'Cretan Institutional Inscriptions' Meets CLARIN-IT

Irene Vagionakis
University of Bologna, Italy
`irene.vagionakis2@unibo.it`

Paola Baroni
CNR-ILC – Pisa, Italy
`paola.baroni@ilc.cnr.it`

Riccardo Del Gratta
CNR-ILC – Pisa, Italy
`delgratta@ilc.cnr.it`

Angelo Mario Del Grosso
CNR-ILC – Pisa, Italy
`angelo.delgrosso@ilc.cnr.it`

Federico Boschetti
CNR-ILC & VeDPH – Pisa, Italy
`federico.boschetti@ilc.cnr.it`

Tiziana Mancinelli
VeDPH – Venezia, Italy
`tiziana.mancinelli@unive.it`

Monica Monachini
CNR-ILC – Pisa, Italy
`monica.monachini@ilc.cnr.it`

Abstract

This paper presents *Cretan Institutional Inscriptions*, a resource in the domain of Digital Epigraphy developed at the Ca' Foscari University of Venice and supported by CLARIN-IT as part of its actions addressed to initiatives, projects and events in the field of Social Sciences and Humanities. The paper begins with a brief outline of the project within which the resource was created and then goes into a more in-depth description of the main methodologies used to develop the resource (EpiDoc and EFES) and of their benefits. The paper then focuses on the cooperation of the project with the Venice Centre of Digital and Public Humanities and the Italian node of CLARIN, also illustrating the dockerization process applied to the resource hosted on the CLARIN-IT servers. Some desiderata for future developments are outlined as well. The paper ends with some remarks about the widening of CLARIN horizons towards Digital Epigraphy and on the role of its K-Centres in this respect.

1 Project Description

The EpiDoc collection named *Cretan Institutional Inscriptions*[1] was created as a part of the Ph.D. research project in Ancient Heritage Studies *Kretikai Politeiai: Cretan Institutions from VII to I century BC*, carried out by Irene Vagionakis at the Ca' Foscari University of Venice (UNIVE) from 2016 to 2019 under the supervision of Claudia Antonetti and Gabriel Bodard. The database, built by using the EpiDoc Front-End Services (EFES), collects the EpiDoc editions of 600 inscriptions shedding light on the institutions of the political entities of Crete from the VII to the I century BC.

The project, which contributes to the landscape of Digital Humanities – in particular to that of Digital Epigraphy, through the creation of a new open access online epigraphic resource – and could hopefully be a forerunner for the inclusion of other digital epigraphy projects in the Common Language Resources and Technology Infrastructure (CLARIN), has been a valuable opportunity for collaboration with the Venice Centre for Digital and Public Humanities (VeDPH) and the Italian node of CLARIN (CLARIN-IT) during its final testing and publication stages.

1.1 Aim of the Research

The Archaic, Classical and Hellenistic history of Crete is a history characterised by a very high level of fragmentation. The numerous silences of the literary sources and the gaps in the epigraphic records have resulted in wide sectors of the island history still being overshadowed and in a similar fate befalling

[1]For a detailed description of the resource, see Vagionakis, 2021.

many of its political entities. In particular, the institutional history of Crete was greatly affected by such fragmentation and also by the bulky presence of the albeit scarce literary sources related to it. In fact, the alleged greater authoritativeness of authors such as Plato, Aristotle or Ephorus often prompted to force what was witnessed by the uneven epigraphic records to relate the contradictory information coming from different areas to a single model, thus flattening the variegations of the multiform landscape of «one hundred-cited Crete» (Hom. *Il.* II 649) in the name of the existence of a single unitary *Kretike politeia*. Within that framework, the Ph.D. research project set itself the goal of collecting systematically the records pertaining to Cretan institutions in order to propose an up-to-date reconstruction of the administrative framework of the island political entities, highlighting the specificity of each context, from the rise of the *poleis* and their first epigraphic records in the Greek alphabet(s) to the Roman conquest of Crete (VII-I century BC). By bringing together these so far scattered records in a searchable digital collection, the project also aimed at facilitating their finding, consultation and reuse.

2 The EpiDoc Collection and the TEI Catalogues

The core of the documentary basis of the research consisted of 600 Greek inscriptions, either directly mentioning institutional elements (as the decree from Knossos *I.Cret.* I 8 12 of the late II cent. BC: l. 1, ἔδοξεν Κνωσίων τοῖς κόσμοις καὶ τᾶι πόλι, 'the *kosmoi* and the *polis* of the Knossians decreed') or hinting at them through a revealing terminology (as the treaty between Hierapytna and an unknown *polis I.Cret.* III 3 6, of the late III or early II cent. BC: ll. 1–2, μηνὸς [- - -] τάδε ἔδ[οξ]εν τ[- - -], 'in the month of [- - -], [- - -] decreed these things').

For each inscription, an XML edition compliant with the TEI EpiDoc epigraphic substandard (Elliott et al., 2020) was created, including a descriptive and a bibliographic lemma, the text of the inscription, a selective apparatus criticus and a commentary focused on the institutional data offered by the document, plus links to other related online resources.

