CA·R·M·E·N

Working Papers
3/2022

Presentation of *CLEO* Database: Digital text edition "creating" specific perceptions of texts

Concepción Fernández Martínez
Universidad de Sevilla
cfernandez@us.es

María Limón Belén
Universidad de Sevilla
mlimon@us.es

Sergio España-Chamorro
Universidad Complutense de Madrid
sergio.espana@ucm.es

Horizon 2020 European Union Funding for Research & Innovation

MARIE CURIE ACTIONS

The CARMEN project receives funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 954689.

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

CARMEN Working Papers 3/2022

Edited by Prof. Dr. Marietta Horster and Dr. Erika Fischer
Historisches Seminar – Alte Geschichte
Johannes Gutenberg-Universität Mainz
Welderweg 18 (Philosophicum)
D-55122 Mainz
Email: carmen-itn@uni-mainz.de; https://carmen-itn.eu

DOI: http://doi.org/10.25358/openscience-8191

CARMEN Working Papers are published under the CC BY-ND 4.0 licence.

Presentation of *CLEO* Database: Digital text edition "creating" specific perceptions of texts

Concepción Fernández Martínez,
María Limón Belén & Sergio España-Chamorro

Research in the humanities is, of course, contemporary in its questions and in its use of technologies. It is our desire as well as our duty to make use of the facilities of our age, to offer keys to the past, and to meet the needs and interests of the users. We thus saw the need to revise our methodologies and to apply an innovative digital approach to the study of Latin epigraphic verse. This approach has allowed us to make our studies and editions as a whole available to the research community and a broader public interested in Roman history and poetry. At the moment, we focus on epigraphic poetry from Hispania, Britannia, and partly from Gaul, although we have an eye on gradually completing a digital corpus of inscriptions in verse encompassing the Roman world as a whole.

1. Our starting point: *CLE-Hispaniae*

In 2013, with the edition and study of all the verse inscriptions from Hispania (276) completed and awaiting their publication in the *CIL* (which had been postponed for various reasons and will finally happen in 2022), we became aware of the need and the advantages of disseminating our work by incorporating emerging technologies. We thus began to design, develop, and maintain an interactive database with the vision to convert it into a basic research tool for the interested scientific community. The result was the website called *Carmina Latina Epigraphica Hispaniae. Portal of Latin Epigraphic Poetry: Iconic and Textual Searches* (*CLE-Hispaniae*: http://cle.us.es/clehispaniae; Fig. 1).

Figure 1

This portal offered – and still offers – our edition and study of each *CLE* in three languages (Spanish, Latin, and English). It provides general but comprehensive information on the different elements of each inscription (chronology, layout characteristics, metrical form, etc.). It also includes a wide range of search options with regard to both the edition of the inscriptions and their philological analysis. Users will be able to perform searches on the database in any of the three languages mentioned above, obtaining results in the language in which their question was entered.

A couple of years ago, we decided that we had to take a step forward, and we started to think about the design of a new, updated, and exportable database, whose results could be linked to those of other international databases. This would truly become a basic research tool for the interested scientific community.

2. What is *CLEO* and how did we create it?

In 2019, the need for updating *CLE-Hispaniae* has motivated the creation of a new website: *Carmina Latina Epigraphica Online* (*CLEO*: https://institucional.us.es/cleo; Fig. 2). This new database not only comprises the former *CLE-Hispania* information but also aims to present the edition and commentary of Latin verse inscriptions that a group of researchers from different Spanish and European universities are working on.

The innovation of *CLEO* is thus twofold: on the one hand, it is a database with a global geographical objective and thus overcomes the geographical restriction to the Iberian Peninsula of the first project; on the other hand, the old database, which consisted of downloadable pdf files, has been adapted to the XML standards for encoding ancient documents. *CLEO* also provides a direct link to other epigraphic databases (*EDCS*, *PETRAE*, *MQDQ*, *EDR* ...) and transforms the available data into importable and exportable information.

