

Digital Methods and Classical Studies

Neil Coffee <ncoffee_at_buffalo_dot_edu>, University at Buffalo
Neil W. Bernstein <bernsten_at_ohio_dot_edu>, Ohio University

Abstract

This essay introduces the articles in the special issue by locating them within current major lines of approach in digital classics.

Introduction

In the 1980s, a classical scholar might have been forgiven for considering digital approaches ancillary. The pioneering projects of the Perseus Digital Library and the Packard Humanities Institute might have seemed to simply provide existing texts in a more accessible form. By now, however, it has become clear that digital methods are transforming how we understand classical antiquity. At the same time that Google Books and the Hathi Trust and related efforts have made the content of millions of books instantly available on the internet to the public and scholars alike, electronic databases and journals like this one have opened new horizons by removing impediments of access. As with the study of the humanities as a whole, so within the field of Classics, new and enhanced tools continue to come online to facilitate existing modes of teaching and research and provide new ones. 1

These efforts are still new enough, however, that the field has not yet fully adapted to make them easily accessible. Classicists could once keep current by reviewing the latest journals and reviews, visiting conference bookstands, and leafing through publisher catalogues. But how would one expect to learn about, say, the LOFTS project for the digitization and analysis of fragmentary ancient texts? Peer review of digital resources exists in some areas of humanities, such as Modnets for American modernist literary and cultural studies, but not as yet in Classics.^[1] There are other good starting points, such as the list of digital projects provided by digitalclassicist.org and the announcements and conversations on the digitalclassicist listserv. Digitalclassicist.org has also helped organize a series of seminars in Berlin, Boston, Göttingen, and London. As elsewhere in the digital humanities, blogs and Twitter feeds serve as other important channels for sharing information. 2

Despite the existence of these fora, there has remained a need to bring digital classicists together to share techniques and results, and to connect them with a broader audience of scholars and educators. The Digital Classics Association (DCA) was created in 2013 in order to advance these aims, primarily in the United States, which, despite the ongoing work of Perseus and others, has nevertheless lagged behind Europe in conversations around digital classics. To help bring these to a wider classics audience, the DCA has organized a series of panels at the largest gathering of classicists in the world, the annual joint meetings of the Archaeological Institute of America and the Society for Classical Studies. As a contribution to the dialogue among those with digital interests and skills, it also staged a standalone conference at the State University of New York at Buffalo in spring 2013, “Word, Space, Time: Digital Perspectives on the Classical World,” with presentations on a variety of areas within the study of classical antiquity. 3

The conference was meant to be maximally inclusive, as indicated by its title: “Word” stood for texts generally, “Space” for material culture and place, and “Time” for historical studies. The ecumenical approach meant foregoing the benefit of drawing together experts in one sub-field prepared to fully critique one another’s work. It was designed rather to take advantage of the inherent interdisciplinarity of digital work and create a forum for conversations across the sub-fields of 4

classics and beyond. The conference provided a snapshot of important work along this spectrum, which, given the emerging and ongoing nature of many digital projects, often meant work-in-progress.

Digitizing Primary Texts

The contributions in this volume, drawn from the 2013 DCA conference, represent some of the major lines of approach current within the areas of digital classics.^[2] The most fundamental effort is to continue the work begun by Perseus and PHI in the 1980s of digitizing materials for reading and analysis. As part of its larger goals, the Open Philology project is working to add substantially to the Perseus corpus of Greek and Latin. Complementary efforts are beginning under the auspices of the Society for Classical Studies, the Medieval Academy of America, and the Renaissance Society of America, for the construction of a Digital Latin Library, which might be considered an expanded and richly augmented descendant of the long-useful Latin Library.^[3] More specialized efforts have made considerable progress digitizing texts from beyond the classical era, including DigilibLT, which hosts many late antique Latin texts, and Camena, which has a Neo-Latin collection. Among parallel efforts to digitize primary materials in the realm of material culture, a notable effort is FASTI, a site that aggregates the records from excavation and conservation of classical archaeological sites.

5

Two of the articles in this volume report on ongoing efforts to expand the fundamental materials available online. They also illustrate how these efforts often follow in the footsteps of Perseus in creating a rich set of tools with which to analyze the materials they have digitized. The first of these efforts is in fact an effort under the umbrella of the Open Philology project, the Leipzig Open Fragmentary Texts Series (LOFTS). In their contribution, Monica Berti, Bridget Almas, and Gregory Crane explain how LOFTS approaches digitizing several sets of classical texts consisting wholly or partly of fragments or quotations.