The EpiDoc markup was especially functional to the research questions in the encoding of the Greek texts inside `<div type="edition">`, where its semantic nature proved to be very helpful for the extraction of the institutional elements and for the analysis of variations in their type and function or sphere of competence, avoiding preconceived generalizations and valuing the specificity of each occurrence. In fact, the markup of the institutional elements was based on the use of `<rs type="institution">` along with a customized combination of focused attributes: `@subtype` for specifying their typology (such as assembly or board), `@role` for specifying their function or field of action (such as voter or dedicant), `@ref` for specifying their political entity (usually their *polis*), `@key` for facilitating their indexing through normalized forms of their names. A complete example of the markup of an institutional element can be found in Figure 1 below.


```
<rs type="institution" subtype="official" role="eponym" key="damiorgos" ref="#olous">ἐπὶ
  <w lemma="δαμιοργός">δαμιοργοῦ</w> <persName type="attested"
  key="Leukos"><name nymRef="Λεῦκος">Λεύκου</name></persName></rs>
```

Figure 1. An excerpt from the XML EpiDoc text of inscription *I.Cret.* I 22 4 A (l. 31)

In addition to `<rs>`, some other core TEI EpiDoc elements were used: `<w lemma="">` for the lemmatization of institutional and other relevant terms, `<placeName type="" ref="">` for toponyms and ethnic adjectives, `<persName type="" key="" ref="">` and `<name nymRef="" type="">` for prosopographic and onomastic elements (for officials, honoured individuals, foreign rulers and theonyms).

Overall, the semantic markup involved 8,162 lemmata (`<w>`), 4,353 institutional elements (`<rs>`), 2,633 toponyms or ethnic adjectives (`<placeName>`), 1,694 anthroponyms (`<name>`) and 1,651 prosopographical elements (`<persName>`).

In addition to the epigraphic collection, the research outputs also include the creation of two TEI catalogues: one relating to the political entities of Crete (*poleis*, *koina*, dependent communities, extra-urban sanctuaries and hegemonic alliances); another one relating to the attested Cretan institutions (comprising assemblies, boards, officials, associations, civic subdivisions, social statuses, age classes, months, festivities and other celebrations, institutional practices, institutional instruments and public spaces).

3 Benefits of Using EFES and EpiDoc

The EpiDoc Front-End Services (EFES) are an open source customisable tool for the online publication of ancient documents in EpiDoc XML, inscriptions in primis (Bodard and Yordanova, 2020)[2]. It is the EpiDoc specialisation of Kiln, an analogous framework for publishing collections of TEI XML documents, from which it was forked in 2017[3]. The main strengths of EFES, as well as of its ancestor Kiln, are its comprehensiveness, ease of use and high customizability, which make it possible to quickly create a Web site provided with indices, textual search and browse facilities even from persons without advanced IT skills. In fact, the aim of EFES is to allow the production of such outputs also from small projects whose teams neither include IT experts devoted to the development of Web sites nor have funds to be dedicated to that purpose.

The specific case of *Cretan Institutional Inscriptions* is particularly emblematic of the benefits deriving from the use of EFES, being it the first project carried out by a single person to have used it. Despite the awareness of the importance and usefulness of a collaborative approach to research, the case was that of an individual doctoral project to be completed in three years with no external support, thus perfectly matching the target of users expected by EFES. The timing of the first release of EFES in September 2017, at the end of the first year of the Ph.D. research, was providential and allowed the creation of the Web site in the remaining two years.

In particular, with some customization of the provided XSLT stylesheets, from the EpiDoc markup of the inscriptions it was possible to generate several custom thematic indexes, recording the occurrences of institutional elements, relevant lemmas, prosopography, onomastics, toponyms, ethnics and theonyms. These indexes, especially the one relating to the institutions, are displayed in a tabular format, where all the pieces of information included in the markup are collected in separate columns and can be easily combined and compared with each other.

Besides the indexes, the EpiDoc encoding allowed the creation of very specific search filters, thanks to which the inscriptions can be browsed not only according to their traditional metadata (type of document, type of support, date, provenance, current location, bibliographic reference), but also on the basis of the name (e.g. *agela*, ≪herd≫), type (e.g. tribe) and role (e.g. decreer) of the institutional elements and of the name of the places and divinities mentioned.

Another benefit deriving from the EpiDoc encoding is the high level of accuracy of the textual searches performed on the collection, which ignore all the extremely frequent diacritics due to the epigraphic editorial conventions and can be further refined by including the lemmatized base forms of the terms.

From a linguistic point of view, lemmatization proved to be a rather significant component of the work, considering the nature of the encoded inscriptions. In fact, the texts present themselves as particularly remarkable from a linguistic and dialectal perspective. The inscriptions, in the Cretan Doric dialect, contain a large number of lemmas – mostly terms attested only in epigraphic sources – that are either difficult to be found or completely absent in the main lexica of ancient Greek (such as the renowned Liddell-Scott-Jones Greek-English Lexicon; they are often mentioned within the corresponding Attic lemmas only)[4], including some *hapax legomena* (such as δωροτελέω, συνβολήτρα) that are unique from

---

a lexical or morphological point of view.

Therefore, the lemmatization carried out was aimed at preserving the dialectal particularities of the Dorian Cretan language, including the presence of letters such as digamma (Ϝ) and qoppa (Ϙ), attested the latter also in alternation with kappa in cases such as ϙόσμος/κόσμος, ϙοσμέω/κοσμέω, ὄρϙος/ὄρκος, πρόϙοος/πρόκοος.

