import made the process a meaningful and dialogical exercise in understanding different scopes, users and aims of resources.

The following is the mapping of the flat XML source to the artwork template

```
{”institution”:[”institution”],
”location”:[”modernplace”],
”notes”:[”cp”],
”photo date”:[”photodate”],
”photo description”:[”id”],
”photographer”:[”author”],
”photo license”:[”licence”],
”place of creation”:[”placeofcreation”],
”references”:[”note”],
”source”:[”repository”],
”title”:[”artwork”],
”gwtoolset-title”:[”artwork”],
”gwtoolset-url-to-the-media-file”:[”file”]}
```



Category:Année Épigraphique

Good pictures

View Wikimedia Commons in italiano [x]

From Wikimedia Commons, the free media repository

Subcategories

This category has the following 117 subcategories, out of 117 total.

- Année Épigraphique 1888 (2 C)
- Année Épigraphique 1889 (4 C)
- Année Épigraphique 1892 (5 C)
- Année Épigraphique 1893 (5 C)
- Année Épigraphique 1894 (2 C)
- Année Épigraphique 1895 (1 C)
- Année Épigraphique 1896 (1 C)
- Année Épigraphique 1897 (3 C)
- Année Épigraphique 1898 (1 C)
- Année Épigraphique 1899 (2 C)
- Année Épigraphique 1900 (1 C)
- Année Épigraphique 1901 (1 C)
- Année Épigraphique 1902 (1 C)
- Année Épigraphique 1903 (6 C)
- Année Épigraphique 1904 (4 C)
- Année Épigraphique 1905 (4 C)
- Année Épigraphique 1906 (3 C)
- Année Épigraphique 1907 (3 C)
- Année Épigraphique 1908 (2 C)
- Année Épigraphique 1909 (3 C)
- Année Épigraphique 1910 (5 C)
- Année Épigraphique 1911 (9 C)
- Année Épigraphique 1912 (3 C)
- Année Épigraphique 1913 (4 C)
- Année Épigraphique 1914 (4 C)
- Année Épigraphique 1915 (6 C)
- Année Épigraphique 1916 (1 C)
- Année Épigraphique 1919 (6 C)
- Année Épigraphique 1920 (1 C)
- Année Épigraphique 1922 (3 C)
- Année Épigraphique 1923 (3 C)
- Année Épigraphique 1924 (3 C)
- Année Épigraphique 1925 (2 C)
- Année Épigraphique 1926 (1 C)
- Année Épigraphique 1927 (4 C)
- Année Épigraphique 1928 (2 C)
- Année Épigraphique 1929 (2 C)
- Année Épigraphique 1930 (2 C)
- Année Épigraphique 1931 (1 C)

cont.

- Année Épigraphique 1932 (1 C)
- Année Épigraphique 1933 (2 C)
- Année Épigraphique 1934 (1 C)
- Année Épigraphique 1935 (5 C)
- Année Épigraphique 1936 (1 C)
- Année Épigraphique 1937 (2 C)
- Année Épigraphique 1938 (4 C)
- Année Épigraphique 1939 (4 C)
- Année Épigraphique 1940 (1 C)
- Année Épigraphique 1941 (5 C)
- Année Épigraphique 1944 (15 C)
- Année Épigraphique 1946 (6 C)
- Année Épigraphique 1947 (4 C)
- Année Épigraphique 1948 (4 C)
- Année Épigraphique 1949 (5 C)
- Année Épigraphique 1950 (1 C)
- Année Épigraphique 1951 (9 C)
- Année Épigraphique 1952 (8 C)
- Année Épigraphique 1953 (3 C)
- Année Épigraphique 1954 (3 C)
- Année Épigraphique 1955 (5 C)
- Année Épigraphique 1956 (3 C)
- Année Épigraphique 1957 (4 C)
- Année Épigraphique 1958 (3 C)
- Année Épigraphique 1959 (8 C)
- Année Épigraphique 1960 (2 C)
- Année Épigraphique 1961 (6 C)
- Année Épigraphique 1962 (13 C)
- Année Épigraphique 1963 (5 C)
- Année Épigraphique 1964 (5 C)
- Année Épigraphique 1965 (8 C)
- Année Épigraphique 1966 (2 C)
- Année Épigraphique 1967 (7 C)
- Année Épigraphique 1968 (6 C)
- Année Épigraphique 1969/1970 (6 C)
- Année Épigraphique 1971 (7 C)
- Année Épigraphique 1972 (11 C)
- Année Épigraphique 1973 (12 C)
- Année Épigraphique 1974 (9 C)

cont.

