

2 Fragmentary Texts and the Digital Revolution

This chapter describes how fragmentary authors and works are represented in the first generation of digital libraries. It also analyzes how hypertextual models have been developing a new dimension, where textual fragments are envisioned as *text reuses* preserved in *contexts* that cover and therefore hide their original form. The first section (2.1) presents the relationship between digital scholarship and textual fragments by describing how fragmentary texts are currently collected and published in digital libraries that still depend on the printed editorial practice. The section introduces digital projects that include Greek and Latin fragmentary texts, devotes specific attention to the *Thesaurus Linguae Graecae* (2.1.1) and to the *Jacoby Online* (2.1.2) with the example of the edition of the fragments of Isthos the Callimachean (2.1.3), and presents digital collections of physical fragments such as inscriptions, papyri, and manuscripts that bear textual evidence (2.1.4). The second section (2.2) describes how editions of fragmentary authors and works are representations of hypertexts and how hypertextual theories in literature and digital media are important for a new publishing model of fragmentary texts in a digital environment. The last section (2.3) presents the concept of *cover-text* that has led philologists to move the attention from decontextualized fragments to the role of the context that preserves quotations and reuses of lost texts. Parallel to this is the implementation of computational techniques for *text reuse* detection that are now also applied to historical data (2.3.1). Given that they are strictly connected to text reuse, the two last sections focus the attention on intertextual analysis (2.3.2) and translation alignment (2.3.3), whose methods are beginning to be experimented with literary texts and historical documents.

2.1 Digital Scholarship and Textual Fragments

The digital revolution has been affecting primary and secondary textual sources of Greek and Latin works.¹ The first generation of digital libraries has converted into a machine readable format the reconstructed text of single editions of Classical works. The goal of the second generation of digital libraries is to publish multiple editions of the same work, reproduce the critical apparatus and all other paratextual elements (prefaces, introductions, indices, bibliographies, notes, etc.), and generate collaborative environments for new born-digital critical editions of Greek and Latin sources.² Fragmentary authors and works are directly involved in this process because they consist of quotations and text reuses preserved by still surviving sources. The problem is that the model according to which fragmentary texts are currently represented in digital libraries is not satisfactory, because it strongly depends on printed editorial practices. In the following pages, I describe varieties of this model according to the most important collections of Greek and Latin sources.

As far as ancient Greek sources are concerned, the *Thesaurus Linguae Graecae* (TLG) is the most extended digital library that also includes many editions of fragmentary authors and works. Given the huge amount of TLG texts and the importance of the TLG *Canon*, I refer to section 2.1.1 for a more specific and detailed analysis of the TLG treatment of fragmentary texts in its databank. The *Perseus Digital Library* and the new *Scaife Viewer* don't contain editions of fragmentary authors and works, but the *Open Greek and Latin* (OGL) project has been offering a first set of digitized versions of printed editions of fragmentary authors and works. In this case, the goal is to generate OCR outputs with a basic TEI XML encoding of printed editions in order to produce machine readable files that allow scholars to create digital versions of printed books and extract data for many different purposes. An example is constituted by the *Digital Fragmenta Historicorum Graecorum* (DFHG) that will be described in chapter 4. A project specifically focused on producing digital editions of Greek fragmentary authors is Brill's *Jacoby Online*, which is continuing in a digital format the work started by Felix Jacoby one century ago for collecting evidence of lost Greek historians. Considering its

1 Berti (2019a).

2 Babeu (2011) part. 2–3 on “several generations of digital corpora in Classics”; Apollon/Bélisle et al. (2014); Pierazzo (2015); Boschetti (2018) 11–12. On the concept of *paratext*, see Genette (1982) 9 and *passim*; Berti (2012) 444.

specificity, the project will be discussed in section 2.1.2.³ As for Latin literature, PHI *Latin Texts* is one of the electronic databases of Greek and Latin sources produced by the *Packard Humanities Institute*.⁴ PHI *Latin Texts* is a digital archive of 836 Latin works from 362 authors up to 200 CE with a selection of sources from later antiquity. The *Canon* of the PHI *Latin Texts* offers an overview of authors and works that are part of the collection with references to the printed editions on which digital texts are based: <http://latin.packhum.org/canon>. Authors and works are identified with numbers and abbreviations. For example, Marcus Tullius Cicero is [0474 Cīc] and Cicero's *In Catilinam* (ed. A.C. Clark 1905) is [0474 013 Cīc].⁵

PHI *Latin Texts* includes collections of fragmentary authors, such as the *Historicorum Romanorum Reliquiae* (HRR) by Hermann Peter.⁶ An example are the fragments of the Latin *Annales* of Fabius Pictor [0061 001 hīst].⁷ In this case, PHI reproduces the text of only three of the six Latin fragments of Fabius Pictor originally edited by Peter under the section *Fabii Pictoris Latini Annales* (3, 4, and

3 On these projects see also section 1.3. There are other still in progress plans for producing digital editions of single Greek fragmentary authors or groups of authors belonging to specific genres, as for example the fragments of Protagoras of Abdera by Tazuko van Berkel at Leiden University, the fragments of Demetrius of Scepsis by Alexandra Trachsel at the University of Hamburg, the fragments of Sceptic philosophers by Stéphane Marchand at the École Normale Supérieure de Lyon, and the fragments of ancient Greek rhetoric and oratory by Jan Heßler at the University of Würzburg. Another project is the collection and edition of fragments and testimonies of historians from late antiquity at the University of Düsseldorf, where they are producing printed editions with an online version. The project is still in beta version and has been conceived as a traditional printed critical edition with a parallel online presence: see Fischer (2017) 267–268.

4 These databases were originally published as CD-ROMs. PHI 5.3 was a collection of Latin texts now freely accessible at <http://latin.packhum.org>: see Kozák (2018). PHI 7 was a database of Greek inscriptions and documentary papyri. Greek inscriptions are online at <https://inscriptions.packhum.org>, while documentary papyri are part of *Papyri.info*. On PHI *Greek Inscriptions*, see p. 69.