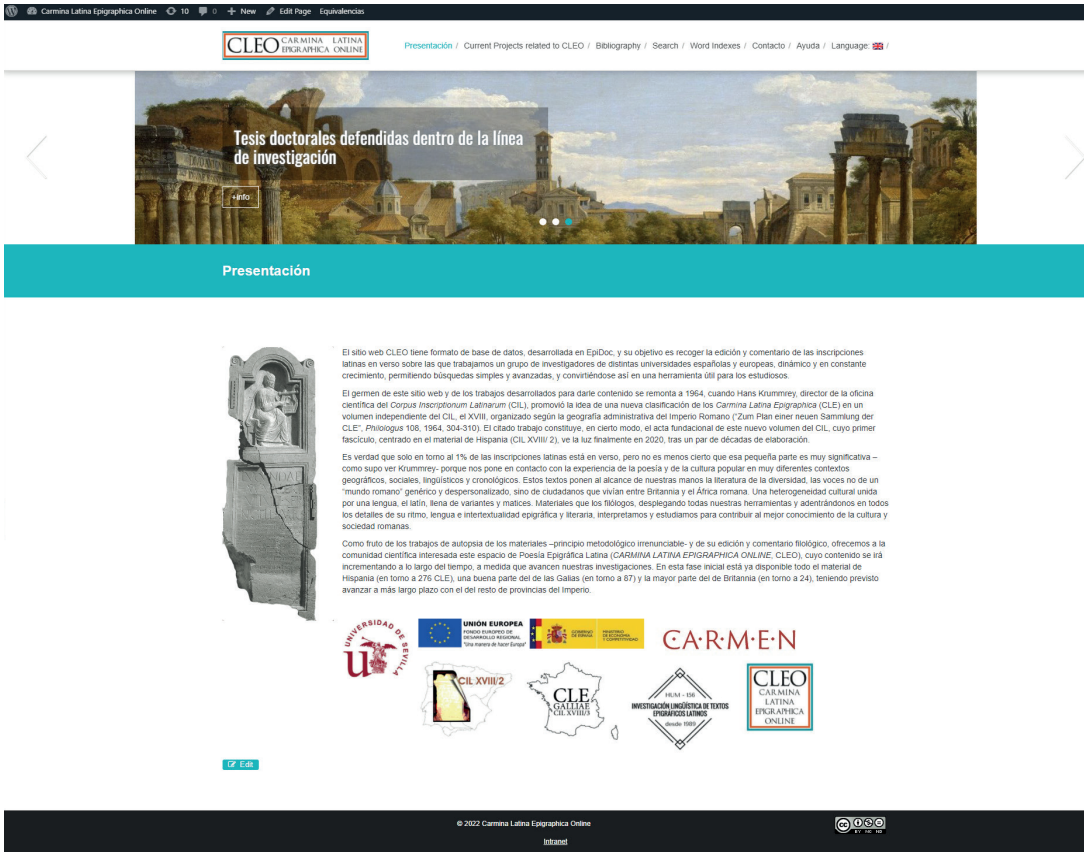

Figure 2

The realisation of *CLEO* involved much more than a simple transformation as it required the adaptation of our contents to the worldwide standards for the digital edition of epigraphic texts. We adopted the *Text Encoding Initiative* (TEI) with *Extensible Markup Language* (XML) proposed by the international EPIDOC project. The exchange of information with other epigraphic databases that use this system, as well as the fusion with any of them if necessary, requires editions in the same format. This ensures that the transfer of data works automatically without the need to adapt the contents again – with the consequential loss of time and need for resources. Large international projects with which we already foster scientific relations use this system, including the *PETRAE* database (*Programme d'Enregistrement, Traitement et Reconnaissance Automatique en Epigraphie* of the Ausonius Institute in Bordeaux), which has been our basic model. With our new *CLEO* database, we are joining these major international projects.

3. What does the *CLEO* project offer researchers today?

As a result of the autopsy of the materials – an essential methodological principle – and their philological editing and commentary, we can offer, in this initial phase, all the material from Hispania (around 276 *CLE*s), a good part of that from Gaul (around 87), and all the material from Britannia (24). Ongoing projects and the continuous updating of the database will allow the inclusion of materials from other provinces such as Germania, Mauretania, or Dacia. In the long term, we plan to move forward with the rest of the provinces of the Empire, for which we will count on the collaboration of our colleagues.

One of the specificities of the materials we work with is precisely the metrical scheme. Our database offers the epigraphic edition with all its diacritical marks and also the text, divided into verses with the metrical signs (Fig. 3). Moreover, we have managed to make the metrical fonts visible. This is an additional benefit, as it makes it possible to assess, at a glance, traditional metric forms or deviations and their possible causes (all of which is explained in the commentary). Perhaps the greatest specificity our database offers is the commentary, which is entirely philological, as if the inscription were any poem by any author, Horace, for example.