- Année Épigraphique 1975 (6 C)
- Année Épigraphique 1976 (8 C)
- Année Épigraphique 1977 (7 C)
- Année Épigraphique 1978 (17 C)
- Année Épigraphique 1979 (9 C)
- Année Épigraphique 1980 (6 C)
- Année Épigraphique 1981 (6 C)
- Année Épigraphique 1982 (6 C)
- Année Épigraphique 1983 (8 C)
- Année Épigraphique 1984 (9 C)
- Année Épigraphique 1985 (14 C)
- Année Épigraphique 1986 (11 C)
- Année Épigraphique 1987 (3 C)
- Année Épigraphique 1988 (3 C)
- Année Épigraphique 1989 (9 C)
- Année Épigraphique 1990 (7 C)
- Année Épigraphique 1991 (17 C)
- Année Épigraphique 1992 (14 C)
- Année Épigraphique 1993 (16 C)
- Année Épigraphique 1994 (16 C)
- Année Épigraphique 1995 (17 C)
- Année Épigraphique 1996 (31 C)
- Année Épigraphique 1997 (18 C)
- Année Épigraphique 1998 (13 C)
- Année Épigraphique 1999 (7 C)
- Année Épigraphique 2000 (25 C)
- Année Épigraphique 2001 (15 C)
- Année Épigraphique 2002 (15 C)
- Année Épigraphique 2003 (24 C)
- Année Épigraphique 2004 (12 C)
- Année Épigraphique 2005 (19 C)
- Année Épigraphique 2006 (23 C)
- Année Épigraphique 2007 (10 C)
- Année Épigraphique 2008 (5 C)
- Année Épigraphique 2009 (1 C)
- Année Épigraphique 2010 (empty)
- Année Épigraphique 2011 (empty)
- Année Épigraphique 2012 (empty)

U

- Unidentified AE number (3 F)

Category: Ancient Roman inscriptions by publication

Figure 7.6: AE Categories in Wikimedia Commons

Results[5] can be seen also from the usage report applied from user Fae to the EAGLE Category: `https://commons.wikimedia.org/wiki/Category_talk:Media_contributed_by_EAGLE/reports`

[5] `https://commons.wikimedia.org/wiki/Category:Drawing_by_G%C3%A9za_Alf%C3%B6ldy`

Figure 7.7: The EAGLE Template

These have been a revelation for epigraphers as well which would have never expected a community of over 500 users to have contributed to photo of inscriptions.

Noticable among this is the presence of some of the largest files ever uploaded which belong to the Inscriptions from Montenegro[6] which have been contributed with Upload Wizard (`https://commons.wikimedia.org/wiki/Special:UploadWizard`) or Commonist (`https://commons.wikimedia.org/wiki/Commons:Commonist`) and with support given by the EAGLE project, by Dr. Olga Pelcer.[7]

EAGLE participation in the Wikimedia Development Task Force in order to produce on the basis of this and the activity with Wikibase on the EAGLE Media wiki, a deliverable submitted at the end of January with formal recommendations.

Recommendation 1: For every Europeana project, considering the possible benefits of a Wikimedia component should be default behaviour.

Recommendation 2: Help to facilitate local connections between GLAMs and Wikimedians.

Recommendation 3: Generate and distribute knowledge about Wikimedia culture among Europeana-staff.

Recommendation 4: Generate and distribute knowledge about Europeana and GLAMs among Wikimedia.

Recommendation 5: Europeana supports efforts in bringing pro-forma policy to partners regarding open licensing of both content and data.

Recommendation 6: Europeana to gather and distribute best practices about measuring impact on the Wikimedia platforms.

Recommendation 7: Make Wikidata a central element of the ´portal to platform´ strategy

Recommendation 8: Europeana should continue to invest in technology that improves the interoperability between GLAMs and Wikimedia platforms.

Recommendation 9: Joint applications for external funding opportunities

[6] E.g. `https://commons.wikimedia.org/wiki/File:Honorary_inscription_for_M._Iulius_Philippus.JPG`

[7] See also among the most edited files the lapidarium of the Ljubljana Catedral `https://commons.wikimedia.org/wiki/File:Ljubljana_-_Dolni%C4%8Darjev_lapidarij_(CIL_III_3853).jpg`