In addition to the linguistic peculiarities, the epigraphic nature of the database is also an aspect that made lemmatization particularly useful. In fact, the high fragmentary nature of the texts led to the presence of very frequent diacritical marks, particularly in relation to the indication of lacunae, additions and uncertain readings (square brackets, dashes, dots, underdots, question marks; see e.g. the text in Figure 2), which severely limits the search possibilities using the simple text-by-string search method. The combined action of EpiDoc markup and lemmatization made it possible to ignore diacritical marks in text searches as well as to perform text searches based on the lemmatized text.

35. Iscrizione edificatoria degli *eunomiotai* di Aptera

Tipologia documentaria: iscrizione edificatoria

Supporto: sconosciuto

Datazione: II secolo a.C.

Provenienza: Aptera

Collocazione attuale: iscrizione probabilmente perduta

Edd. Haussoullier 1879, p. 436, n. 10; *SGDI* 4949; Guarducci 1933, n. 5; *IC* II 3 21 ↗PHI.

```
          [- - - - - - - - - -]
          [- - -]ν Εὐρυμήδης Ἀνδι[- - -],
          [- - -]χος Ἀρχέτω, Ὀρσικλῆ[ς - - -],
          [- - -]σκος Ὀξυμ[άχ(?)]ω,
          [- - - Ἀ]λκιμένη ἐπεμελῆθε[ν - - -]
     5    [- - -]ρσιος καὶ τᾶν λοιπᾶν πα[σᾶν - - -]
          [- - -] μέστα ἐπὶ τ[ . ] εὐνομιῶτ[αν - - -]
          [- - -]ον.
```

6: εὐνομιῶτ[αν] *IC* in apparato; εὐνομιωτ[- - -] *IC*.

Il collegio degli *eunomiotai*, attestato a Lato, Olous e forse Knossos con la denominazione di εὐνομία, a Polyrrhenia con il nome di συνευνομιῶται (cf. *IC* II 23 9), anche ad Aptera – come a Lato – si occupa dell'edificazione o della manutenzione di strutture pubbliche della città (cf. Guarducci 1933, pp. 201-205, Chaniotis 2008, pp. 114-116). Le competenze religiose che l'istituzione mostra di avere altrove – a Lato e Polyrrhenia – suggerirebbero che ciò che è stato costruito o ristrutturato dagli *eunomiotai* sia uno o più edifici sacri, come sembra indicare l'espressione τᾶν λοιπᾶν πα[σᾶν], verosimilmente riferita all'oggetto di ἐπεμελῆθεν.

Quanto alla composizione del collegio, nella parte conservata dell'iscrizione è possibile identificare almeno cinque *eunomiotai*; il loro numero, tuttavia, potrebbe essere maggiore, similmente a quanto avviene a Lato, dove l'iscrizione completa *IC* I 14 2 ne attesta nove, mentre il testo quasi completo di *IC* I 16 21 ne ricorda sette.

Elementi istituzionali o altri termini rilevanti: *epimeletes* (epimeleomai), *eunomiotai*.

Figure 2. An inscription of the collection, *I.Cret.* II 3 21

Istituzione	Termine attestato	Individuo	Tipologia	Ambito / Ruolo	Località	Periodo	Occorrenze
Damiorgos	δαμιοργέω	Eteon f. Archetos	Magistrato o funzionario	Dedicante	Aptera	E	seg_60_984.2
Damiorgos	δαμιοργός	A-	Collegio	Eponimo	Olous	E	seg_23_548.2
Damiorgos	δαμιοργός	Arsias	Magistrato o funzionario	Eponimo	Olous	E	[ic1_22_4.B.61]
Damiorgos	δαμιοργός	Autosthenes	Magistrato o funzionario	Eponimo	Olous	E	[ic1_22_4.B.1] ic1_22_4.B.19
Damiorgos	δαμιοργός	Botrynos	Collegio	Eponimo	Olous	E	[seg_23_549.1]
Damiorgos	δαμιοργός		Collegio	Eponimo	Kydonia	E	[seg_41_731.3]
Damiorgos	δαμιοργός		Collegio	Eponimo	Polyrrhenia	E	ic2_23_7.B.1
Damiorgos	δαμιοργός	Leukos	Magistrato o funzionario	Eponimo	Olous	E	ic1_22_4.A.31 ic1_22_4.A.35
Damiorgos	δαμιοργός	Onasandros f. Parmenon	Collegio	Eponimo	Polyrrhenia	E	ic2_23_7.A.1
Damiorgos	δαμιοργός	Sosos f. Tasskos	Collegio	Eponimo	Polyrrhenia	E	[ic2_23_8.1]

Figure 3. An excerpt from the index of institutional elements

4 *Cretan Institutional Inscriptions* at VeDPH

The Venice Centre for Digital and Public Humanities (VeDPH) was inaugurated in 2019 and belongs to the Department of Humanities of the Ca' Foscari University of Venice (UNIVE-DSU). The mission of the centre is the promotion of interdisciplinary methodologies for "the collaborative development of durable, reusable, shared resources for research and learning" (https://www.unive.it/pag/39289).