5 The *Canon* of the PHI is now ingested in a new project for publishing and curating critical editions of Latin texts, which is under development as the *Digital Latin Library* (DLL) and which is also going to include fragmentary authors and works: <https://digitallatin.org>. The HTML code of the PHI *Canon* embeds tags and attributes that can be used for extracting data and reusing it for other purposes. The DLL has extracted this data and integrated it in its catalog, which is a *Linked Open Data* (LOD) resource built according to the best practices of library information science: <https://catalog.digitallatin.org>. For example, Marcus Tullius Cicero is cataloged as DLL #A5129 and its entry is available at <https://catalog.digitallatin.org/dll-author/a5129>. For a presentation of the DLL project, see Huskey (2019) and *Digital Latin Library Project* (2021).

6 Peter (1870–1914).

7 See <https://latin.packhum.org/author/61>. This data is ingested in the DLL Catalog with a DLL identifier (DLL #W2649): <https://catalog.digitallatin.org/dll-work/w2649>. Fragments of Fabius Pictor have been encoded in TEI XML as part of the *Digital Fabius Pictor* project developed at the University of Leipzig as a result of a collaboration between the Institute of Computer Science and the Historical Seminar: see Straßburger (2018).

Authors	Word Search	Concordance	About	PHI Latin Texts
Fabius Pictor, Annales 3.1				Betacode
LATINI ANNALES EX LIBRO I				
NON. 518M				
Et simul uidebant picum Martium.			3.1	
SERV. DAN. A. 8.630				
spelunca Martis			4.1	
EX LIBRO IV				
GEL. N.A. 5.4.3				
Quapropter tum primum ex plebe alter consul factus est, duouice-			6.1	
simo anno postquam Romam Galli ceperunt.				

Figure 2.1. PHI: Fabius Pictor, fragments of the *Latini Annales*

6).⁸ Moreover, PHI doesn't publish the critical apparatus, the commentary, and the context of the witnesses as Peter does, but reproduces only the words that can be attributed to the original lost texts of Fabius Pictor (fig. 2.1).

An identical treatment of the Latin fragments of the *Annales* of Fabius Pictor is available in the *Bibliotheca Teubneriana Latina* (BTL) and in the *Library of Latin Texts Series A* (LLT-A). The BTL online provides electronic access to all printed editions of the *Bibliotheca Teubneriana Latina* with a database that includes about 13 million word forms.⁹ The BTL publishes the same portions of fragments 3, 4, and 6 as in PHI *Latin Texts* (fig. 2.2). It allows to export a PDF file with the text of the fragments and also different citation formats with a permanent URL.¹⁰ The *Library of Latin Texts Series A* (LLT-A) is part of a cluster of full-text Latin databases and dictionaries which contains over 78 million Latin words from more than 3,800 works attributed to ca. 1,200 authors.¹¹ LLT-A provides the same portions of fragments 3, 4, and 6 of Fabius Pictor as in PHI *Latin Texts* and in the BTL, and allows users to export a PDF file with information about the author, the texts, and the reference edition (figg. 2.3, 2.4, and 2.5).¹² PHI *Latin Texts* includes under the name of Fabius Pictor also the fragments of the *Iuris Pontificis Libri* ([0061 002 iur]) according to the edition *Iurisprudentiae Anteiustinianae Reliquiae* by P.E. Huschke, E. Seckel, and B. Kübler (vol. I, 1908). LLT-A collects

8 Peter (1870–1914) I 112–113 (Greek fragments are at 5–39). PHI seems to publish only the fragments that Peter attributed to Quintus Fabius Pictor and not those that he questioned as being of Numerius Fabius Pictor. On these problems of attribution, see also FHG III 80–93, FGrHist (BNJ) 809, Cornell (2013) I 163–166, and Woodman (2015) 4–22.

9 The BTL is accessible only through a library subscription at <https://doi.org/10.1515/btl>.

10 See <https://www.degruyter.com/document/database/BTL/entry/AFAPIANNA/html>. For the fragments of the *Iuris Pontificis Libri*, see below.

11 The project started in 1991 and the collection is now accessible with an institutional subscription on the BREPOLIS website. Since 2009, LLT-A is supplemented by LLT Series B.

12 See http://clt.brepolis.net/LLTA/pages/TextSearch.aspx?key=AFAPIANNA_.

some of these fragments (*Iuris Pontificis fragmenta*) under *Fabius Pictor quidam* (an *potius Quintus Fabius Maximus Seruilianus*), but according to the edition of Peter (1870–1914) I 114–116.

Fabius Pictor quidam

Annalium fragmenta (in aliis scriptis seruata)

De Gruyter | 2011

Entry Type	publication
Entry Language	Latin
Author	Fabius Pictor quidam [fragmenta in aliis scriptis seruata]
Work	Annalium fragmenta (in aliis scriptis seruata)
Aetas	Antiquitas
Genre	prosa
TLL Code	FAB. PICT. hist.
LLA	LLA 157

Saeculum

Century	s. 2 a.c. (dubium)
Chronology	8

Memento

Fabius Pictor quidam [fragmenta in aliis scriptis seruata]

s. II a. Chr. ?

Annalium fragmenta (in aliis scriptis seruata) - s. 2 a.c. (dubium) - prosa

LLA 157 - TLL FAB. PICT. hist.

Teubner (H. Peter, 2da ed. 1967) | *Historicorum Romanorum reliquiae*, vol. II, p. 112–113

Summa formarum: 22

Summa formarum dissimilium: 22

Hinweise

Die Fragmente 3, 4 und 6 wurden aufgenommen.

Notes

Fragments 3,4 and 6 are included.

Table of Contents ▼

fragm. 3, p. 112 l. 7

et simul uidebant picum Martium.

fragm. 4, p. 113 l. 3

spelunca Martis

fragm. 6, p. 113 l. 11

Quapropter tum primum ex plebe alter consul factus est, duouicesimo anno postquam Romam Galli ceperunt.