Type of verse

- **Type of verse:** Dactílico (dístico elegíaco)
- **Verse/line correspondence** Si
- **Prose/verse distinction** Si

Epigraphic edition

a) Mussia ▲ Agele
bis ▲ denis ▲ Agele ▲ florentibus ▲ annis
et ▲ specie ▲ et ▲ vita ▲ femina ▲ prima ▲ fuit
hunc ▲ sortita ▲ locum ▲ miserae ▲ sunt ▲ ossa ▲ puellae
5 hic ▲ raptam ▲ matri ▲ consociavit ▲ humus
sedibus ▲ aeternis ▲ Agele ▲ non ▲ laesa ▲ quiescas
et ▲ bene ▲ conposita (!) ▲ sit ▲ tibi ▲ terra ▲ levis
b) Mussia ▲ Ↄ (= Gaiae) ▲ l(iberta) ▲ Rosia
hic ▲ s ▲ t ▲ t ▲ l
10 c)L ▲ Postumius ▲ Barnaeus
hic ▲ s(it) ▲ t(ibi) ▲ t(erra) ▲ l(evis) ▲ Ↄ (= Gaiae) ▲ l (ibertus) ▲

Text divided into Verses

<Mussia> bis denis Agele florentibus annis –⏑⏑l–/–l–/⏑⏑l–/–l–⏑⏑l–⏓
et specie et vita femina prima fuit. –⏑⏑l––l–ll–⏑⏑l–⏑⏑l~
hunc sortita locum miserae sunt ossa puellae: ––l–⏑⏑l–/⏑⏑l–/–l–⏑⏑l–⏓
hic raptam matri consociavit humus. ––l––l–ll–⏑⏑l–⏑⏑l~
5 sedibus aeternis, Agele, non laesa quiescas, –⏑⏑l––l–/⏑⏑l–/–l–⏑⏑l–⏓
et, bene conposita, sit tibi terra levis. –⏑⏑l–⏑⏑l~ll–⏑⏑l–⏑⏑l~

Figure 3

Of course, we work *ex novo* but not *ex nihilo*. In each of our *schedae*, we have reviewed all the previous bibliographies, collections, editions, and general and specific studies in order to create a critical apparatus and to provide a documented commentary.

4. Some technical issues

As already mentioned, the main novelty of the *CLEO* database is the use of the international collaborative system called EpiDoc. This system of text encoding tools (Text Encoding Initiative's standard) is essential for the online publication of editions of documents from the past in general and of inscriptions in particular. Numerous epigraphic databases have already adapted their data to this system as it ensures the preservation of data and allows it to be exported and imported.

The code parameters follow the set standards. On the technical support side, the vocabularies developed by the Europeana Eagle project are applied. In terms of geography, we follow the system developed by the Trismegistos project and we included the ID that this project assigns to each document. This will allow a reliable geolocation on the general map that will be programmed in the second phase of the project (*CLEO* 2.0; see below).

We also collaborate with the Digital Epigraphic community (https://epigraphy.info) in order to join general projects and to enhance the digital epigraphic databases in general.

The *CLEO* intranet is organised by means of a form into which the data is entered (Fig. 4).