VeDPH not only promotes and funds new projects, but is also in charge of legacy projects developed at UNIVE-DSU over the past decades.

Since its foundation in 2019, the Venetian Detached Research Unit (URT) of the Institute for Computational Linguistics «A. Zampolli» of the National Research Council of Italy (CNR-ILC) has been working in collaboration with VeDPH within the *Archipelago DPH* project in order to ensure that the creation of new digital resources and the maintenance of the legacy ones are effectively durable, reusable and shared.

According to this vision, CLARIN-IT provides the necessary know-how through webinars and seminars at VeDPH, a state-of-the-art technological infrastructure to develop and test the new prototypes created by VeDPH affiliates, a suitable Web infrastructure for (permanently or temporarily) hosting the legacy projects developed at UNIVE-DSU and all the CLARIN tools and strategies to make new data and legacy data as FAIR[5] as possible.

In this context, *Cretan Institutional Inscriptions* gave the opportunity to test this model of collaboration between VeDPH and CLARIN-IT through its Executing Institution, namely CNR-ILC.

---

[5]FAIR is an acronym for Findable, Accessible, Interoperable and Reusable (https://www.go-fair.org).

5 *Cretan Institutional Inscriptions* at CLARIN-IT

The Italian Consortium CLARIN-IT[6] has a strong interest in the field of Digital Classics and aims at including a large part of resources for historical languages in its repositories (for an overview of the consortium at a whole see Nicolas et al., 2018). The Repository[7] of the ILC4CLARIN Centre[8] already contains important resources, such as the ALIM archive (Ferrarini, 2017), presented also at the CLARIN Conference 2020 (Boschetti et al., 2020), as well as many resources[9] from the ERC project "LiLa: Linking Latin"[10]. The deposit and description of *Cretan Institutional Inscriptions* follow this path and increase the number of resources for Ancient Greek available in the CLARIN Virtual Language Observatory (VLO). Indeed, if VLO is queried for some of the main keywords used to describe *Cretan Institutional Inscriptions* in the ILC4CLARIN Repository (such as, for instance, *epigraphy* or *epigraphic*), only few resources are returned. The authors hope that *Cretan Institutional Inscriptions* pave the way for other similar initiatives to be described in Italy and other countries belonging to CLARIN-ERIC.

5.1 Organizational Aspects

In this section, the organization of the *Cretan Institutional Inscriptions* resources within the Italian node of CLARIN is described. The Web site dedicated to the dataset created within the Ph.D. project (`https://www.clarin-it.it/cretaninscriptions`)[11] is hosted by CLARIN-IT in the framework of the activities supporting projects and events in the Social Sciences and Humanities sector. The Web application to interact with the dataset (`https://ilc4clarin.ilc.cnr.it/cretaninscriptions`)[12] is offered as a service of ILC4CLARIN, the first CLARIN B-Centre of CLARIN-IT. The Persistent Identifiers (DSpace Handles) relating to the description sheets drawn up by the authors for the dataset [13] and the Web application[14] are available under a free license in the ILC4CLARIN Repository.

The strategy behind this organization is the following one: the *Cretan Institutional Inscriptions* collection has its GitHub repository (`https://github.com/IreneVagionakis/CretanInscriptions`), which contains the dataset, the software for the search engine, some customization, the licenses of use and the releases of the dataset. The authors decided to periodically deposit the various releases of the dataset in the ILC4CLARIN Repository as well, so that scholars can access the complete data without using the search engine. On the one hand, this approach guarantees the versioning of the dataset and the long term preservation of the data; on the other hand, it shares *Cretan Institutional Inscriptions* with the CLARIN community.

6 Dockerization

In the last decades a growing number of humanists exploited the digital ecosystem to improve their research and disseminate their results. Textual scholars learnt to build digital resources, to develop computational tools and to use research infrastructures, as well as to reuse generic frameworks, originally implemented for other domains of knowledge, but adapted to specific needs in Digital Humanities projects. Before the advent of the Docker system, two scenarios characterized the technological choice for new digital initiatives: 1) adopting resources and an infrastructure provided by some technological centre; 2) adopting proprietary resources and technologies within a single project. Both scenarios have advantages and disadvantages. As far as the first scenario (based on technological centres) is concerned, benefits lie mainly in soundness, availability, homogeneousness, maintainability and security concerns about the

---

[6]`https://www.clarin-it.it`.

[7]`https://dspace-clarin-it.ilc.cnr.it/repository/xmlui`.

[8]`https://ilc4clarin.ilc.cnr.it/en`.

[9]The resources described are available in the Virtual Language Observatory (VLO): `https://vlo.clarin.eu/search?4&q=CIRCSE`.

[10]`https://lila-erc.eu`.

[11]The PID for the URL is `http://hdl.handle.net/20.500.11752/1002`.

[12]The PID for the URL is `http://hdl.handle.net/20.500.11752/1003`.

[13]`http://hdl.handle.net/20.500.11752/OPEN-548`.