6

The first challenge lies in the selection of texts to digitize. Classicists have the advantage that the editions of their primary texts, and even some secondary texts, are often old enough to be free of copyright restrictions. Among the digital editions LOFTS is producing is one the *Fragmenta Historicorum Graecorum (Fragments of Greek Historians, or FHG)* of Karl Müller. The volumes of *FHG* were published from 1841 to 1870, and became the basis of the authoritative print edition of Felix Jacoby, *Die Fragmente der griechischen Historiker*, whose first volume appeared in 1923, and which has been continued now in online subscription format as Jacoby Online. The choice to digitize Müller's earlier *FHG* was partly by necessity, since some if not all of Jacoby is under copyright, and so not available for an open access edition. This sort of situation is faced by many digital classics efforts. Classicists may have full and free use of editions of most texts, just not the most recent editions. As part of the Open Philology project, the leaders of LOFTS have developed a best-practices strategy for confronting this problem, which includes signalling differences from more recent editions. Also, earlier editions, even when superseded in some respects, often have virtues that later editions do not. The *FHG*, for example, is more concise and manageable than Jacoby's more comprehensive work.

7

The second challenge is how to take advantage of the digital medium to enrich the texts. Addressing this challenge requires answering three questions. What additional information should be included together with the texts in the form of XML markup, including Linked Open Data identifiers to allow it to be identified and used by other resources? How should the marked-up texts be presented online? And what additional research tools should be joined with the texts?

8

The LOFTS team has approached all of these questions in an exemplary way. They have developed sophisticated methods for representing texts that are missing author and context, and even (parts of) the text itself, as when it is known only through paraphrase and reference. LOFTS also integrates its editions admirably with other resources. The project has used Canonical Text Services standards to identify texts and parts of texts, so that they can be easily linked to and analyzed.

9

The contribution of Caroline T. Schroeder and Amir Zeldes ("Raiders of the Lost Corpus") describes an analogous effort at digitization, but one that confronts its own distinct challenges. The goal of their Coptic Scriptorium project is to develop a digital corpus of Coptic texts. Coptic is the form of the ancient Egyptian language that emerged in the Roman period, with a largely Greek alphabet used for grammar and vocabulary evolved from ancient Egyptian. Critical as they are for the understanding of late Egyptian culture, as well as the contemporary cultures of Greece and Rome, many

10

Coptic texts have never been edited and published at all, much less in digital form. The project they describe thus takes as its starting point the creation of a collaborative publication and editing space for presenting these texts.

In both of these projects, the creation of digital editions is supplemented by tools with which to study them. LOFTS is the integrating its texts with the emerging Perseids editing platform for textual analysis and annotation. The Coptic Scriptorium includes morphological and syntactical tools for language analysis, as well as with taggers that allow entities within the text to be linked to other online resources, including those dealing with geographical and prosopographical information. The conjoining of texts and tools is characteristic of current efforts in digital classics, again following the example of Perseus.

11

Digitizing Derived Textual Information

Separate efforts have been underway focusing just on tools, apart from the publication of digital texts. One basic task is to derive entities from texts as independent objects of study, and for use in abstract models. One set of such entities are unique signifiers in the ancient world. Pleiades has created unique digital identifiers for all known *places* in antiquity. These include not just names of towns, cities, and geographical features, but also places that are manmade (Hadrian's wall), mythical (Atlantis), or contested (the birthplace of Homer). SNAP-DRGN is creating a uniform index with unique identifiers for ancient *people*. The Perseus Catalog is creating an index of all classical *texts*, using Canonical Text Services standards for the creation of unique identifiers to refer to texts, textual editions, and specific textual locations. By using Linked Open Data standards, these projects are establishing the infrastructure for highly sophisticated queries. In the future, for example, it should be relatively simple to discover how often Roman knights are mentioned in connection with Mediterranean ports, or what parts of foreign lands different Greek authors typically referred to at the opening of their works.

12

In his article "Extracting Citation Networks from Publications in Classics," Matteo Romanello demonstrates his method for automatically identifying and extracting canonical references to classical works in secondary scholarship. Through his efforts, for instance, one can now gather every reference to Vergil's *Aeneid* 3.452 in searchable scholarship, whether the title was given as "*Aeneid*" or abbreviated to "*Aen.*" or "*A.*". Romanello can then demonstrate, by means of network graphs, the density of reference to certain passages over others. His work will make it possible to precisely identify trends in scholarship, and reveal how scholarly conversations have affected how we see the ancient world.

