Digital Classics:
A Survey on the Needs of Ancient Greek Scholars in Italy

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

Anika Nicolosi
Dip. DUSIC
Parma University, Italy
anika.nicolosi@unipr.it

Alberto Stefanini
Novareckon
Novara, Italy
alberto.stefanini@virgilio.it

Abstract

This paper presents and discusses the findings of a survey carried out in order to assess the use of digital resources and digital technologies with respect to work in Ancient Greek scholarship, as well as to identify the factors that are likely to constrain its use and to elicit needs and requirements of Ancient Greek scholars in Italy. The survey is in line with the principles behind the recent user engagement strategy developed by CLARIN-ERIC and constitutes one of the national efforts undertaken by CLARIN-IT to contribute to the wider impact of CLARIN on Digital Classicists.

1 Motivation and Background

Interest in language technologies is never been as strong as it is now. The main conferences in Digital Humanities are seeing more participation by computational linguists while at Computational Linguistics conferences the humanities and social sciences are now an important line of research. The necessity to meet the needs of an audience of different users opens new challenges for language technologies: easily usable tools, adaptable to different types of content become crucial. Texts in the humanities can often be heterogeneous and the quality of resources, in particular the quality of digital editions of texts, is receiving increasing attention.

CLARIN-ERIC (www.clarin.eu), the European research infrastructure for the Humanities and Social Sciences, was created to promote the development of technological solutions aimed at making language resources available to scholars, researchers, students of the Humanities and Social Sciences through a unified and standardized mode of access to data and computational tools. This involves making available digital repositories where data, corpora, lexicons, tools and language services are catalogued, stored, retrieved and used in a simple and intuitive way by users. It represents the perfect solution for bringing together producers and developers of language technology with its users. The idea that user needs should be a central part of the design and development process of any ITC infrastructure is not necessarily new, and pointing out that the CLARIN audience are Humanists and Social Scientists is not enough. There is a wide range of people working within the academic sector; one set might be more proficient in one area but less in another, so there are different needs and different ways of engaging with users.

In order to study the current interest in the use of digital resources and related tools in one specific context of the Humanities, Ancient Greek scholarship, we performed a survey based on a questionnaire to ascertain the current practice and the related needs of a group of practitioners in the field. At a larger scale, the work represents one of the first attempts undertaken within the context of CLARIN-IT (Monachini and Frontini, 2016) to contribute to the wider impact of CLARIN on the specific Italian community of those interested in the application of Digital Humanities to the field of Classics and to ancient world studies. Similar surveys were sporadically performed during the last decade, pertaining to fields that do not strictly concern Digital
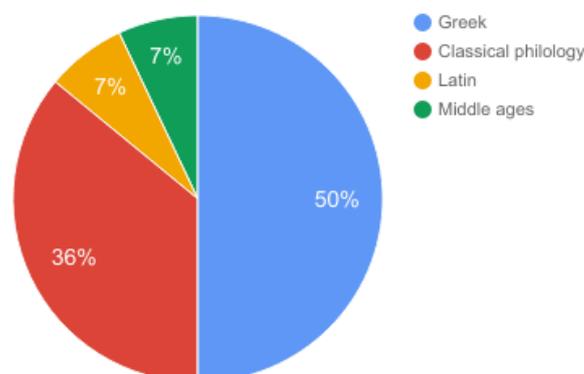
Classics (DC) – although the general outcome of their remarks is relevant, and it is evident that they may also be applied to DC resource design. For our purposes it is enough to quote the key outcomes of the studies by (Toms and O'Brien, 2008), (Audenaert and Furuta, 2010) and (Warwick et al., 2008).

A point worth remarking is that the above studies concern a wide spectrum of scientific interests within the digital humanities' realm and involve scholars from several countries, and mostly in the anglophone world. For this reason, our study focuses on the specific scientific community dealing with Digital Classics; specifically on a restricted sample of Italian digital humanists with focus of interest on ancient Greek philology.