[14]`http://hdl.handle.net/20.500.11752/OPEN-550`.

technologies adopted. These advantages have the inevitable cost of experimenting new (if not immature) approaches, methods and tools. Actually, on the technological side, research activities were severely limited by these constraints. Indeed, the digital framework was often defined a priori and, therefore, hardly negotiable. For instance, relational databases were used alongside applications written in specific programming languages to build the Web Graphical User Interfaces (GUIs). As for the second scenario (proprietary technologies and resources), the flexibility offered in experimenting new methodologies, practices, models and resources often gave profitable outcomes. However, the innovative results were usually confined to a single initiative and almost never acknowledged within other similar initiatives. These last disadvantages were mainly due to the lack of well-defined standards, stable protocols, sufficient reliability and online availability of services.

A new scenario to build, share and run applications emerged recently. Indeed, Docker technology enables new DH applications to be packaged, published and deployed in single flexible units, called Docker images. The introduction of the Docker environment keeps the advantages offered by the aforementioned legacy scenarios but, at the same time, gets rid of their disadvantages. In fact, Dockerization provides a way to guarantee the availability of services, the maintainability of tools, the portability of applications, the flexibility of the environment, the scalability of the architecture and the consistency and reusability of the results of a digital project. In addition, the adoption of the Docker tool as the main deployment technology has two other advantages: 1) the DevOps methodology for agile development and continuous integration; 2) microservices for creating coherent components exposing data and services. Thanks to these properties, the typical problems of updating technical and data dependencies are successfully overcome, significantly improving data security.

Within this context, CLARIN-IT supported the *Cretan Institutional Inscriptions* team to adopt the DevOps methodology for the development, building and deployment of its digital resource (i.e. the collection of epigraphic texts encoded according XML/EpiDoc schema) as well as the Web application to interact with it (i.e. the EFES Web platform). This methodology increases *speed* of the development process of *Cretan Institutional Inscriptions* in each of its phases (text encoding, software implementation, testing and deployment on production servers). It also reduces *unexpected issues*, mainly due to different operative systems or different versions of software libraries. In addition, the DevOps methodology facilitates the *visualization* of the *Cretan Institutional Inscriptions* running components[15] through a Web container manager. Thus, the CLARIN-IT team adopts Docker technology in order to implement a sound development workflow as well as long term preservation policies for the applications hosted, increasing their collaborative implementation and portability.

With the aim of maximizing the benefits of this methodology, CLARIN-IT runs the *Cretan Institutional Inscriptions* EFES application as a *stack of dockers containers*, managed through the Rancher environment, an open source Web container manager[16]. In this way, the different technologies and devices used (such as the Operative System, the Java Virtual Machine and the Web server) are separated from the technologies and devices of other applications running on the same server.

7 *Desiderata* for Future Developments

As with all work, although a lot has been done, a lot remains to be done. Among the main *desiderata* for the future there is undoubtedly the integration of data visualization systems through maps, relational graphs and timelines. This could be allowed by several existing open source tools and JavaScript libraries, among which Leaflet, Cytoscape.js and, above all, Palladio stand out.

Leaflet[17], an open source JavaScript library for creating interactive maps, can be easily integrated inside EFES[18], providing a filterable map view of the various institutional elements marked-up in the inscriptions.

---

[15]https://goto.docker.com/rs/929-FJL-178/images/20150731-wp_docker-3-ways-devops.pdf.

[16]https://www.docker.com, https://rancher.com.

[17]https://leafletjs.com.

[18]A precedent for this is the integration of Leaflet into the EFES-based site of *Fiscal Estate in Medieval Italy: Continuity and Change (9th-12th centuries)* project, *Fiscus*, https://fiscus.unibo.it.

Cytoscape.js[19], an open source JavaScript library for creating interactive graphs, can also be easily integrated inside EFES[20], providing – through the usage of its force-directed fCoSE layout[21] – highly dynamic relational networks.

Compared to the JavaScript libraries just mentioned, Palladio[22] has the advantage of being an open source application developed specifically for the visualization of complex data resulting from historical research. Its visualization options include a Map view, a Graph view, a Table view and a Gallery view. In particular, its Map view is remarkably rich: in addition to the map itself, it includes the possibility of filtering data also according to a chronological criterion through the Timeline and Timespan filters. Such functionalities could allow to enhance the graphical performance of what can be obtained through a combination of multiple EFES search filters, which already allow a cross-search of geographical and temporal data and information relating to the institutions mentioned (as well as other variables relating to the epigraphic sources and their contents). The following figures (Figs. 4-8) offer some examples of what types of visualizations can be obtained trough the Palladio Web application, using respectively its Map view (Fig. 4), its Map view with a Timespan filter (Figs. 5-6) and its Graph view (Figs. 7-8) applied to some case studies relating to the institutions attested as decreers and dedicants and to the full picture of the attestations of tribes and *agelai*.

In addition to the Map, Graph and Timeline views, in the future it would be useful to develop some other features, such as the implementation of an API interface allowing an export of the data also in RDF[23] or JSON formats and, from the point of view of contents, the inclusion of the English translations of the inscriptions.

Figure 4. Geographical distribution of the attestations respectively of *kosmoi* (top left), *boule* (top right), *demos* (bottom right), and *polis* (bottom left) as decreers (visualization on Palladio - Map view)

---

[19]https://js.cytoscape.org.

[20]A precedent for this, again, is the integration of Cytoscape.js into the EFES-based site of *Fiscal Estate in Medieval Italy: Continuity and Change (9th-12th centuries)* project, *Fiscus*, https://fiscus.unibo.it.