Source

Title	Bibliotheca Teubneriana Latina (BTL) Online
Edited by	De Gruyter
Publisher	De Gruyter 2009

Figure 2.2. BTL Online: Fabius Pictor, fragments of the *Latini Annales*

Musisque Deoque is a digital archive of Latin poetry from its origins to the Italian Renaissance supported by a critical and exegetical electronic apparatus. The collection includes also fragmentary works, as for example the fragments of Quintus Ennius. Figure 2.6 shows lines 105–109 (book 1) of the *Annales*. The text is based on the edition of Otto Skutsch (*The Annals of Quintus Ennius*, 1985) and has been manually digitized by Paolo Mastandrea and Silvia Arrigoni.¹³ The page provides the critical apparatus, references to the source texts (with concordances to the editions of Johannes Vahlen and Enrico Flores), complete bibliographical and

13 In this case the permalink is <http://www.mqdq.it/texts/ENN|anna|105>. Texts and critical notes of the edition of Otto Skutsch have been digitized and are now available online on the website of the *Oxford Scholarly Editions Online* (OSEO) service (subscription required) (fig. 2.8).

The screenshot shows the 'Library of Latin Texts - Series A' interface. At the top, there are navigation links: 'Search Screen', 'Table of Contents', and 'Distribution of Word-forms'. A 'Last Update: 2014-12-17' is displayed on the right. Below the navigation is a 'Selection' section with a search bar containing 'E' and a list of search results. The first result is 'Fabius Pictor quidam [fragmenta in aliis scriptis seruata]' and the second is 'Annalium fragmenta (in aliis scriptis seruata) - Fabius Pictor quidam [fragmenta in aliis scriptis seruata] - s. 2 a.C. (dubium)'. Below the search results are two columns: 'Contents' and 'Contexts'. The 'Contents' column lists three fragments with their respective page and line numbers. The 'Contexts' column provides detailed information for each fragment, including the text itself and the source edition.

Figure 2.3. LLT–A: Fabius Pictor, fragments of the *Latini Annales*

This screenshot displays detailed information for the search results. It includes the title 'Fabius Pictor quidam [fragmenta in aliis scriptis seruata]', the author 's. II a. Chr.?', the work title 'Annalium fragmenta (in aliis scriptis seruata) - s. 2 a.C. (dubium)', and the edition 'LLA 157 - TLL FAB. PICT. hist. Teubner (H. Peter, 2da ed. 1967) [Historicorum Romanorum reliquiae, vol. I], p. 112-113'. It also provides statistics: 'Summa formarum: 22', 'Summa formarum dissimilium: 22', and 'Media uerborum longitudo: 5,82'. A 'Permalink' is provided: 'http://clt.brepolis.net/LLTA/pages/TextSearch.aspx?key=AFAPIANNA_'. Below this, there are two paragraphs of text in French and English, providing background information and acknowledgments to the University of Liège and the Teubner edition.

Figure 2.4. LLT–A: Fabius Pictor, fragments of the *Latini Annales* (background on the text)

The screenshot shows a PDF export titled 'Export from the Library of Latin Texts - Series A'. It includes the export date '2018-05-19 09:01 (CET)', the publisher '© Brepols Publishers, Turnhout, 2014', and the website 'http://www.brepolis.net'. The document is divided into three sections: 'Results', 'Background on the text', and 'Background on the text'. The 'Results' section contains the same search results as Figure 2.3. The 'Background on the text' section contains the same background information and acknowledgments as Figure 2.4.

Figure 2.5. LLT–A: Fabius Pictor, fragments of the *Latini Annales* (PDF export)

textual-critical information, and the metrical scansion of the Latin text through *Pede certo*.

The screenshot shows a web interface for the MQDQ project. At the top, there are navigation links: 'About', 'Witnesses', 'Metres', 'Metrical scansion', and 'Close'. The main title is 'Ennius *annalium fragmenta* 105-109'. Below this, there is a reference basis text: 'Reference basis text: D. Skutsch (1985)', 'Editing of the digital edition: P. Mastandrea, S. Arrigoni (2014)', 'Data insertion and check: S. Arrigoni', and a permalink: 'http://www.mqdq.it/texts/ENN[anna]105'. There are also 'Other sections' and a search bar. The main content area displays the text of fragment 105: '105 Pectora ** tenet desiderium; simul inter Sese sic memorat: "O Romule, Romule die, Qualem te patriae custodem di generum! O pater, o genitor, o sanguem dis oriundum! Tu produxisti nos intra luminis oras"'. To the right, there is a metrical scansion: '110-114 V² = 114-118 Fl.' and a commentary: 'Cic. rep. I, 64. Iusto quidem rege cum est populus orbatus, "pectora" diu (diu corrector) "tenet desiderium", sicut ait Ennius post optimi regis obitum, "simul... oriundum", non eros nec dominos appellabant eos quibus iuste parerant, denique ne reges quidem, sed patriae custodes, sed patres, sed deos; nec sine causa: quid enim adiungunt? "tu produxisti nos intra luminis oras", uitam honorem decus sibi datum esse iustitia regis existimabant. Lact. inst. I, 15, 30: nam Romulum desiderio suis fuisse declarat Ennius, apud quem populus antium regem dolens haec loquitur: "O Romule... generum: tu... oras, o pater o patriae o sanguem dis oriundum". Prisc. gramm. II p. 250, 12 vs.: ueteres hoc sanguem dixerunt [...] ut ait Ennius (scen. 26) [...] idem in II § liber "peruētustus" Columae anni "o genitor... oriundum".'

Figure 2.6. *Musisque Deoque*: Quintus Ennius, *Annales*, II. 105–109

As far as Latin fragments are concerned, I also point to *Grammatici disiecti: sources fragmentaires pour l'histoire de la grammaire latine* (<https://gradis.hyptheses.org/>). The project is directed by Alessandro Garcea and is “a research blog dedicated to gathering, for the first time, all Latin grammatical texts which are preserved exclusively in fragmentary form.” The “primary purpose is to publish bio-bibliographical sketches of the authors of such texts, be they grammarians, teachers, erudite writers or any other author who may have written works on grammar, regardless of their position in society or their linguistic activity.”¹⁴ According to the description, *Grammatici disiecti* provides through a WordPress website a list of Latin fragmentary grammarians dated between the 3rd century BC and the 4th century CE. As soon as they are ready, the project publishes separate pages with complete bio-bibliographical presentations of the authors of the collection.