2 Consultation

Our survey aims to evaluate the impact of digital techniques within the specific reference community of Ancient Greek scholars in Italy. It is a relatively small field but this area of study retains great interest in Italy where it also looks back to a great tradition; it also includes university students and schoolteachers. Moreover, Italian Ancient Greek scholars are an active part of an international community (including especially Europe, North and South America) and this field of studies has great perspective because its research spectrum involves also Latin, Ancient History and Philosophy, and Classics in general. Finally, it is important to remember that Ancient Greek studies are an essential part of our Occidental Cultural Heritage, and it is crucial that the highest number of people knows these texts and their contents. For these reasons, the Italian Ancient Greek community is an excellent field to test new opportunities offered by Digital Humanities about knowledge and quality in transmission of ancient texts. The questionnaire was sent to a selected group of Italian researchers whose main focus of study - although their interests span over a broader area, encompassing Greek and Latin literature - is research into and teaching of the Ancient Greek language. The survey was performed from May to September 2016 and it is now available on-line at http://www.clarin-it.it/it/content/sondaggio-current-practice-digital-classics-tools. Our survey was supplementary to a master degree thesis discussed at the University of Parma (Stefanini, Nicolosi and Monachini, 2017), for this reason we needed quick answers and targeted people known to be active in Digital Classics with focus on Ancient Greek language. At this research stage, the sample is numerically consistent with the survey target because of its specific expertise (Ancient Greek) and of its number, that is about 10 percent with respect to the initial potential target population (see Figure 1).

Figure 1 - Main interest

2.1 Questionnaire key points

The survey focused on the digital resources and tools needed to support a excellent and usable digital edition of an ancient text. For this reason, first we ask to applicant to identify and judge the tools they use and know. The questionnaire included four sections:

- Current practice with Digital Classics tools (where applicants are asked to specify which typology of tools they use in their studies)

- Advanced functions (where applicants are asked to specify which available/unavailable function is/would be more useful)

- Utility of digital resources and tools for philology (where applicants are asked to qualify the importance of four key functionalities):

  - to match the different versions of a text with mention of the text sources and provide critical apparatuses
  - to make available diverse interpretations supported by computer-based linguistic/stylistic analysis
  - to make available a digital copy of the primary source (code, papyrus, epigraph etc.)
  - to provide one or more translations of a referred text in contemporary languages

  An evaluation of the usefulness of testing/experimenting with one or more of these functions is required. Then, we ask the specific expertise of the applicant.

- Quantitative evaluation: the questionnaire asks to assess in a 1-5 scale the usefulness of a set of functions considered crucial for digital editions dealing with classical studies (Table 1 shows the functions list and their average resulting judgment).

Function	Average score
primary sources	4.00
variants	3.57
critical editions	3.57
collaborative hypotheses	3.36
logic analysis	3.36
conjectures	3.29
metric analysis	3.14
translation	3.00
reviewing variants	3.00

Table 1: Quantitative Evaluation

2.2 Overview of replies

The key outcomes of the survey may be summarized as follows.

Digital Libraries. All respondents make use of Perseus and TLG. Trismegistos, Perseids and Alpheios, Musisque Deoque and PHI are often mentioned. All these resources receive some criticism concerning their coverage and/or their availability.

Functionalities. Search by lemma, morphologic analysis (and availability of on-line dictionaries) are deemed the most useful functions, although they are not available in most cases.

Syntactic Analysis. There is a requirement to improve syntactic analysis, improve search mechanisms, such as syntax-based search, and make available Application Programming Interfaces (API) to several functions. Some users ask to provide hypertext links to the referenced texts.

Variants. Testing and experimentation is most required for provision of several text variants, logic-syntactic analysis, text search, digital copies of the primary source, translation in contemporary languages. One of the replies says rather optimistically that those activities are current practice already.

Primary Sources. Quantitative replies highlight the need to provide primary sources, followed by a set of functions with a close score (variants, logic analysis, conjectures and text hypotheses, metric analysis).

Alia. The rather low score obtained by translations and interactive support to reviewing variants may be due to the bearing of the sample towards research rather than teaching and is to some extent contradicted by qualitative replies.