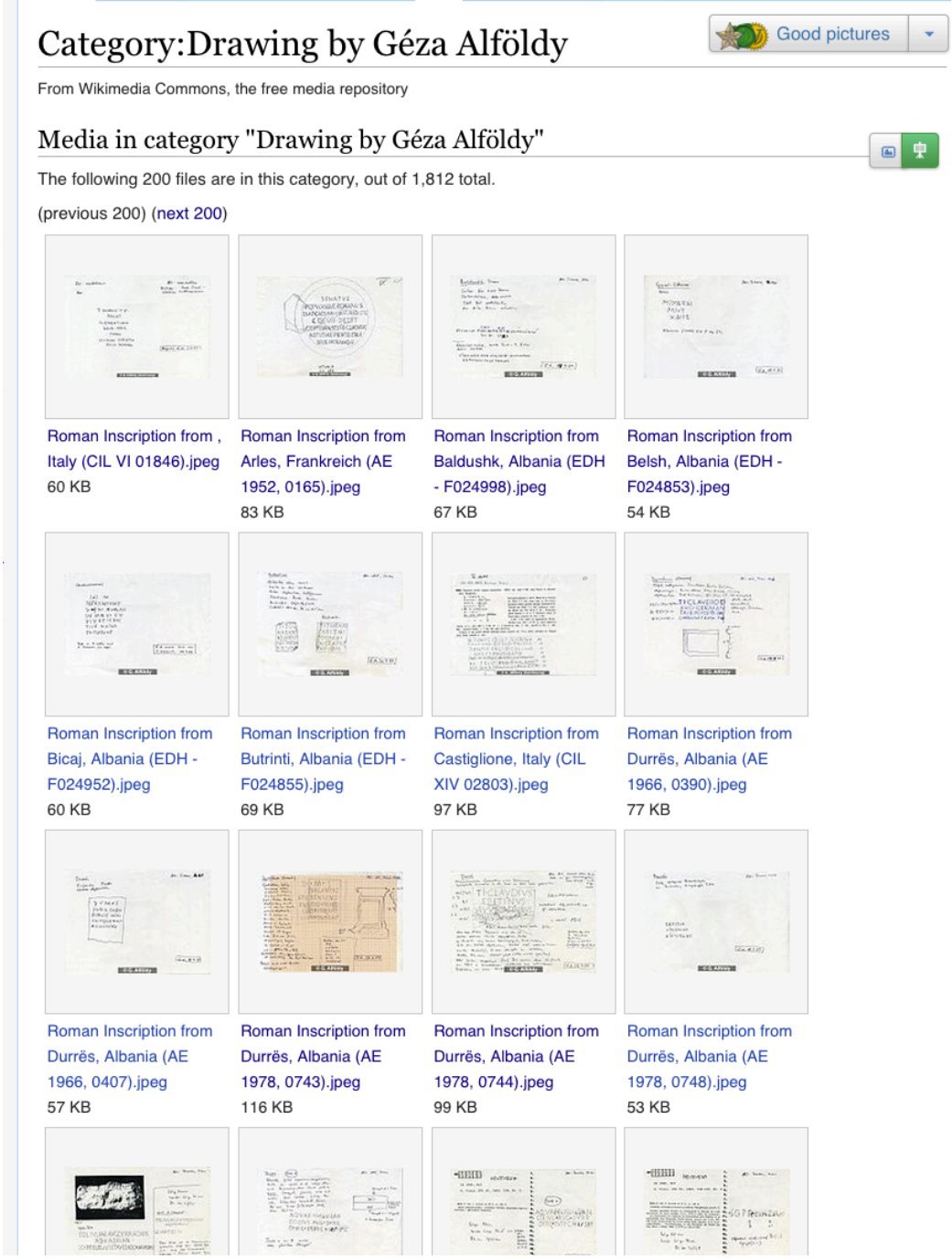


Figure 7.8: Drawings of Inscriptions By Prof. Dr. Géza Alföldy

Ancient Roman inscriptions in Italy (2,293) · Inscriptions from ancient Roma (1,904) · Drawing by Géza Alföldy (1,812) · CC-Zero (399) · Media lacking author information (305) · License migration redundant (220) · Inscriptions from ancient Latium et Campania (108) · Inscriptions from ancient Venetia et Histria (104) · Ancient Roman inscriptions in Syria (171) · Inscriptions from ancient Africa Proconsularis (145) · Ancient Roman inscriptions in Tunisia (138) · Inscriptions from ancient Syria (137) · Ancient Roman inscriptions in Algeria (96) · Ancient Roman inscriptions in Museo Nacional de Arte Romano (59) · Inscriptions from ancient Etruria (41) · Media with locations (71) · Ancient Roman inscriptions in the Land of Valencia (41) · Ancient Roman inscriptions in Germany (73) · Dis Manibus inscriptions (56) · Inscriptions from ancient Numidia (69) · Inscriptions from ancient Arabia (69) · Inscriptions from ancient Macedonia (69) · Inscriptions from ancient Aemilia (33) · Ancient Roman inscriptions in Albania (64) · Ancient Roman inscriptions in Austria (61) · Ancient Roman inscriptions in Turkey (57) · Inschriften vom Obergermanisch-Rätischen Limes (57) · Ancient Roman inscriptions in Galicia (35) · Ancient Roman inscriptions in D. Diogo de Sousa Museum (28) · Ancient Roman inscriptions in Extremadura (29) · Ancient Roman inscriptions in Jordan (44) · Ancient Roman inscriptions in France (44) · Inscriptions from ancient Germania superior (43) · VSLM inscriptions (37) · Ancient Roman inscriptions in Spain (42) · Inscriptions from ancient Provincia incerta (25) · Ancient Roman inscriptions in Castile and León (28) · Inscriptions from ancient Umbria (20) · Inscriptions from ancient Asia (40) · Ancient Roman inscriptions in Greece (37) · Inscriptions from ancient Germania inferior (37) · Ancient Roman inscriptions in Andalusia (23) · Inscriptions (35) · Religious dedications by ancient Roman military personnel (27) · Cultural heritage monuments in Spain with known IDs (16) · Legio VII Gemina (14) · Latin inscriptions without nomenclature (22) · Ancient Roman inscriptions in Hungary (26) · Pontevedra museum (20) · Ancient Roman inscriptions in the Republic of Macedonia (24) · Inscriptions from ancient Samnium (12) · Inscriptions (22) · Images from Wiki Loves Monuments 2013 (15) · Ancient Roman inscriptions in Romania (23) · CIL VI 031256 (11) · AE 1957 (11) · Norba Caesarina (16) · AE 1958 (11) · CIL VI 041140 (11) · CIL VI 001187 (11) · Inscriptions from ancient Baetica (22) · Inscriptions from ancient Moesia superior (22) · Images from Wiki Loves Monuments 2011 (17) · Inscriptions from ancient Belgica (21) · Ancient Roman aras in Germany (20) · Inscriptions from ancient Pannonia inferior (20) · Ancient Roman aras in Portugal (12) · Inscriptions from ancient Venetia et Histria (Regio X)? (10) · Sculptures in the Museu de Belles Arts de València (11) · Inscriptions from ancient Dacia (20) · Ancient Roman inscriptions in Bulgaria (19) · Inscriptions from ancient Transpadana (9) · License migration completed (15) · Miliaria in Spain (10) · Images from Wiki Loves Monuments 2013 in Spain (10) · RIU 0575 (9) · Inscriptions from ancient Latium et Campania (Regio I)? (9) · CIL VI 040771 (9) · PD Old (15) · Ancient Roman inscriptions in Split archaeological museum (17) · AE 1981 (8) · CIL VI 037133 (8) · AE 1948 (8) · Inscriptions from ancient Etruria (Regio VII)? (8) · AE 1962 (8) · AE 1989 (8) · CIL III 011019 (8) · RIU 0573 (8) · CIL VI 041052 (8) · AE 1948 (8) · RIU 0529 (8) · AE 1991 (8) · Files by DerHexer (15) · Ancient Roman inscriptions in the United Kingdom (15) · Inscriptions from ancient Mauretania Caesariensis (15) · Ancient Roman inscriptions in Switzerland (15) · Steles in Galicia (9)

Figure 7.9: Categories mostly used on items curated or uploaded by EAGLE partners

Recommendation 10: Europeana should investigate becoming the first Wikimedia "Movement Partner"

Collaborating with Wikimedia Projects is in itself a best practice which everyone believes should be adopted for GLAMS. This is proven when looking at the results given from the BaGLAMa 2 Tool for the pageviews of items in the Category Media Contributed by EAGLE.

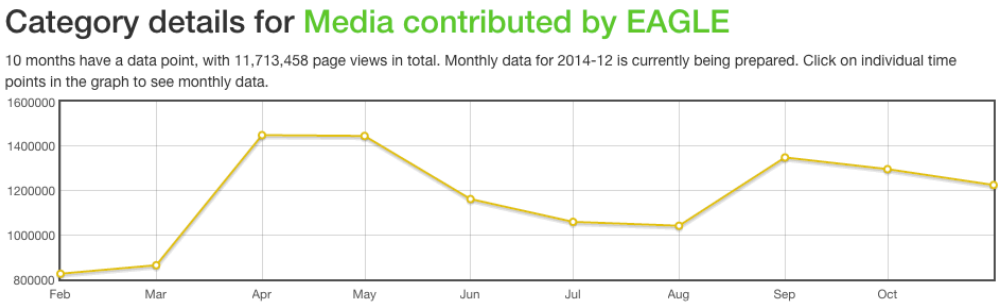


Figure 7.10: Page views for the files contained in the Category Media Contributed by EAGLE

Page views in 2014-9

Total monthly page views: 1,348,722. Download this table.

Site		Pages	Views
English Wikipedia	Details	118	650,189
Spanish Wikipedia	Details	147	219,547
German Wikipedia	Details	191	146,206
French Wikipedia	Details	99	124,878
Portuguese Wikipedia	Details	36	38,462
Italian Wikipedia	Details	41	34,789
Russian Wikipedia	Details	16	28,159
Polish Wikipedia	Details	4	26,343
Romanian Wikipedia	Details	26	14,041
Bulgarian Wikipedia	Details	24	13,459
Dutch Wikipedia	Details	27	10,352
Chinese Wikipedia	Details	5	6,027
Czech Wikipedia	Details	5	5,397
Hungarian Wikipedia	Details	10	5,216
Norwegian Wikipedia	Details	8	3,802
Slovenian Wikipedia	Details	12	3,634
Japanese Wikipedia	Details	8	3,351
Serbian Wikipedia	Details	9	2,157
Galician Wikipedia	Details	18	1,799
Catalan Wikipedia	Details	20	1,235
Arabic Wikipedia	Details	2	1,168

Figure 7.11: Page views for the files contained in the Category Media Contributed by EAGLE (values above 1000 by country)

Key Best Practices Summary for this chapter

BP 16: Synchronization with other major projects It is always beneficial to understand what interaction there can be and to evaluate where it is better to build on existing resources instead of running parallel projects.