[21]https://github.com/iVis-at-Bilkent/cytoscape.js-fcose.

[22]*Palladio. Visualize complex historical data with ease*, http://hdlab.stanford.edu/palladio, https://github.com/humanitiesplusdesign/palladio-app.

[23]Kiln and EFES already provide some basic functionalities for handling RDF data: see https://kiln.readthedocs.io/en/latest/tutorial.html#querying-rdf and https://kiln.readthedocs.io/en/latest/rdf.html.
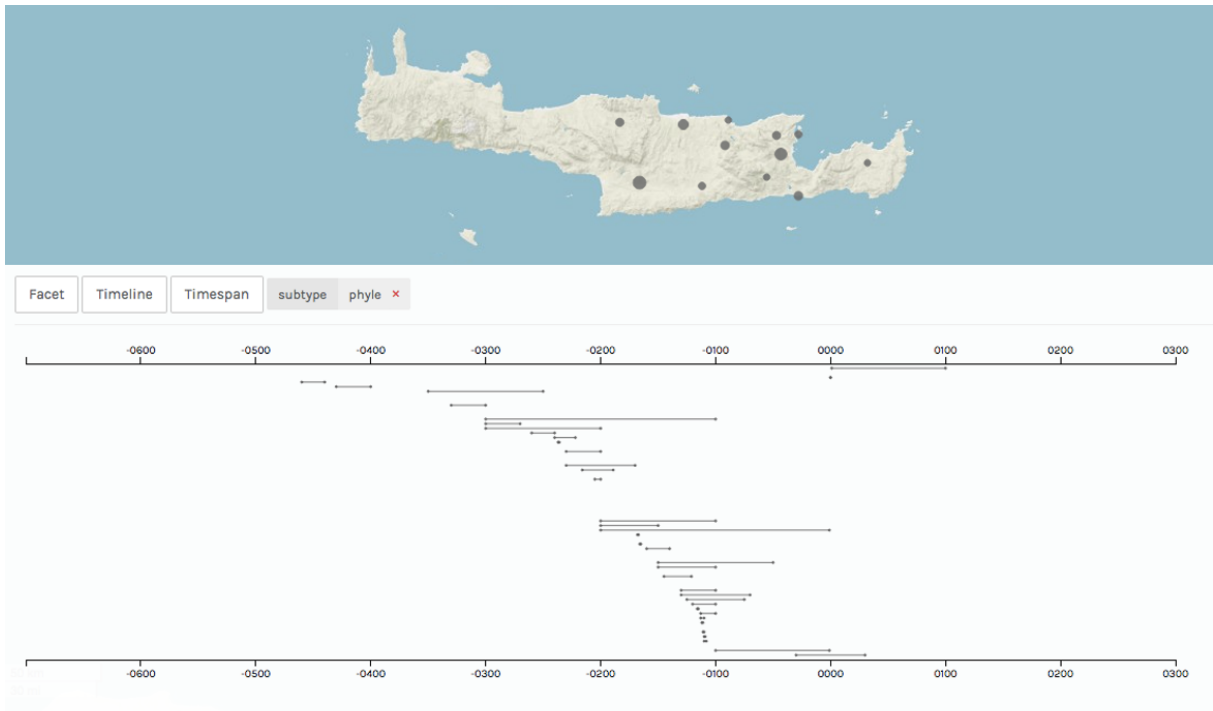
Figure 5. Geographical and chronological distribution of the attestations of tribes (visualization on Palladio - Map view including a Timespan filter)