Beyond the projects mentioned in these pages, there are also other resources for accessing in many different ways digital versions of printed editions of fragmentary authors and works. For example, publishing companies have been offering this kind of service through online subscriptions for individual and institutional customers. The *Digital Loeb Classical Library* offers “an interconnected, fully searchable, perpetually growing, virtual library of all that is important in Greek and Latin literature.”¹⁵ The online collection counts more than 520 vol-

14 A new edition of the corpus is in preparation for *Les Belles Lettres* editions (*Collection des Universités de France*). As part of an interest in Latin grammarians, the project is connected to the *Corpus Grammaticorum Latinorum* (CGL): see Garcea/Cinato et al. (2010).

15 See <https://www.loebclassics.com>. As of 2021, the annual subscription for individual users is priced at \$170 for the first year and \$70 for subsequent consecutive years. Prices for academic and public libraries depend on the size of the institutions interested in the subscription.

umes of Latin, Greek, and English texts, which are available in an interface that allows readers to browse, search, bookmark, annotate, and share content. The Loeb online has also editions of fragmentary works, as for example the Greek epic fragments published by Martin West in 2003 (fig. 2.7).¹⁶ Other examples that are accessible with a subscription or direct payment are the *Oxford Scholarly Editions Online* (OSEO), that enables readers to search across the texts, navigate through reference forms and look up words in the *Oxford Latin Dictionary* (fig. 2.8),¹⁷ or the online *Bibliotheca Scriptorum Graecorum et Romanorum Teubneriana* that has put online PDF files of its editions including fragmentary authors and works.¹⁸



Figure 2.7. Digital Loeb Classical Library: M.L. West, *Greek Epic Fragments*

A different kind of service is offered by mass digitization projects like *Google Books*, *Internet Archive*, and *HathiTrust*, which have been giving free access to millions of scanned copies of books that are in the public domain. These collections contain also many volumes about Classical sources, allowing scholars to *rediscover* past editions of ancient authors that comprise fragmentary ones.¹⁹

16 West (2003).

17 <https://www.oxfordscholarlyeditions.com>

18 <https://www.degruyter.com/serial/BT-B/html>

19 For a selection of ancient Greek and Latin texts in *Google Books*, see <https://www.google.com/googlebooks/ancient-greek-and-latin.html>. On how *Google Books* is "reshaping" the way scholars do research, see Findlen (2013) and Graham/Milligan et al. (2016) 38–44.

Oxford Scholarly Editions Online

LOCK PANELS READ WITH

Main Text **Notes** **Extras**

Otto Skutsch (ed.), *The Annals of Quintus Ennius*

[-] Front Matter

Book 1

Book 2

Book 3

Book 4

Book 5

Book 6

Book 7

Book 8

Book 9

Book 10

Book 11

Book 12

Book 13

Book 14

Book 15

Book 16

Book 17

Book 18

[+] Fragments of the Annals of uncertain origin

Fragments of unknown works perhaps attributed to the Annals

Dubious Fragments

Vestigial Fragments

Spurious Fragments

[+] End Matter

ager *Romulus primam*
diuitias in partu erit, a quo
tribus appellata Titentium,
Raminum, Lucorum:
nominate, ut ait Ennius,
Titentes ab Titio,
Raminetes ab Romulo,
Luceres, ut hantia, ab Lucumone.

109 104 O Tite, tute, Tati, ubi tanta,
tyrannae, tulisti

110 105 Pectora... tenet desiderium;
simul inter
Saepe sic memorant: 'O Romule,
Romule die,
Qualem te patriae custodem di
generum!
O pater, o genitor, o sanguem dis
ortusdum!
Tu prodixisti nos intra luminis
oras

111 110 Romulus in caelo cum dis
genitalibus aeuom
Degit

112 (derui) c' ma

105, inter: the synsphaeia which permits a proposition to occupy the end of the line, though rare even in scolic metres, seems to belong there rather than to the hexameter; see Soph. OC 899 f.; Anon. Fig. 6. 1 (II 414 Kock). Plautus does not know it but Terence has it, with *ae Andr.* 659; *Eam.* 631; 859; with *ae* probably *Eam. prol.* 7. Ennius seems to have it in *trag.* 142 and may have transferred the practice to the hexameter; *inter* is so found later in *Ving. genev.* 3. 229; *Ant.* 10. 899; *Hor. sat.* 1. 7. 11; and, in the dactylic, *Prop.* 4. 2. 3. See Norden, *Aen.* 273–402; Wackernagel, *Syll.* II 196; 1; Marouzeau, *REJ.* 25 (1947) 305 ff.

106 memoriam, to die: *ae ut* die *ae ut die ut diem cond.* *Lact.*

106. Romule, Romule die: the doubling of the name here probably has nothing to do with the doublings found in prayer and magic, discussed by Norden on *Aen.* 6. 46. Whilst the doubling there emphasizes the correctness or importance of the name or word itself, here the emphasis seems to belong to the attribute added in the repetition.

die: on the meaning of the word see the note on 60. The original Latin meaning 'of the sky' may be thought to creep through in *Lucret.* 1. 22 *alios in luminis oras* ('bright' Bailey);

On 1 f. 54 see Farrell and Nelis, *Augustan Poetry and the Roman Republic* (2013), 197

On 1 f. 54 see Farrell and Nelis, *Augustan Poetry and the Roman Republic* (2013), 181 n. 4

On 1 f. 54 see Green, *Discourse and Discretion in Roman Astrology: Manilius and His Augustan Contemporaries* (2014), 155–7

On 1 f. 54 see Green, *Discourse and Discretion in Roman Astrology: Manilius and His Augustan Contemporaries* (2014), 167

On 1 f. 110 see Green, *Discourse and Discretion in Roman Astrology: Manilius and His Augustan Contemporaries* (2014), 155–7

On 1 f. see Goldschmidt, *Shaggy Crocus: Ennius' Annales and Virgil's Aeneid* (2013), 40

On 1 f. see Goldschmidt, *Shaggy Crocus: Ennius' Annales and Virgil's Aeneid* (2013), 214

On 1. 1 see Elser and Hernández Lobato, *The Poetics of Late Latin Literature* (2016), 320

On 1. 34 see Elser and Hernández Lobato, *The Poetics of Late Latin Literature* (2016), 331

