Introduction

This book is the publication of the Habilitation monograph that I wrote to describe the results of new experimental research in Digital Classics carried out at the University of Leipzig as part of the *Open Philology* project of the Alexander von Humboldt Chair of Digital Humanities.

The aim of this monograph is to present a new model for producing digital editions of historical fragmentary texts, by which I mean texts that are now lost in their original form and transmitted only through quotations and reuses in later works. Being a Classicist, in this research I analyze ancient Greek and Latin sources with a main focus on historiography. The choice of this topic is due to three main circumstances: 1) an interest in Greek fragmentary historiography that led me to explore this genre and publish papers and critical editions of fragmentary authors,[1] 2) a related interest in the transmission of Classical texts and in the philological contribution of the Alexandrian Library,[2] and 3) the impact of the digital revolution on Greek and Latin that brought me to work with the *Perseus Project* at Tufts University and with the *Open Greek and Latin* initiative at the University of Leipzig.[3]

The meaning of the term *edition* in the title of this book has to be explained, given that this word is the focus of many scholarly debates in recent publications concerning the Digital Humanities.[4] In my work, the expression *digital edition* refers not to the publication of a new reconstructed text of Greek historical fragments that is the result of autopsies and new readings of manuscripts and other historical media, but to the *critical* selection, production, analysis, interpretation and annotation of digital data about fragmentary authors and works. Scholars

---

1 This work began as part of a collaboration with the Italian series *I Frammenti degli Storici Greci* at the University of Roma Tor Vergata and as part of my contribution as a copy editor and author of the Brill's *Jacoby Online* project. Results of these activities are available in the following publications: Berti (2009a); Berti (2009b); Berti (2010); Berti (2012); Berti (2013a); Berti (2013b); Berti/Jackson (2015); Martin/Berti (2017).

2 Berti/Costa (2010); Berti/Costa (2013); Berti (2014a); Berti (2014b); Berti (2015a).

3 *Open Greek and Latin* (OGL) is part of the *Open Philology* project developed by the Alexander von Humboldt Chair of Digital Humanities at the University of Leipzig under the direction of Gregory R. Crane: Baumgardt/Berti et al. (2014); Berti (2019a). Beside this monograph, results of my research on digital editions of historical fragmentary texts are available in the following publications: Berti/Romanello et al. (2009); Romanello/Boschetti et al. (2009); Almas/Berti (2013); Berti/Almas (2013); Büchler/Geßner et al. (2013); Berti/Almas et al. (2014–2015); Berti (2015b); Yousef/Berti (2015); Berti/Almas et al. (2016); Berti/Bizzoni et al. (2016); Berti/Blackwell et al. (2016); Berti/Daniels et al. (2016); Berti (2017a); Berti (2018); Berti (2019b); Berti (2019c).

4 Many bibliographic resources on these debates have been produced and are easily accessible by searching the *Internet*. The most recent and significant contributions are Sahle (2013), Apollon/Bélisle et al. (2014), Pierazzo (2015), and Sahle (2016).

with enough experience in digital data know that the computational environment is bringing new questions and solutions to the treatment and preservation of historical texts. Centuries of philological practice have been contributing in an impressive manner to the advancement of our knowledge of the Greco-Roman world thanks to the technology of the printed book. If the goal of the digital turn is to preserve these results and continue these exegetical efforts, it is indisputable that the digital medium is very different from the printed medium and that it requires a different approach.

This monograph describes practical and concrete problems that scholars have to deal with when trying to digitally represent and analyze textual fragments of lost authors and works. Publishing *printed* historical fragments is a very difficult and complex task. Publishing *digital* historical fragments is an even more difficult and complex task, because, if we always need to answer critical research questions, we also have to navigate in a new dimension where we can't rely on the work of our predecessors and we can't make use of conventions and standards established by philology in the printed age.

This is the reason why results described in this book are experimental and address new issues that still have to be properly discussed and solved if we want to take full advantage of digital technologies and fulfill our responsibility to preserve the Greek and Latin textual heritage. When I started my work on this topic, the technology was different and many digital resources for analyzing Greek and Latin were missing. A lot of time was spent on producing data and rethinking our approach to historical sources.[5] My future work and the work of future generations will certainly help go beyond naiveties and mistakes of this first phase of research in Digital Classics.

Experiments and results of my research are also profoundly influenced by and indebted to very different scholary environments I have been working for. My Italian academic background taught me the principles of historiographical research. My activities at the *Perseus Project* represented a fundamental moment to move from analog to digital philology and concretely experiment with the production of digital textual fragments. My research and teaching duties in the Institute of Computer Science at the University of Leipzig have given me the possibility to deepen the computational aspect and produce new resources.

The outcome is a very interdisciplinary product that also reflects what is happening not only in Classics, but also in Computational Linguistics and Digital Humanities, where we can see the birth and growth of many projects that apply new technologies to the study of historical languages and of the ancient world.[6]

---

5   Cf. Jannidis/Kohle et al. (2017) for a decription of the work of a humanist today, who needs to combine expertise both in Computer Science and in his/her relevant humanities field.

6   For a description of the current state of the art of Digital Classical Philology, see Berti (2019a) and Chronopoulos/Maier et al. (2020).