BP 18: Wikimedia Developments It is good practice to collaborate as much and as thoroughly as possible with Wikimedia projects at local and national level.

HE 8: Metadata Curation Given the existing efforts by users of Wikimedia Commons we believe that it is best practice to contribute to a common effort, unrelated to project's lifetime in order not only to give accessible content better quality but also to enhance.

HE 9: GLAM Toolkit upload images with several control stages and the support of the Wikimedia Community, different from the original one and able to provide fit for purpose advice

HE 10: Perseids The integration with Perseids and the set up of a board of editors allows to have full control while maintaining the system entirely open to contributions

Chapter 8

Best Practices for the Future of Epigraphy and Digital Epigraphy

8.1 The EAGLE network as a Best Practice

The EAGLE project has achieved the application of most of the standards agreed by the Electronic Archive for ancient Greek and Latin Epigraphy, the preexhisting consortium who met in 2008 to discuss these, as requirements for a single portal (Mambrini 2013):

- alignment to an EpiDoc export of most of the available resources for use and reuse (which allows for redesign of these resources to become full digital editions);
- connections and software tools to move whole systems to highest available data standards to the maximum extent possible;
- alignment of controlled vocabularies to a unified but non substitutive resource for epigraphy;
- support for further documented enlargement of the network outreach to other sectors and their user needs;
- creation of several common reference resources as translations of inscriptions and a general epigraphic bibliography.

A new discipline revolves around EAGLE and the activities it fosters and supports. More then 200 participants[1] per day attended at the Paris conference on Digital Epigraphy,[2] attesting that the years of work on the databases are part of a consistent and steady movement towards full digital editions of epigraphic material. These participants were in large part members of the EAGLE project Best

[1] Number of participant registered on Eventbrite, where the event was sold out.

[2] The program is available here: `www.eagle-network.eu/about/events/eagle2014/` while the proceedings are available here: `http://www.eagle-network.eu/wp-content/uploads/2015/01/Paris-Conference-Proceedings.pdf`

Practice Network which include a variety of members from different traditional disciplines:

- latin and greek epigraphy;
- other epigraphies (arabic, mayan, etc.);
- museum partners and archaeologists;
- linguists and image experts;
- encoders and IT specialists (EpiDoc and CIDOC international communities);
- Digital Humanities scientists.

This is a network of specialists who are committed to the best for digital epigraphy as a discipline in its own, as a support for other researchers from different context at the highest standard currently available of data architecture and management. The *Epigraphic Database Clauss Slaby* is the last among all existing epigraphic databases whose collaboration is still to be confirmed, although some interest from them in joining the project has been notified. Also the *Inscriptiones Graecae* team and the *Christian Inscriptions* database of the Berlin Academy, as well as several members of the team of the *Corpus Inscriptionum Iudaea/Palestina* have already expressed interest in contributing to the common efforts pursued by the EAGLE consortium. Every day an increasing number of projects in their beginning or in their latest phases contact EAGLE to join in.[3]

While this is an unavoidable development the detachment of Epigraphy and Digital Epigraphy is not something that can be simply allowed to happen, because it would be of detriment to the effort of all. This new movement needs to substitute or fuse together with the traditional epigraphies whose objective it pursues with new editorial and workflow techniques. A renewed bond and a full integration needs to take place so that a student of Epigraphy will be always studying also XML and won't be scared of using a web editor and an online repository instead of a WYSIWYG editor and a local database.

8.2 Users needs in 2014

Digital Epigraphy is not any more at the stage of pilot projects, it is a mature discipline bringing contributions to its own field, to the disciplines who have given contribution to it and to others disciplines it serves by definition. It has its own questions and problems and its critical points. The same might be said for users of epigraphic databases who expect everyday more and use with more knowhow the resources they are presented with. Nevertheless, today, users need to check two to five databases to be actually sure that they have found and seen all available information online, i.e. information about printed material and all what is

[3]The *AXON* project from the University of Venice for example and the *British Museum Inscription's* database for example among the latest.