Finally, many respondents pointed out that important research needs in the field are models and software for authoring, editing, indexing and presenting a digital edition, and which way to link them together with the available resources and improve them. Digital editions may provide scholars with copious, much complete materials to ease their research and their studies, with a deeper insight into useful research methods.

3 Conclusions

To be successful, tools and services must be compatible with researchers' workflows, which are often discipline-specific, and they must be easy to use. Researchers must be primary actors describing their expectations from tools and services in support of their studies, and public feedback is subordinate to the community's interests.

As the PARTHENOS (www.parthenos-project.eu) Report on User Requirements (2016) also states, it is worth remembering that the humanities' community shows enormous heterogeneity Identifying user requirements is crucial for enhancing existing resources and create new resources and tools, to. For this reason, our survey defines needs and requirements of Greek scholars (and, in the future, of Classical scholars at large) thus addressing a set of R&D priorities priorities that may be the base for establishing a consistent research and innovation agenda for Digital Classics. Moreover, our study may also help identifying gaps and drive the development of new technologies for ancient studies at large.

Research infrastructures may indeed provide significant developments in the field of Digital Humanities because they offer opportunities to store, develop and share data. In addition, e-infrastructures are the perfect framework with which to spread knowledge about good practices related to a discipline (also with attention to possible educational aspects) and offer the opportunity to share tools. CLARIN delivers services that, once integrated at an earlier stage of the process, may facilitate research tasks, and provide users with a variety of analitic tools to process the selected corpus; there are useful tools that need minor adaptaion only (see, for example, CophiWordNet Search Interface 1.0 and Collaborative Philology Lab - Overview). CLARIN should organize the infrastructure so as to serve Ancient Greek scholars in a better way and help them to improve their studies. A concrete action plan should offer a new workbench and a data repository in which to insert text in a simple and intuitive way and visualize its encoding with specific TEI transcription; provide apparatus, literature and translation, link together primary sources and lexica, provide textual (and metrical) analysis and commentary, and offer search tools. To sum up, it is crucial to improve ancient studies by developing a common platform that responds to the desiderata of the scholars themselves. As we have argued, CLARIN will be able to support classical studies thanks to the application of concepts, methods and tools from the Digital Humanities to the field of classics and the study of the ancient world; thanks to CLARIN, we could test and store textual data for Ancient Greek scholars.

Finally, our survey may improve CLARIN in fostering new and sustaining existing knowledge in digital research in the humanities. Within the Humanities, informing research communities is fundamental and CLARIN will play a crucial role in disseminating results to the relevant academic, cultural, industrial communities and the interested public.

References

[Audenaert and Furuta2010] Neal Audenaert and Richard Furuta. 2010. What Humanists Want: How Scholars Use Source Materials. *JCDL '10: Proceedings of the 10th Annual Joint Conference on Digital Libraries*, 283–292. ACM, New York, NY. http://dx.doi.org/10.1145/1816123.1816166.

[Monachini and Frontini2016] Monica Monachini and Francesca Frontini. 2016. CLARIN, l'infrastruttura europea delle risorse linguistiche per le scienze umane e sociali e il suo network italiano CLARIN-IT. *JCoL - Italian Journal of Computational Linguistics*, vol. 2, 2, 11–30. Special Issue on NLP and Digital Humanities.

[Stefanini, Nicolosi and Monachini2017] Alberto Stefanini, Anika Nicolosi and Monica Monachini. 2017. Indagine e sperimentazione sulle pratiche d'uso di risorse e strumenti digitali nell'ambito della filologia greca. *Diss. - University of Parma*, 24, march, 2017.

[Toms and O'Brien2008] Elaine Toms and Heather L. O'Brien. 2008. *Journal of Documentation*, 64, 102–130. http://dx.doi.org/10.1108/00220410810844178.

[Warwick et al.2008] Claire Warwick, Isabel Galina, Melissa Terras, Paul Huntington and Nikoleta Pappa. 2008. The Master Builders: LAIRAH research on Good Practice in the Construction of Digital Humanities Projects. *Literary and Linguistic Computing*, 23, 383–396. http://discovery.ucl.ac.uk/13810/