13

The contribution of Francesco Mambrini, *Trebanking in the world of Thucydides. Linguistic annotation for the Hellespont Project*, takes us back to the primary sources themselves. Resources like the Classical Languages Toolkit now provide versions of basic and essential tools for linguistic analysis long available for English, such as tokenization, part-of-speech tagging, stemming, and even scansion. A need yet remains, however, for a resource of great interest to linguists as well as philologists, a set of fully syntactically annotated Greek and Latin sentences. Mambrini's paper reports on the Hellespont Project, an effort to perform syntactic analysis on the text of Thucydides, as a step toward creating a method for automatic syntactic parsing of ancient Greek generally. "Trebanking" is the common technical term for full morphological and syntactical analysis, and based on the treebank they created, Mambrini and his collaborators set out to answer compelling questions such as whether there is a general difference between the actions of the Athenians and the Spartans in the period between the Persian and Peloponnesian Wars. This is extremely challenging work, as it means trying to fit precise distinctions to the subtleties of syntax, but the article shows that real progress can be made.

14

Large-Scale Analysis

Just as "big data" has followed from digitization generally, so in digital humanities has the promulgation of digital resources and first-order tools been followed by attempts at large-scale analysis. This procedure has been given different names, such as, for literary studies, the "distant reading" Franco Moretti contrasts with traditional "close reading," [Moretti 2013] or the "macroanalysis" Matthew L. Jockers contrasts with "microanalysis" [Jockers 2013]. The promise of large-scale analysis is to gather together in great numbers elements of works that have been digitized and detect larger or otherwise hard-to-discern patterns in the form, contents, and locations of cultural artefacts.

15

In his contribution to this volume, Matthew L. Jockers himself brings this larger perspective on the digital humanities to bear in an interpretation of an aspect of the classical tradition. Here he focuses on the question of what the distribution of geographical place names in a corpus of 3,500 19th-century American, British, and Irish novels can tell us. After laying out his methods in ways that can serve as a guide to others pursuing similar approaches, he comes to a variety of telling observations. These include, for example, the fact that the Mediterranean is characterized as a more wild and adventuresome place, described with words coming under the themes of “Outlaws and Robbers,” “Female Heroines,” “Men with Guns.” This contrasts with the characterization of cities and civilizations situated on the Mediterranean such as Egypt, Greece, and Jerusalem, which are associated with more sublime themes like “Art and Beauty” and the “Holy and the Sacred.”

16

Education

Pedagogy and outreach have been major goals of work in digital classics from its beginnings. In the 1980s, Perseus and the Vergil Project were founding efforts to establish texts and study materials for students. More recently, the Dickinson Classical Commentaries represent a leap forward in classical language pedagogy that takes full advantage of students’ increased access to the internet through phones and tablets. The site successfully integrates multimedia offerings with traditional pedagogical aids such as customizable vocabulary lists and live readings by the instructor. The series includes authors commonly read in the introductory Latin language curriculum (Caesar, Cicero, and Ovid) as well as those not commonly encountered by undergraduates such as Callimachus and Sulpicius Severus.

17

More generally, making basic materials and analytical tools more easily available to all has provided additional benefits for students. In addition, teachers have created or adapted new digital pedagogical materials and techniques for the use of students. Two of the examples presented in this volume include Jeffrey Rydberg-Cox’s Greek pedagogy site and Rebecca Schindler’s initiative on spatial literacy.

18

Jeffrey Rydberg-Cox’s paper reports on his initiative to make self-guided pedagogical materials freely available for students of ancient Greek. He adapted John Williams White’s *First Greek Book* (1896), a textbook with 80 lessons based upon the language and vocabulary of Xenophon’s *Anabasis*, for a variety of mobile platforms. Sentence alignment exercises, in which users match individual Greek words with English words and phrases, help students make the transition from drilling exercises into reading authentic Greek. Spaced repetition enables students to practice material over multiple sessions. The initiative has been successful both for Rydberg-Cox’s students at University of Missouri-Kansas City and other students of ancient Greek. The site boasted 7,800 unique users in its first twelve months, a significant number for a less-commonly taught language.