natively online. And this, provided that they are aware of the existence of all the available databases Most of the resources developed are of a small scale, although maybe rich in content, and cannot afford to invest in search engine optimization, or are simply unaware of what would need to be done to make them selves visible to crawlers and users or of what the possibilities are. But in the last years projects of extremely high quality did instead manage to gain the visibility of a printed publication and international recognition. Other projects are simply not publicized enough, making them pointless in most cases, because the mere existence of a resource in the web is not enough to claim for it usefulness. Too many time the purpose of making accessible by ”putting it online” is flagged without this basic need of dissemination, aggregation, exposure aspects. There is no guarantee that the same search in major engines will return the same results due to current search algorithms. But there is also no thought about long term stability when the deep web contains most of what has been developed in tables, databases and containers which are very practical to build, functional perhaps, but unaccessible to crawlers and therefore not archivable, hardly reusable and invisible, to the detriment also of usability of each site, whose content are hidden to the user who can only search via an interface but cannot browse for example. The EAGLE consortium has been working in the framework of the EAGLE project to put an end to this by aggregating and adding stable and unique IDs to texts and contents, giving data a structured framework and bringing them as much as possible to the surface as raw and transformed data. But the path to go forward is very long indeed and compared to the desirable outcomes this is only a beginning.

It cannot be denied that online epigraphic edtions and data, like photos or translations, are nowadays stably integrated in the workflow of researchers and users; most often, it can be claimed, scholars start their researches by resorting to online rather than printed collections. It is therefore very important that the qualified information provided online is visible, aware of possible meaningful connections, but most of all reliable and curated from the beginning. Relying on exports and aggregators is for now perhaps a good temporary solution but not a long term sustainable one. If nobody takes care of that, a series of chain reaction misunderstandings, silently hindering research beyond the scope of researcher responsibility, might occur. It is therefore of the outmost importance that the information online is curated by specialists who take continuous care to provide the best possible quality of data and contents.

For example no update is systematically made in the current bibliographic repertoires like AE of digitally born editions which are published. It would not probably be possible. But a digital repertoire could do both instead. Sometimes newly published or updated materials in digital editions online is overlooked entirely to the detriment of research in epigraphy as a whole. A unique overseeing body to keep track of all projects and feed in newly published and updated materials in printed and digital editions would therefore be strongly necessary but is currently absent.

8.3 Methodology: old and new issues

The first generations of epigraphers used to accomplish a series of tasks

- Autopsy of the inscribed object;
- write letters to other scholars and researchers;
- research on published material in print;
- draw the objects and the texts;
- make squeezes for further analysis;
- collect documentation in archives;
- build indexes and concordances;
- publish printed editions;
- publish new complete corpora;
- keep bibliographic repertoires updated for research.

This methodology of work implied a series of issues:

- Very long publication time;
- Costs of print publication for author and editor;
- Cost of print publication for readers;
- Editorial limits: apparata;
- Editorial limits: mark up of text;
- identification;
- time paradoxes in publications;
- no photos;
- no translations;
- niche effect.

Today the list of tasks for Epigraphers has increased abundantly and calls for a long series of side and extra competencies. Here is a non comprehensive list:

- Autopsy of the inscribed object;
- emails to other scholars and researchers;
- research on published material in print;
- research on published material online;
- drawings;
- squeezes;
- professional photos;
- archives;
- new printed editions with photos;
- new corpora with photos;
- repertoires (only of printed publications);
- databases (local and online);
- digital re-editions;
- native digital editions;

- 3D scans and reconstructions;
- laser scans;
- Mathematical models;
- projects on related topics;
- geographical data handling;
- identification.

Times have changed, some issues have been mitigated by evolution of technology and some possibilities arose for the first time only recently or just became affordable.

The researcher can restitute its work to the community with new media, as the web, but this means also new techniques and editorial tools to give back research results. Nevertheless the fundamental questions stay the same, and choices differ on these regards as it happened before both in printed publications and in digital editions and related projects: How to describe the object and the text? How to represent it? What is useful for the reader? What limits does the collection have? Decide or let decide and to what extant? How much ready made analysis should be available? Quality? Quantity? which relevance does the research have? How frequently it should be updated? Which parts can be delegated to other researchers?

We have now large amounts of data, which enables sophisticated analysis, we have the possibility to easily automate indexes, to interoperate items and cross-check papyri, literature, archeological data, digital libraries, digital projects' contents, potentially with one only access point (whichever we choose among many). But still most of the previously mentioned issues persist:

- not accessible photos;
- not accessible squeezes;
- few translations;
- only limited piloting of techniques;
- volatility of databases;
- multiplication of places to check;
- limited integration of digital and printed;
- even more time paradoxes;
- fast update of digital editions.

Some have changed and new ones rose, and an all set of new questions where brought in from the studies of Digital Humanities:

- format;
- metadata format;
- only partially machine readable format and metadata;
- many and with very different rationale;
- apparent digital divide (people prefer to spend months not to learn a popular database software, rather then face a mark-up language);

- necessary Team Work for research and for publication: no simple author-editor interaction;
- Cost of digital edition and database;
- Non-trivial amount of developmen know-how required;
- sustainability (for the author and for the reader).