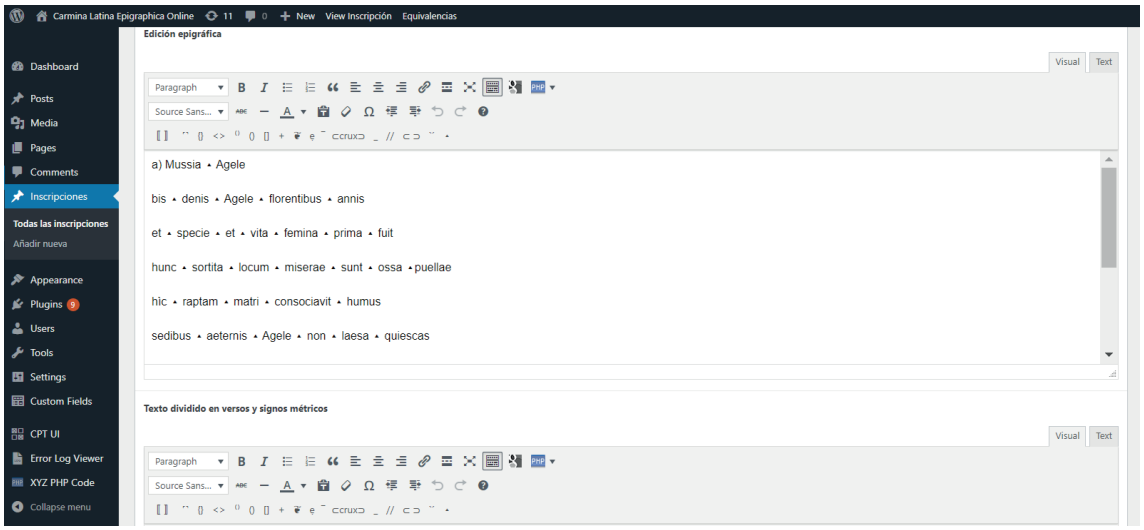

Figure 4

Once the file is published, XML code is instantly available as an automatic process (Fig. 5).

Figure 5

5. Internal mode of operation for data entry

When the database was designed, we aimed at making EpiDoc easy to use for anyone. Thus, we designed a form with all necessary boxes that automatically converts the entered data into EpiDoc language. There are text boxes, dropdown menus, and a text editor to insert the Latin text of each inscription. The editor includes buttons for quick and easy insertion of graphic signs such as *hederae* or punctuation, parentheses, and everything else necessary when editing texts of this type. Like the text in the rest of the form, the text inserted in this box is automatically transformed into EpiDoc language.

As we said, our vision was to optimise and facilitate the process so that a specific mastery of XML or EpiDoc to insert data is not necessary. In this sense, we are very satisfied with the result.

6. Search processes and possibilities

The search tool is based on the old *CLE-Hispaniae* system but incorporates the new search engine and allows searching in several records. On the one hand, there is the simple search option (Fig. 6), which allows direct search for any term in all the boxes of the records already included in the database.

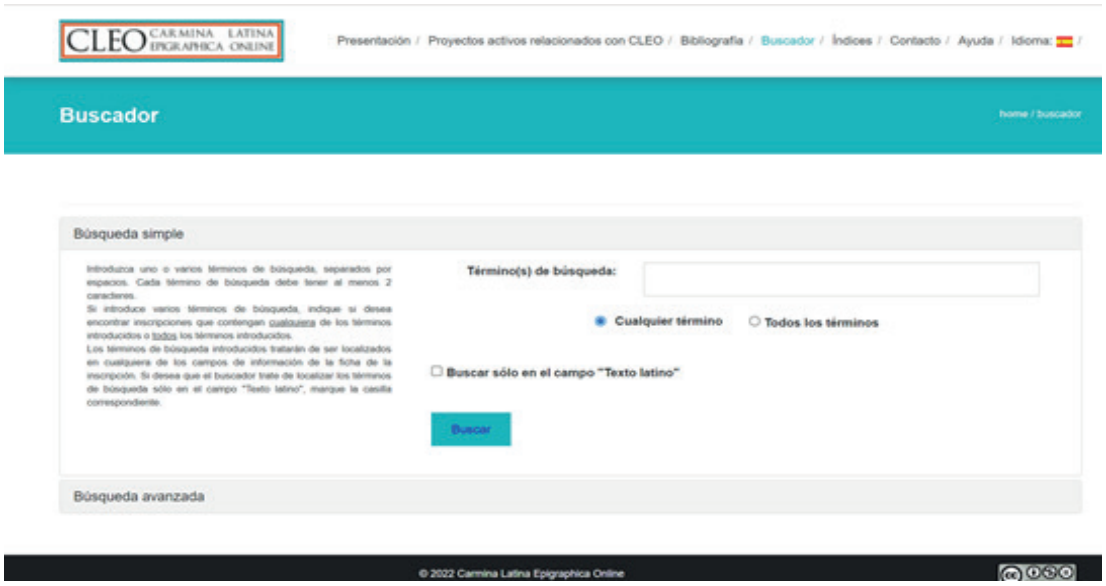

Figure 6

On the other hand, there is, like for many other databases, an advanced search option (Fig. 7). It offers a limitation to four specific parameters: geography, characteristics of the epigraphic support, dating, and text. These can be used individually or in combination.