Figure 6. Geographical and chronological distribution of the attestations of *agelai* (visualization on Palladio - Map view including a Timespan filter)
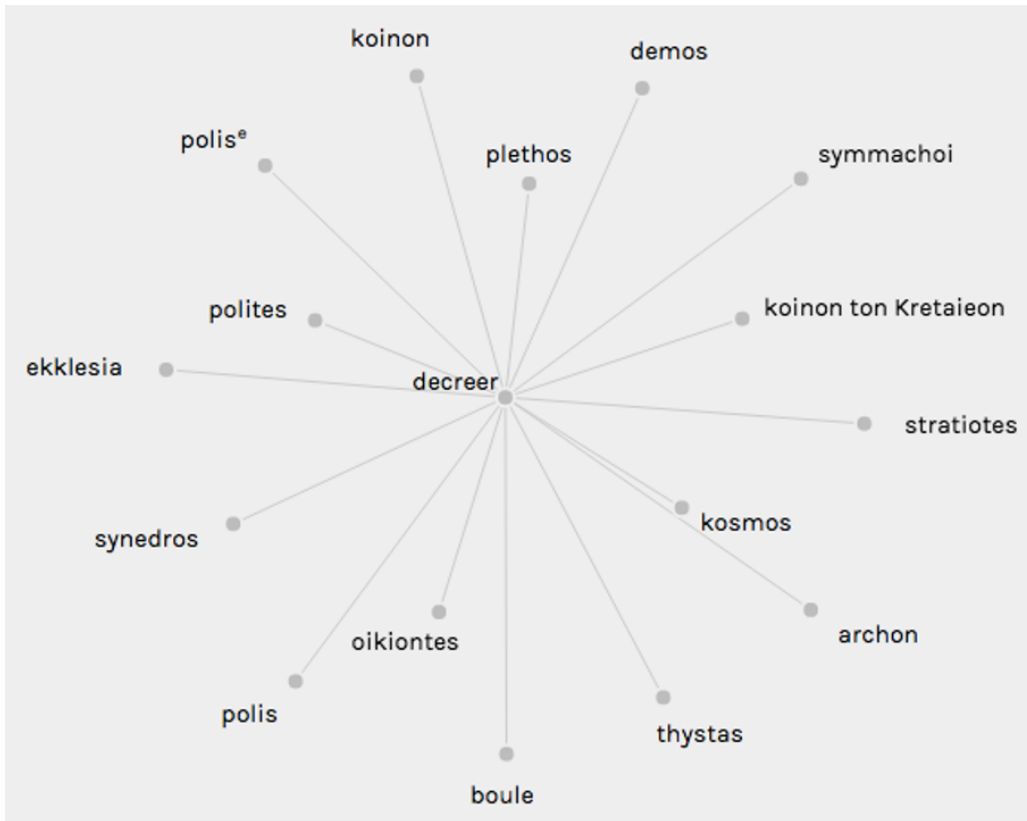
Figure 7. Institutions attested as decreers (visualization on Palladio - Graph view)
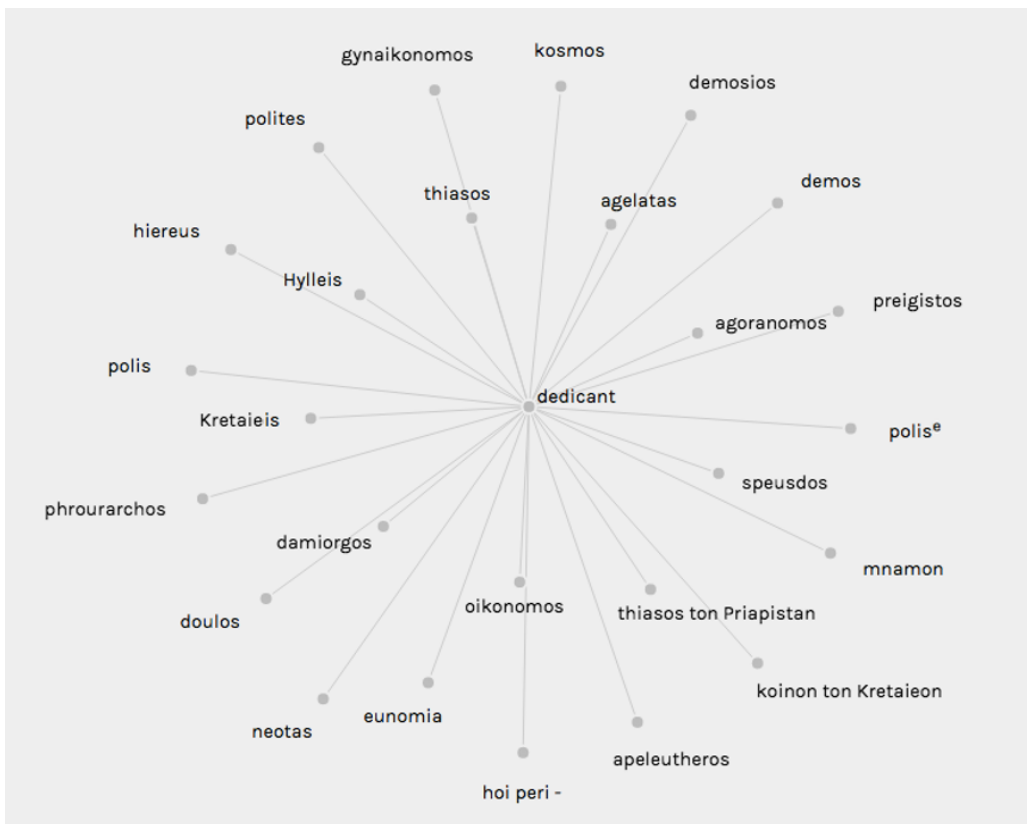

Figure 8. Institutions attested as dedicants (visualization on Palladio - Graph view)

Selected papers from the CLARIN Annual Conference 2021

8 Digital Epigraphy and the CLARIN Community: The Role of K-Centres

The original focus of CLARIN is on language resources, such as dictionaries and treebanks, and textual resources, such as literary corpora, used in particular as a test to study linguistic phenomena in a context. Recently, however, the interest of the CLARIN community in the domain of Digital Humanities significantly increased, as it emerged in the last CLARIN Annual Conferences, in CLARIN initiatives such as the CLARIN Café series and in initiatives supported by CLARIN such as the last Annual Conference of the Italian Association of Digital Humanities and Digital Culture (AIUCD)[24].

By widening its horizons, the CLARIN community has enlarged its interdisciplinary competence towards the digital representation of cultural artifacts, such as ancient and modern epigraphs, Medieval manuscripts, ethnographic oral archives etc. The processing of fragmentary texts, Information Retrieval applied to the text of scholarly editions with variant readings in the critical apparatus and the mapping of linguistic annotations on facsimile of historical documents are just a few examples of the new challenges that the CLARIN Infrastructure has to face.