Digital editions run ahead in many terms but are sometimes pulled back by trying to keep always up with printed publication and its times, especially when they do not aim at a publication but at an archival function.

So there are many more problems than there used to be. Where to go then? Back to a simpler approach and faster on getting research done? Or forward, accepting new challenges and a slow but complete change in the methodology with a promise for further possibilities? Stick to the basic problems or gather also the new ones?

Moving forward presents challenges, opportunities and risks.

Challenges:
- harmonize without standardizing;
- open access;
- facilitate work with enough tools to meet the demands of users;
- advocate for digital editions;
- advocate for appropriate recognition of digitally published research work, and of contribution to collaborative efforts.

Opportunities:
- Interoperate resources to highlight unforeseeable connections;
- New results from Networking and Team work;
- Greater Outreach;
- Exploit project's results and contribute - get more interesting finds than statistics on contents (e.g. number of items from one place or with a particular feature).

Risks:
- different philosophies and editorial techniques;
- different purposes;
- misunderstandings of interoperability (DE is not a supplement to E, E are not users of DE resources);
- Digital Epigraphy and Epigraphy split up.

Things can move forward not just faster but for better research with the slow and problematic (i.e. critical and analytical) digital approach. EAGLE has done to this end a ground networking activity which also means bringing together techniques and making common efforts instead of individual attempts and has started to pave the way accomplishing initial tasks towards wider objectives, like devising metadata structures fit for purpose, providing training for text encoding, making and making available photos, building applications to facilitate reuse and new ways to use the data, linked data techniques to unveil potential connections otherwise unexpected. The EAGLE BPN has also devoted attention to curation and

contribution to other projects, as we cannot just require others to collaborate, we need to contribute to each others projects in the best way we can, and where needed close gaps with sub-projects which never miss the chance to experiment new tools and potential further developments, focusing on the dynamic of progress for research rather than on its speed. The EAGLE BPN believes in building resources which are always open to further input and sustainable because they are reusable. Reuse of the data produced, rather than (just) future funding, is the crucial factor to ensure to a project a long-lasting life. It is by reuse that the data produced will increase chances of its survival. And this is a lesson we learn from the past, from manuscripts for example and copyist monks, as well as the first publishers, which unanimously fought a battle to preserve by making as many editions as possible of the manuscript they had.

8.4 The future of digital epigraphy

The EAGLE aggregator will be hosted in Sapienza Università di Roma,[4] although the aggregation process will slow down in 2016, together with the networking activity. The databases will continue their existence together with the EAGLE project and all efforts will be made not to loose the meaningful infrastructure behind the aggregation process. This is the reason why a lot of effort is being put into making aggregated content well known to potential re-users because that will give continuity and sustainability to some parts of the work which will be done. Only by being reused and shared continuously the content produced and curated stands a chance of survival and relevance in the near and far future.

With the end of funding for the EAGLE project approaching, we are already working towards the long term sustainability of the many projects in the consortium and best practice network, so that the products developed will not be lost in few years and will continue to be useful for the researchers.

Keeping one stable decision board will enable each database to set aside internal competitions to move ahead together for the benefit of the entire community and will still allow to maintain a commonly coordinated coverage of the materials with responsibilities given on the basis of competence and quality. A continuous process of alignment to each other best practices by moving towards the new aggregation standards will instead enable a general improvement of resources which cannot invest in long term improvement due to project time constraints and lack of a possible planning horizon.

Digital Epigraphy has gained fame for being always ahead and opening ways and path for other disciplines. The scholars and researchers taking part in it keep up with the evolving environment and do not fall behind. The path chosen and

[4] Further and more detailed information will be found in the specific deliverable on this subject.

taken in the last thirty years could then be exploited and taken forward to the next thirty, advocating for the best in a coordinated and networked manner.

Bibliography

AA.VV (2013). *AIM Infrastructure Specification*. Deliverable 4.1. Eagle Project (325122).

Berners-Lee, T. (2006). *Linked Data - Design Issues*. Design Issues. URL: http://www.w3.org/DesignIssues/LinkedData (visited on 03/08/2012).

Bizer, C., T. Heath, and T. Berners-Lee (2009). "Linked Data - The Story So Far". In: *International Journal on Semantic Web and Information Systems, Special Issue on Linked Data* 5 (3), pp. 1–22. ISSN: 1552-6283, URL: http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf.

Buckland, M. and L. Lancaster (2014). "Combining Place, Time, and Topic: The Electronic Cultural Atlas Initiative". In: 10.5. URL: http://www.dlib.org/dlib/may04/buckland/05buckland.html (visited on 03/11/2014).

Creative Commons (2012). *About The Licenses - Creative Commons*. URL: http://creativecommons.org/licenses/ (visited on 05/15/2012).

Description, o. W. (2013). *Europeana Network of Ancient Greek and Latin Epigraphy*. Description of Work. Eagle Project (325122).