On 1. 72 see Kraglund, *Roman Historical Drama* (2014), 43

Copyright © 2018. All rights reserved. Cookie Policy Privacy Policy Legal Notice Credits

Access is brought to you by Monica Berti. Log out

Oxford Index About the Index Show related books Search across all sources in Oxford Index

Figure 2.8. OSEO: O. Skutsch (ed.), *The Annals of Quintus Ennius*, II. 105–109

Needless to say that this is the first step toward an extraordinary contribution to the preservation of an inestimable patrimony of past scholarship, which is often neglected, not only because it is considered old and out-of-date, but also because in many cases it is difficult to locate and consult in traditional libraries. An example, among many others, is represented by the five volumes of the *Fragmenta Historicorum Graecorum* (FHG) that will be described in chapter 4. These projects offer not only the possibility to consult and read scholarly printed editions, but also to get source files for experimenting with Optical Character Recognition (OCR) techniques. As for ancient Greek sources, this kind of experimentation is being performed by the *Open Greek and Latin* (OGL) project and by *Lace: Greek OCR*, which is directed by Bruce Robertson at Mount Allison University.²⁰ The project has an online catalog with an updated list of OCR'd texts including editions of fragmentary authors and works. Through the list, it is possible to access single books, visualize the alignment of the image of each page with its OCR stages, download the relevant files, and also contribute with manual OCR post-correction.

²⁰ On OGL, see p. 30. On OCR for ancient Greek and on *Lace: Greek OCR*, see Robertson (2019).

2.1.1 Thesaurus Linguae Graecae (TLG)

As I mentioned in chapter 1, since the beginning the TLG has addressed the problem of dealing with lost authors in the *Canon* and in the online collection. As far as the catalog and the publication of textual fragments are concerned, the TLG follows standards and conventions of printed editions of fragmentary authors and works.²¹ An example is the fragmentary historian Hecataeus of Miletus, who corresponds to `tlg0538`. In the TLG Hecataeus is classified as *historicus*, is dated between the 6th and the 5th century BC (6–5 B.C.), and has the geographical epithet *Milesius*. In this case, the reference edition is FG_RHist 1 and the TLG follows the arrangement of Felix Jacoby with his distinction between 25 *testimonia* (`tlg0538.001`) and 373 *fragmenta* (`tlg0538.002`).²²

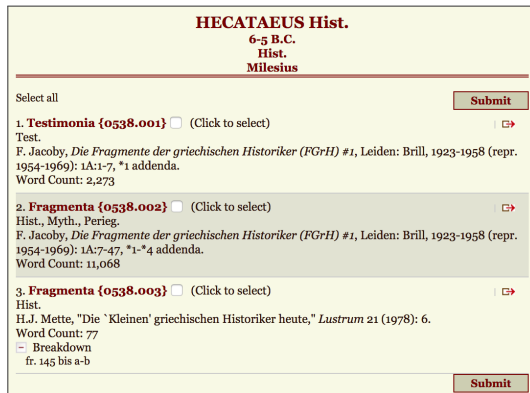


Figure 2.9. TLG: Hecataeus Milesius (`tlg0538`)

The TLG adds two other witnesses to Jacoby's fragment 145 that were published by Hans Joachim Mette as 145bis (a and b) (`tlg0538.003`).²³ Figure 2.9 shows the pop-up window of Hecataeus with bibliographic references and the word count of the texts of testimonies and fragments, excluding the apparatus criticus and the commentary that are not reproduced in the TLG. Figures 2.10 and 2.11 show how Hecataeus' fragment 145bis was originally published by Mette and how is replicated in the online TLG, which closely reproduces the layout of the printed page and adds symbols to mark lines in its files: `l1` marks the line with the reference to the witness until the colon, `1` marks the line with the Greek text of the

21 For a list of epithets and work classifications that identify fragmentary authors and works belonging to different literary genres in the TLG, see pp. 21 and 23.

22 For the *fragmenta* of Hecataeus, the TLG provides further classifications (*Historica*, *Mythographa*, and *Periegesis*).

23 Mette (1978) 6.

fragment, and @1 stands for *page end*.²⁴ The TLG provides a link to cite the page with the edition of Mette, which incorporates author and work numbers (0538 and 003): <http://stephanus.tlg.uci.edu/Iris/Cite?0538:003:0> (fig. 2.12).²⁵ As part of the online TLG services, it is also possible to perform morphological analyses and obtain statistics about the words of the texts of the fragments (see below).

145 bis a) HERODIANOS Καθολική προσωοδία 7, Cod. Vindob. Hist. Gr. 10 fol. 6^r. ed. H. Hunger, Jahrb. Österr. Byz. Gesellsch. 16, 1967, 16 (10): ... ταῦτα μὲν ἔφαμεν παροξύνεσθαι, λέγω δὲ τὸ 'ἀπλόος', 'διπλόος', 'τριπλόος' καὶ ὅσα ἐστὶ τοιαῦτα. ὅτι γὰρ οὐκ ἐστὶ σύνθετα, ἐν τῇ εἰς '-ους' καταλήξει ἔροῦμεν. τὰ μέντοι κύρια, εἰ καὶ σπάνια εὐρέθῃ, προπαροξύνεται, ὥσπερ ἔχει τὸ 'Σίγγροος'· ἐστὶ δὲ πόλις, ὡς Ἐκαταίος Περιηγήσει Εὐρώπης.

b) ETYM. MAGN. 613, 30 Gaisf. ('Herodian.' I 127, 14 Lentz [hier ასოთოს]): τὰ διὰ τοῦ '-οος' ὀνόματα προπαροξύνεται, οἷον 'Σίγγροος', 'Πείροος'.

Figure 2.10. Mette (1978) 6 = Hecataeus, FGrHist 1 F 145bis

Figure 2.11. TLG: Hecataeus, FGrHist 1 F 145bis = Mette (1978) 6

Figure 2.12. TLG: link to cite Hecataeus, FGrHist 1 F 145bis = Mette (1978) 6

Another example is Hellanicus of Lesbos, who is represented in a similar way in the TLG with a reproduction of the *testimonia* and the *fragmenta* from the FGrHist and Mette (1978). In this case, the TLG allows users to select fragments from the different FGrHist sections of Hellanicus arranged by Felix Jacoby, and

24 @1: “marks end of page in source text regardless of whether page is part of the citation system. All files end in @1”. This method seems to be not consistent in the TLG corpus. Cf. the example of Hellanicus below.