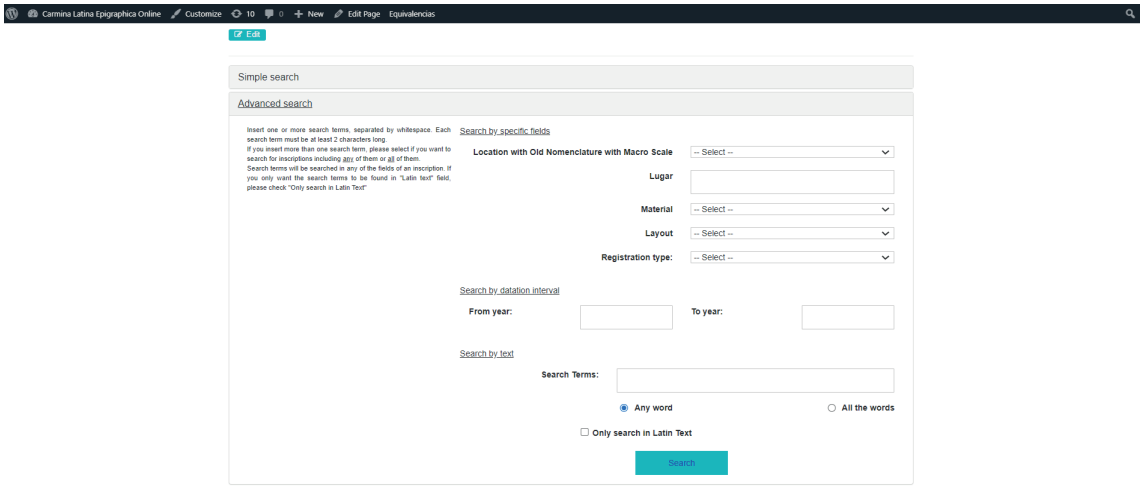

Figure 7

Within the geographic section, the search can be narrowed down by two more parameters: It is possible to search by province or by place. For the province also appear additional dropdown menus in case of internal subdivisions in several established levels such as provincial groups (like Gallia, Hispania, Germania, Moesia, …) or sub-provincial groups (mainly for the Hispanic *conventus iuridici*). The box for a specific place fills in automatically, which allows searching for place names such as cities or micro-toponyms of specific archaeological sites, villages, and very specific places.

The search by characteristics of the epigraphic support allows searching by material, type of support, or registration type. Once again, this typology follows the Europeana Eagle vocabularies (https://www.eagle-network.eu/resources/vocabularies/). The search by dating makes it possible to search for specific time intervals or in between certain dates. Lastly, the text search allows to enter specific words (as in the simple search) that can be combined with the other search parameters mentioned.

The search results appear all on one page instead of on different tabs and show some basic data on the record and the photograph (Fig. 8).

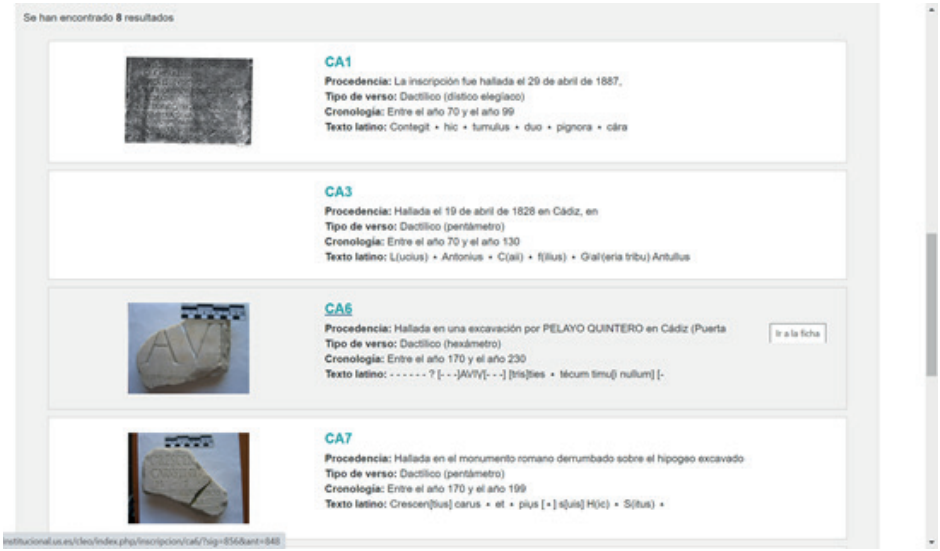

Figure 8

7. Our next steps

Now that the database and the website are operating, we are preparing *CLEO* 2.0. We will mainly work on implementing the import-export tool in order to directly save, protect, copy, and share our data. Additionally, one of our priorities is to create a map on which users can easily see the available material. We have already included the Trismegistos number that allows us to insert this data in a Web-GIS environment such as *PETRAE*, *EDH*, or *RIB*.