Elliott, T., G. Bodard, E. Milonas, S. Stoyanova, C. Tupman, and S. Vanderbilt (2007). *EpiDoc Guidelines: Ancient documents in TEI XML*. URL: http://www.stoa.org/epidoc/gl/latest/.

Feraudi-Gruenais, F., ed. (2010). *Latin on Stone: Epigraphic Research and Electronic Archives*. In collab. with S. A. Takacs, F. Feraudi-Gruenais, W. Stenhouse, R. R. Benefiel, C. Witschel, G. Bodard, S. Evangelisti, and A. Kropp. Lexington Books. ISBN: 0739145908.

Ginouvès, R. (1998). *Dictionnaire méthodique de l'architecture grecque et romaine. Tome III. Tome III.* Rome; Athènes (GR): École française de Rome ; École française d'Athènes. ISBN: 2728305293 9782728305292.

Guarducci, M. (1987). *L'epigrafia greca dalle origini al tardo Impero*. Roma: Istituto Poligrafico e Zecca dello Stato, Libreria dello Stato.

Harper, C., G. Dunsire, D. Hillman, and J. Phipps (2012). "Linked data vocabulary management: infrastructure support, data integration, and

interoperability". In: *Information Standards Quarterly* 24 (2/3). ISSN: 1041-0031.

Harpring, P. (2010). *Introduction to Controlled Vocabularies Terminology for Art, Architecture, and Other Cultural Works*. URL: http://www.getty.edu/research/publications/electronic%5C_publications/intro%5C_controlled%5C_vocab/pdf.html.

Heath, T. and C. Bizer (2011). *Linked Data: Evolving the Web into a Global Data Space*. 1.0. Morgan & Claypool. ISBN: 9781608454310. URL: http://linkeddatabook.com/editions/1.0/ (visited on 03/15/2012).

Hughes, N. (2011). *Constructing the Open Data Landscape*. ScraperWiki Data Blog. URL: http://blog.scraperwiki.com/2011/09/07/constructing-the-open-data-landscape/ (visited on 09/26/2011).

Isaac, A., R. Clayphan, and B. Haslhofer (2012). "Europeana: moving to Linked Open Data". In: *Information Standards Quarterly* 24 (2/3), pp. 34–40. ISSN: 1041-0031.

Lassère, J.-M. (2005). *Manuel d'épigraphie romaine. 2, 2,* Paris: Picard. ISBN: 2708407325 9782708407329.

Liuzzo, P. M., R. Santucci, and C. Prandoni (2013). *Networking infrastructure and terms of reference*. Deliverable 2.1. Eagle Project (325122).

Mambrini, F. (2013). *EAGLE portal and services requirements and specifications*. Deliverable 5.1. Eagle Project (325122).

Manzella, I. D. S. (1987). *Mestiere di epigrafista: guida alla schedatura del materiale epigrafico lapideo*. Edizioni Quasar. ISBN: 9788885020795.

Martin, X. (2014). *First release of the flagship mobile application and SDK*. Tech. rep. Eagle Project (325122).

Open Data Commons (2012). *Licenses | Open Data Commons*. URL: http://opendatacommons.org/licenses/ (visited on 05/15/2012).

Orlandi, S., R. Santucci, V. Casarosa, and P. M. Liuzzo, eds. (2014). *Information Technologies for Epigraphy Information Technologies for Epigraphy Information Technologies for Epigraphy and Cultural Heritage*. Studi Umanistici. EAGLE BPN. http://digilab-epub.uniroma1.it/index.php/Antichistica/art Sapienza Università Editrice.

Piso, I. (2001). *Inscriptions d'Apulum*. Inscriptions de la Dacie Romaine - III 5).

Prandoni, C., N. Alfarano, and V. Casarosa (2014). *EAGLE portal*. Tech. rep. Eagle Project (325122).

Rivero Ruiz, E., M. Andrea, and V. Vassallo (2013). *EAGLE metadata model specification*. Deliverable 3.1. Eagle Project (325122).

Stevenson, J. (2012). "Linking LIves: Creating an end-User Interface Using Linked Data". In: *Information Standards Quarterly* 24 (2/3), pp. 14–23. ISSN: 1041-0031.

Talbert, R. J. (2000). *Barrington Atlas of the Greek and Roman World: Map-by-map Directory*. Vol. 1. Princeton University Press, p. 680. ISBN: 0691049459.

World Wide Web Consortium (2012). *Semantic Web - W3C*. Semantic Web - W3C. URL: http://www.w3.org/standards/semanticweb/ (visited on 04/05/2012).

Zeng, M. L. and L. M. Chan (2006). "Metadata Interoperability and Standardization - A Study of Methodology Part II". In: *D-Lib Magazine* 12 (6). ISSN: 1082-9873. DOI: 10.1045/june2006-zeng. URL: http://www.dlib.org/dlib/june06/zeng/06zeng.html (visited on 03/09/2012).