25 On TLG author and work numbers, see p. 19 n. 55. The TLG doesn’t provide identifiers for single fragments: in this case, 003 refers to the edition of Mette and not specifically to fragment 145bis with its two witnesses.

there is also a specific reference to fragments on papyrus (figg. 2.13 and 2.14).²⁶ As far as the layout of the FGrHist page is concerned, the TLG tries to replicate it. Figures 2.15 and 2.16 show how Hellanicus' fragments 2–5 (FGrHist 323a) are represented in the printed edition of Jacoby's FGrHist and in the online TLG. The TLG reproduces the text of the fragments including references to other source texts and follows the FGrHist in printing with spaced-out letters those parts of the fragments that seem to be direct quotations (cf. p. 36). Missing elements are the critical apparatus and therefore line numbers for the text of the fragments, references in round brackets to the corresponding numbers of the fragments in other parts of the FGrHist and in the FHG, references in angle brackets to other fragments of the same FGrHist section, chronological data in the margin of the page, and references to fragmentary work titles (in this case Ἰστορίαι) and to book numbers (in this case B) according to which fragments are arranged in the collection.

HELLANICUS Hist.
5 B.C.
Hist.
Lesbius

Select all

1. **Testimonia {0539.001}** (Click to select) |
 Test.
 F. Jacoby, *Die Fragmente der griechischen Historiker (FGrH) #4, #323a, #687a*, Leiden: Brill, 1923-1958 (repr. 1954-1969): 1A:104-107; 3B:40-41; 3C:412.
 Word Count: 1,374

2. **Fragmenta {0539.002}** (Click to select) |
 Hist., Myth.
 F. Jacoby, *Die Fragmente der griechischen Historiker (FGrH) #4, #323a, #601a, #608a, #645a, #687a*, Leiden: Brill, 1923-1958 (repr. 1954-1969): 1A:107-152, *6-*8 addenda; 3B:41-50, 732-733; 3C:1-2, 190, 412-414.
 Word Count: 17,074
 Breakdown
 fr. 124b (PSI 1173): vol. 1A, p. *6 addenda
 fr. 189 (P. Oxy. 10.1241): vol. 1A, p. 150
 fr. 201 bis (P. Giss. 307v): vol. 1A, p. *7 addenda

3. **Fragmentum (P. Oxy. 26.2442) {0539.003}** (Click to select) |
 Hist.
 H.J. Mette, "Die 'Kleinen' griechischen Historiker heute," *Lustrum* 21 (1978): 7.
 Word Count: 58
 Breakdown
 fr. 133 bis

Figure 2.13. TLG: Hellanicus (tlg0539)

Considering that the TLG is not only a digital collection of texts but also a catalog of ancient Greek works based on printed editions, it would have been better to include at least the reference to fragmentary work titles and book numbers provided by Felix Jacoby, because they are a fundamental element to understand

26 PSI X 1173; P.Oxy. X 1241; P.Giss. 307v; fr. 133bis (= P.Oxy. XXVI 2442). It is not clear the criterion used by the TLG for selecting these papyri, because there are other fragments of Hellanicus preserved on papyrus but not mentioned in the breakdown: FGrHist 4 F 19b (P.Oxy. VIII 1084), F 68 (P.Oxy. XIII 611), and F 197bis (= PSI XIV 1390). There is also an inscription among the *testimonia*: FGrHist 4 T 30 (IG II/III² 2363). On the evidence of these physical fragments, see section 2.1.4.

Volume-Jacoby#-F 1a,4,F
 3b,323a,F
 3b,601a,F
 3c,608a,F
 3c,645a,F
 3c,687a,F

Fragment line

HELLANICUS, *Fragmen*

1a,4,F.

(1a) SCHOL. APOLL. RHOD. III 1179: **περὶ τῆς Κάδμου εἰς Θῆβας** παρουσίας Ἀντιμάχου ἐν τῇ Συνομιλῇ τῶν Θηβαίων Παράδοξον (III) ἰστορεῖ, καὶ Ἕλληνας ἐν α' Φορωνίδος ἰστοροῦν ὅτι καὶ τοὺς ὀδόντας ἐσπειρε τοῦ δράκοντος κατὰ Ἄρεος βούλησιν, καὶ ἐγένοντο πέντε ἄνδρες ἑνοπλοῖ, Οὐδαῖος Χθόνιος Πέλιος Ὑπερήνωρ Ἐχίων. (1b) —III 1186: τοὺς περιλειφθέντας ἐν τῇ μάχῃ σὺν Κάδμῳ κατοικήσαι ἐν Θῆβας φησί. λέγει δὲ καὶ Ἕλληνας ὅτι Κάδμος ἐξελὼν τοῦ ὄφεως τοὺς ὀδόντας ἐσπειρεν, ἐκ δὲ αὐτῶν πέντε ἄνδρες ἐφύσαν, Οὐδαῖος Χθόνιος Ὑπερήνωρ Ἐχίων Πέλιος, καὶ ὁ μὲν Ἕλληνας μόνους φησὶ <στοιχεῖ> βεβλιωτημέναι. @1 (5)

(2) ATHENAI. IX 410 F: τὸν δὲ τοῖς γεννηθῆσι δάνατα πατὴρα διδόντα κατὰ χειρὸς Ἡρακλεῖ ἴδιον, ἂν ἀπέκτεινε ὁ Ἡρακλεῖς κωνδύλαι, Ἕλληνας μὲν ἐν ταῖς Ἱστορίαις Ἀρχιαν φησὶ καλεῖσθαι· δι' ὃν καὶ ἐξεχώρησε Καλυδῶνος, ἐν δὲ τοῖς δευτέρῳ τῆς Φορωνίδος Χαυριαν αὐτὸν ὀνομάζει Ἡρόδοτος δ' ἐν ἑπτακαδεκάτῃ τοῦ καθ' Ἡρακλεῖα λόγῳ (31 F 3) Ἐνόμον. (5) καὶ Κίτιον δὲ τὸν Πύλιτος μὲν υἱόν, ἀδελφὸν δὲ Ἀντιμάχῳ ἀπέκτεινε ἄνω Ἡρακλεῖς οἰκοῦντοτα αὐτοῖ, ὡς Νικάνθορος ἰστορεῖ ἐν δευτέρῳ Οἰτακῶν (F 17 Sch.), ὡς καὶ ἀνεπίδοι φησι τέμενος ὑπο τοῦ Ἡρακλεῖος ἐν Πρωσίῳ, ὃ μέχρι νῦν προσαγορεύεσθαι Οἰνοζόου.