This monograph is structured in five chapters that introduce the new domain of digital fragmentary literature and describe two related projects that I have been working on in the last five years: the *Digital Fragmenta Historicorum Graecorum* and the *Digital Athenaeus.*

Chapter 1 (*Fragmentary Texts and Print Culture*) explains the meaning of the word *fragment* in Classical studies and individuates the most important scholarly phases that have been producing modern collections of fragmentary texts. Statistics are offered to quantify the amount of fragmentary authors and works on the basis of available digital data, and a description of the characteristics of printed editions of historical fragmentary texts is provided to show the role of the technology of the printed book in shaping the field of fragmentary historiography in the last two centuries.

Chapter 2 (*Fragmentary Texts and the Digital Revolution*) describes the first generation of digital libraries where fragmentary texts are collected and published in a way that still depends on the printed editorial practice. The chapter analyzes how hypertextual theories are important for a new model of fragmentary texts in a digital environment and describes the concept of cover-text, that helps philologists move the attention from the isolated fragment to its context of transmission. The chapter describes also experimental implementations of computational techniques that in the future will be hopefully applied to the domain of fragmentary literature, such as text reuse detection, intertextual analysis, and translation alignment.

Chapter 3 (*Distributed Annotations of Fragmentary Texts*) explains the new idea of conceiving fragmentary texts as annotations of textual elements about lost authors and works. It also describes two resources for producing canonical citations and annotations of historical fragments: the *CITE Architecture* and the *Perseids Fragmentary Texts Editor.*

Chapter 4 (*Digital Fragmenta Historicorum Graecorum*) is the description of the *Digital Fragmenta Historicorum Graecorum* (DFHG), which is the dynamic and expanded version of the printed edition of the *Fragmenta Historicorum Graecorum* (FHG) edited by Karl Müller.

Chapter 5 (*Digital Athenaeus*) is the description of the *Digital Athenaeus*, which is a project that provides an inventory of authors and works cited in the *Deipnosophists* of Athenaeus of Naucratis and implements a data model for identifying, analyzing, and citing uniquely instances of text reuse, in order to produce a text-based and annotated catalog of Greek fragmentary authors and works.

The *Digital Fragmenta Historicorum Graecorum* is not the mere digitization of a printed collection, but the extraction, expansion, and deep analysis of its data to produce further resources for the study of Greek fragmentary historians and their works. The reasons for choosing this collection are fully explained in the

following pages. Nevertheless and given the topic of this monograph, a reader could ask why I didn't produce a new born-digital edition of Greek fragmentary historians. I have two main answers to this question.

The first answer is that many extant sources that preserve Greek historical fragments are still missing in an open and digital format. A comparison with the first volume of the indices of *Die Fragmente der griechischen Historiker* by Pierre Bonnechère shows how many sources of fragments and editions are not yet digitized.[7] This situation will be certainly overcome in the next decades, but today is a significant limit considering that a complete collection of source texts preserving testimonies and fragments is the preliminary condition to produce an edition of a fragmentary author.

The second answer is that the time at my disposal for this project would have allowed me to work only on one fragmentary author or on a small group of authors with a limited number of fragments. The adverb *only* in this sentence may seem provocative, because the study of a single fragmentary author can be extremely difficult and time consuming, as I know by experience. But, if we change our perspective from a traditional philological point of view to a computational point of view, this adverb makes better sense. Focusing on one author or on a small group of authors would have resulted in a work very similar to what I edited in a printed format, without a proper exploration of new computational possibilities. The goal of this research was not to replicate what we can still achieve with the technology of the printed book, but to experiment with new forms of analysis in a digital environment, where a single fragmentary author or a few hundreds fragments are not enough as an object of study.[8]

This book shows that digital libraries are transforming the concept of textual fragments. If in printed books textual fragments are chunks of texts extracted from their original context, in digital editions they are annotations of textual elements pertaining to lost authors and works. The *Digital Athenaeus* project was born from the need to produce this model and move from the perspective of the quoted author to that of the quoting author, which will be the prevailing perspective in the future development of digital fragmentary literature. The choice of the *Deipnosophists* is not only due to historical and philological reasons that are extensively described in the following pages, but also to the fact that its main editions are openly available in a digital format and that data can be extracted and structured from them in a sustainable way.

This monograph is written in English for two main reasons. The first reason is that English is not only the language of the Digital Humanities, but also

---

7   Bonnechère (1999).
8   For important recent contributions on digital data and literary studies, see Piper (2018), Eve (2019), Lemercier/Zalc (2019), and Underwood (2019).

the natural language that in the last decades has produced models and programming languages to analyze textual data. If our responsibility as humanists and philologists is to translate and express these models into other languages, we can't renounce to describe the results of our research in English, given that this language permeates methods and standards of Digital Philology.

The second reason is that I was employed as an Academic Assistant of the Alexander von Humboldt Chair of Digital Humanities at the University of Leipzig to develop an English program for teaching and research. I achieved this task by contributing to the creation of a new Bachelor of Science and a new Master of Science in Digital Humanities, where I currently teach courses in Digital Philology and Digital Classics.[9] I also fulfilled this task by promoting research projects, conferences, and publications that include data described in this book.[10]

---

9   Part of this task was also the creation of *Sunoikisis Digital Classics*, which is an international consortium of Digital Classics programs developed in collaboration with the Harvard's Center for Hellenic Studies and the Institute of Classical Studies at the University of London: Berti/Crane et al. (2015); Berti (2016b); Berti (2017c); Berti (2017b).

10   Bibliographic resources are collected in the *Bibliography* at the end of the volume with DOIs and dates of access. The volume cites many links that are not provided with stable identifiers, but are important to show the current state of research. The last access to these links was on July 23, 2021. The *Index* at the end of the volume lists Digital Humanities and Digital Classics projects not explicitly mentioned in chapter and section titles.