19

Rebecca Schindler’s paper discusses the integration of spatial literacy skills in the classical studies curriculum. Spatial literacy is now valued alongside the more traditional skills of writing, numeracy, and critical thinking. Classical studies has tended to fall behind other disciplines in terms of teaching spatial literacy, as courses in GIS for classicists are relatively rarer. Schindler describes how she and her colleagues have integrated spatial literacy into classical archaeology courses, using examples from the Collaboratory for GIS and Mediterranean Archaeology (CGMA) at DePauw University.

20

Conclusions

Three years after the original DCA meeting, the future of digital humanities in classical studies continues to look promising. Both philologists and archaeologists will benefit from a series of new, publicly available projects. One volume cannot capture the diversity of projects discussed at the DCA conference, much less the broader field of classical digital humanities. To enlarge the scope further somewhat, however, we conclude with a survey of two projects whose leaders presented at the conference but do not have articles in this volume, to give a sense of other important directions in the field.

21

Network analysis has been a major line of approach among digital humanists. Research in social network analysis considers how associations of individuals and groups map onto geographical space. Representative of such efforts is the work of Maxim Romanov on toponyms used in the early Islamic world.^[4] By gathering locating names such as “Al-

22

Baghdadi” and “al-Basri” from the 14th-century *History of Islam* written by the Damascene historian al-Dhahabī on maps of the geography where Islam was born and grew, Romanov shows how the frequency of mentions of regions changes. From this information, he is able to sketch a kind of social geography of the early Islamic world as one indicator of the shifting cultural and religious importance of different regional centers over the centuries.

Traditional philological studies have also been transformed by the new digital tools. Neil W. Bernstein described a method for gauging the internal connectedness, through allusion and intertext, of the works that make up the corpus of classical Latin epic poetry. Bernstein and his collaborators took the individual, often subtle borrowings of language from one epic author to another revealed by the search provided by the Tesseract Project web tool. They conducted a rigorous, large-scale study that compares the overall level of intertextuality among authors writing epic.^[5] The result was a relatively objective measure of which authors were most influential in the tradition, and which on the other hand were most interested in alluding to and borrowing from the language of their predecessors and contemporaries. The result is a kind of genealogy of influence that was impossible to conduct before the advent of digital tools.

23

The diversity of projects introduced at the DCA meeting reflected the interdisciplinarity of classical studies, which has traditionally integrated the study of literary and archaeological sources. So whereas the notion of “big tent” digital humanities, including not just developers but users and others, has had to be argued for in the broader field of digital humanities [Gold 2012], participants in the 2013 DCA conference demonstrated a strong sense of common purpose in developing technologies and adapting them for particular ends. The continuing development of large-scale digital infrastructure in classics described above should provide an improved common basis for such efforts, allowing for increased collaboration around existing tools and data archives, and so easier creation and dissemination of new analytical and pedagogical applications.

24

Notes

[1] The most-read journal in classics, the *Bryn Mawr Classical Review* book reviews, once had a companion review of digital resources, but no longer.

[2] An account of the origins of digital classics is provided by [Crane 2004], of digital humanities, by [Hockey 2000].

[3] Like Latin Library, the Lacus Curtius website is helpful for offering clean, simple texts, though Lacus Curtius has Greek in addition to Latin, as well as English translations for many of its texts.

[4] Romanov, M. forthcoming “Toward Abstract Models for Islamic History.” In *The Digital Humanities + Islamic & Middle Eastern Studies* (pdf of submission).

[5] Published separately in *DHQ*: Bernstein, N. W., K. Gervais, and W. Lin. 2016. “Comparative rates of text reuse in classical Latin hexameter poetry.” *DHQ* 9.3.

Works Cited

Crane 2004 Crane, G. 2004. “Classics and the Computer: An End of the History.” In *A Companion to Digital Humanities*. Eds. S. Schreibman, J. Unsworth and R. G. Siemens. Oxford, Blackwell: 46-55.

Gold 2012 Gold, Matthew, ed. 2012. *Debates in the Digital Humanities*. Minneapolis: University of Minnesota Press.

Hockey 2000 Hockey, S. M. 2000. *Electronic Texts in the Humanities: Principles and Practice*. Oxford, Oxford University Press.

Jockers 2013 Jockers, M. L. 2013. *Macroanalysis: Digital Methods and Literary History*. Champagne, IL, University of Illinois.

Moretti 2013 Moretti, F. 2013. *Distant Reading*. London: Verso.



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.