On the one hand, CLARIN helps the disciplinary communities of digital philologists, digital epigraphers and digital historians to share best practices relating to digitalization processes in Digital Humanities. In particular, these best practices concern the digitization of images, the transcription of primary sources, the encoding of texts, the annotation of linguistic, stylistic and philological phenomena, the visualization of scholarly editions, the creation of search engines and the publication and reuse of literary and documentary data. Indeed, the use of language resources is transversal to different disciplines. New epigraphic corpora enrich the repertories of attested inflected forms, which feed up the spellcheckers used to improve the OCR post-processing applied to corpora of secondary sources etc. On the other hand, CLARIN has the opportunity to extend the infrastructure in order to meet the specific needs of the humanists, in particular the interrelation between text and document, the non-sequentiality of the variorum edition and the diachronic perspective.

CLARIN Knowledge Centres[25] are suitable instruments to share interdisciplinary competence that, in order to accomplish each phase of a digital project, has to be coordinated. Among them, the new Knowledge Centre for Digital and Public Textual Scholarship (DiPText-KC)[26], which is the second CLARIN K-Centre of CLARIN-IT, is specifically devoted to the scientific representation of literary and documentary texts as well as to their elaboration and publication.

Cretan Institutional Inscriptions gave us the opportunity to test the organization of DiPText-KC in order to put the project leader in close contact with a team of experts in the DevOps technologies distributed among Pisa, Venice and other European DH Centres. Indeed, Digital Epigraphy is an interesting test bench for verifying the cross-fertilization between language infrastructures and Digital Humanities, since, in order to study the historical context etc., it involves the description of the material support, the study of (possibly fragmentary) texts and references to the secondary literature.

Being a solid project based on EpiDoc, *Cretan Institutional Inscriptions* adopts standard protocols for encoding its textual resources. DiPText-KC addressed the project leader to publish the software source code and textual data separately and under open licenses as well as to make data and metadata compliant to the FAIR principles in order to maximize their reusability.

9 Conclusion

With this paper the authors aimed to highlight how CLARIN-IT is opening up to areas of Digital Humanities that, until few years ago, were not central to the CLARIN world. Indeed, CLARIN has always demonstrated interest not only in the living language but also in literary texts, as it is evidenced by the high number of corpora in the CLARIN repositories. However, its interest in Ancient Greek and Latin Digital Epigraphy and Papyrology is more recent. A search in VLO with the keyword "epigraph* Greek OR Latin" or with the keyword "papyr* Greek OR Latin" shows that most of the records is very recent (around 2020). These disciplines give CLARIN the possibility to reason about new use cases to extend

---

[24]https://aiucd2021.labcd.unipi.it/en/home-english.
[25]https://www.clarin.eu/content/knowledge-centres.
[26]https://diptext-kc.clarin-it.it.

the metadata set with information needed by epigraphers and papyrologists, such as the geolocalization of the document, the presence or not of a facsimile in the collection, the literary genre to which texts belong, if a text is in poetry or prose etc.

Federated Content Search can also benefit from an expansion to new types of documents, especially in the way of citing the occurrences returned by a query and in the way of dealing with fragmentary texts.

The authors hope that a project such as *Cretan Institutional Inscriptions*, which contain both a critical edition and a specific visualization tool, can contribute to widen the bridge between purely linguistic interests and other areas of the Humanities inside the CLARIN world, where this important connection is often missing. It is to this end, in fact, that the authors are implementing a DevOps methodology and a Docker infrastructure aimed at hosting such a kind of initiatives.

10   Acknowledgment

References

Bodard, G. and Yordanova, P. 2020. Publication, Testing and Visualization with EFES: A Tool for All Stages of the EpiDoc XML Editing Process. *Studia Universitatis Babeș-Bolyai Digitalia*, 65(1):17–35, Dec.

Boschetti, F., Del Gratta, R., Monachini, M., Buzzoni, M., Monella, P., and Rosselli Del Turco, R. 2020. "Tea for Two": The Archive of the Italian Latinity of the Middle Ages Meets the CLARIN Infrastructure. In *CLARIN Annual Conference 2020*, pages 121–125. CLARIN-Virtual Edition.

Elliott, T., Bodard, G., Mylonas, E., Stoyanova, S., Tupman, C., and Vanderbilt, S. e. a. 2020. EpiDoc Guidelines: Ancient Documents in TEI XML (Version 9.2).

Ferrarini, E. 2017. Alim ieri e oggi. *Umanistica Digitale*, 1(1).

Liddell, H., Scott, R., and Jones, H. S. 1925-1940. *A Greek-English Lexicon. Ninth Edition*. Oxford Clarendon Press.

Liddell, H., Scott, R., and Jones, H. S. 2011. *The Online Liddell-Scott-Jones Greek-English Lexicon*. Thesaurus Linguae Graecae.

Nicolas, L., König, A., Monachini, M., Del Gratta, R., Calamai, S., Abel, A., Enea, A., Biliotti, F., Quochi, V., and Stella, F. 2018. CLARIN-IT: State of Affairs, Challenges and Opportunities. In *Selected Papers from the CLARIN Annual Conference 2017*.

Vagionakis, I. 2021. Cretan Institutional Inscriptions. A New EpiDoc Database. *Journal of the Text Encoding Initiative*.