(3) HARPOKR. SUID. s. Στεφανηφόρος: ... Στεφανηφόρου ἠρώων, ὡς ἔδουεν, ἦν ἐν ταῖς Ἀθήναις, εἴη δ' ἂν ὁ Στεφανηφόρος ἦτοι τῶν

Figure 2.14. TLG: Hellenicus' *fragmenta* (tlg0539.002)

44 XI. ATHEN

*Ἄρεος δέ, ἐπεὶ τὰ φρονεῖ διαίκα, ὃ δὲ Ἄρης ἐπὶ τῶν φόνων ἢ ἐπὶ ἐπιπέ τὸ δόμο ἐκεί ὃ Ἄρης ἐν τῇ πρὸς Ποσειδάωνα ὑπὲρ Ἀλεξοῦ διηγεῖ, ἐπε ἀπέκτεινε αὐτὸν βασιλευσὶν Ἀλιππῆρ, τὴν αὐτοῦ καὶ Ἀργαῶος τῆς Κίερος ποσειδάων, ὡς φησὶ Ἕλληνας ἐν ἄ.

5 2 (39; 65) HARPOKR. s. v. Παναθήναια: Διμοσθένης Φιλιστοῦς (4, 35), διὰ τὰ Παναθήναια ἤγετο Ἄθηναι, τὰ μὲν καθ' ἑαυτὸν ἐπιεικῶς, τὰ δὲ καθ' ἑκαστοῦ ἐπιεικῶς, ἄπειρ καὶ μεγάλα ἐκάλου· Ἰσοκράτης Παναθηναϊκῆς (12, 17), ... ἤγαγε δὲ τὴν ἑστῆν πρῶτος Ἐργυρόβουλος ὃ Ἡρακλεῖος, καθὰ φησὶν Ἕλληνας τὰ κατ' Ἀθήνας (324 F 2), ἐκείνους 10 ἐν ἄ Ἀθήνας, πρὸ τούτου δὲ Ἀθήναι ἐκαλεῖτο, ὡς βεβλιωσεν Ἰστῆρος ἐν γ τῶν Ἀττικῶν (334 F 4).

3 (40; 66) HARPOKR. s. v. Φορωνίδος: Ἰππερὶς ἐν τῷ Κατὰ Περσολόγῳ (F 145 Bk1.), ἐπὶ τῷ Ἀθήναι Φορωνίδος ἀνομιλῆσθαι ἀπὸ Φίβραντος βασιλεύσαντος Κοσμήτων καὶ ἰπ' Ἐργυρόβου ἀνακρίνεσθαι, βεβλιω- 15 καν Ἄνδρον ἐν γ τῶν Συγγενῶν (10 F 1), ἦν δὲ Ποσειδάωνος υἱὸς ὃ Φέρβας, καθὰ φησὶν Ἕλληνας ἐν ἄ Ἀθήνας.

4 (41; 101) PRIOR. BEOL. p. 53, 21 ReI (SYNAS. LEX. p. 362, 24 Bk1): Ἄλιον τὸ κρὸς ὀδόντας· Ἐπαιτός δὲ παντός (I F 167) καὶ Διονύ- 20 σιος (687 F 3) καὶ Ἕλληνας ἐν ἄ Ἀθήνας καὶ Τίμας (566 F 76) καὶ Ἐρῆδος (p. 75 G1s.).

<F 14-19 Thesaur>
 <F 20-21 Trojanischer Krieg>
 <F 22 König Demophon>
 <F 23 Melanthos, Kodros, Medon>
 25 <F 11 Ionische Wanderung>

B


5 (42; 71) a) HARPOKR. (SUDA; SCHOL. DEMOSTH. 18, 107a) s. v. 5221+ Μουσική: Διμοσθένης ἐν τῷ Ἰππὶ Κερκυράντος (18, 107), τόπος παρα- 30 θαλάσσιος ἐν τῇ Ἀττικῇ. Ἕλληνας δὲ ἐν β' Ἀθήνας ὀνομάσθαι φησὶ ἀπὸ Μουσίου τοῦ βασιλέως τῶν Περσολόγῳ. b) SCHOL. DEMOSTH. 18, 107b: Μουσική τόπος περὶ τὴν Πηραιά, εἶθε ἐπὶ τὴν Μουσικήν Ἀγρέτιδος . . .

1 ἐπεὶ — Ἄρης Suid ἐπὶ τὰ φρονεῖ, διαίκα F ὃ Ἄρης Σ (B1) 2 ὃ Ἄρης om. Suid
 4 Φορωνίδος Suid Φορωνίς Σ τὴν Φορωνίδα Ἄρεος καὶ ἰπ. K. Φορωνίς Et 7 πανθη-
 ναϊκῶς A Epit Suid πανθηναϊκῶς r 8 πρῶτος Epit Suid 4 Harp 12-16 ohne die
 citate Epit Suid Et. M. p. 298, 28 12-13 Ἰππερὶς ἐν τῷ K. II. ὀνομάσθαι λέγει
 ἀπὸ κτ. Et. 14 ἰπ' ὄσο BC 19 ἐκ ἄ — Τίμας om. Σ 29 ἦν β' Ἄρη, om. Suid
 30 Μουσίου Suid Suid Lex. rhet. p. 279, 23 Bk1 μουσικήν Harp Παναθηναϊκῶς
 (metr. BC) Harp Suid (metr. BC) Lex om. Schol

Figure 2.15. Hellenicus, FGrHist 323a 42

Volume-Jacoby#-F fragment line

[Prev](#) | [Next](#)

HELLANICUS, *Fragmenta*. {0539.002} 

λο[υ]μένην *| αρτιδ[.] ἐ(τι) Πτερό(ην) πῖν τ[*] / Πτερόα δ(ἐ) καλεῖτ(αι) ἢ μη[*] / θεν ε(ίς) Πτερό(αν *) / ε(ίς?) Κάροισαν δ(ἐ) ἐπλευσε[*] / οἱ ἐν (10) Περσίδι κ(αι) οἱ ἄλλη[*] *| πῖν ἂν ὀφ(θῆ) μ[ε]ρ[ὸ]ν ἀρχεμόνα / Ἀθην(α)ίως ἐκεῖ γ(άρ) .[*] / τ(ῆς) ἀλογίας (?) ἢ Ἀθη[ν]ά *| / ὠν ἀ(πό) Αἰξ(ω)ν(ῆς?) ἐκεῖ*** @1 (201 ter) NATAL. COM. *Myth.* 7, 2 (p. 706 ed. Geney.): *fuit enim roetatum consuetudo, ut fluvios tauri similes effingerent, quoniam cum impetu irrupentes tauris similem edant mugitum; vel ut Hellenicus sensit, quia terram sulcare tanquam boves apparent; vel ut aliis placuit, quia circa ripas fluminum tauri mugire ob uberiora pascua audiantur.* (5)

(202) JOSEPH. AJ I 107 (EUSEB. P E IX 13 p. 415 D): Ἡσίοδος τε καὶ Ἐκαταῖος (1 F 35) καὶ Ἑλλάνικος καὶ Ἀκουσίλαος (2 F 46) καὶ πρὸς τοῦτους Ἐφορος (II) καὶ Νικόλαος (II) ἱστοροῦσι τοὺς ἀρχαίους ζήσαντας ἐτη χίλια.

3b,323a,F.

(1*) SYNAG. LEX. p. 444, 1 Bkr Ἄρειος πάγος· δικαστήριον Ἀθήνησιν ἐκλήθη δὲ Ἄρειος πάγος ἦτοι ὅτι ἐν πάγῳ ἐστὶ καὶ ἐν ὕψει τὸ δικαστήριον. @1 Ἄρειος δέ, ἐπεὶ τὰ φονικά δικάζει, ὁ δὲ Ἄρης ἐπὶ τῶν φόνων· ἢ ὅτι ἐπῆξε τὸ δόρυ ἐκεῖ ὁ Ἄρης ἐν τῇ πρὸς Ποσειδῶνα ὑπὲρ Ἀλωροθίου δίκῃ, ὅτε (5) ἀπέκτεινε αὐτὸν βιασάμενον Ἀλκίπτην, τὴν αὐτοῦ καὶ Ἀργαύλου τῆς Κέρκροπος θυγατέρα, ὡς φησὶν Ἑλλάνικος ἐν ᾱ.

(2) HARPOKR. s.v. Παναθήναια· Δημοσθένης Φιλίπποιος (4, 35). διττὰ Παναθήναια ἤγετο Ἀθήνησι, τὰ μὲν καθ' ἕκαστον ἐνιαυτὸν, τὰ δὲ διὰ πενταετηρίδος, ἅπερ καὶ μεγάλα ἐκάλουν· Ἰσοκράτης Παναθηναϊκῶι (12, 17) ἦγαγε δὲ τὴν ἑορτὴν πρῶτος Ἐριχθόνιος ὁ Ἥφαιστου, καθὰ φησὶν Ἑλλάνικος τε καὶ Ἀνδροτίων (324 F 2), ἐκάτερος (5) ἐν ᾱ Ἀτθίδος. πρὸ τούτου δὲ Ἀθήναια ἐκαλεῖτο, ὡς δεδήλωκεν Ἴστρος ἐν γ τῶν Ἀττικῶν (334 F 4).

(3) HARPOKR. s.v. Φορβαντεῖον· Ὑπερίδης ἐν τῷ Κατὰ Πατροκλέους (F 145 B1—Je.). ὅτι τὸ Ἀθήνησι Φορβαντεῖον ὀνομάσθη ἀπὸ Φόρβαντος βασιλεύσαντος Κουρήτων καὶ ὑπ' Ἐρεχθῆος ἀνααιρεθέντος, δεδήλωκεν Ἀνδρῶν ἐν ἡ τῶν Συγγενειῶν (10 F 1). ἦν δὲ Ποσειδῶνος υἱὸς ὁ Φόρβας, καθὰ φησὶν Ἑλλάνικος ἐν ᾱ Ἀτθίδος. (5)

(4) PHOT. BEROL. p. 53, 21 Rei (SYNAG. LEX. p. 362, 24 Bkr): Αἶμον τὸ ὄρος οὐδέτερος· Ἐκαταῖος διὰ παντός (1 F 167) καὶ Διονύσιος (687 F 3) καὶ Ἑλλάνικος ἐν ᾱ Ἀτθίδος καὶ Τίμαιος (566 F 76) καὶ Εὐδοξος (p. 75 Gis.).

(5a) HARPOKR. (SUDA; SCHOL. DEMOSTH. 18, 107a) s.v. Μουνυχία· Δημοσθένης ἐν τῷ Ὑπὲρ Κτησιφώντος (18, 107). τόπος παραθαλάσσιος ἐν τῇ Ἀττικῇ. Ἑλλάνικος δὲ ἐν β̄ Ἀτθίδος ὀνομάσθαι φησὶν ἀπὸ Μουνίχου τινὸς βασιλέως τοῦ Παντακλέους. (5b) SCHOL. DEMOSTH. 18, 107b: Μουνυχία τόπος περὶ τὸν Πειραιᾶ, ἔνθα ἐστὶν ἱερὸν Μουνυχίας Ἀρτέμιδος @1

Figure 2.16. TLG: Hellenicus' *fragmenta* (FGrHist 323a)

the content of the fragments and the editorial structure of the FGrHist.²⁷ Nevertheless, the online TLG provides a very powerful search engine and useful tools for performing morphological analyses and getting statistics and n-grams, which represent an invaluable addition to printed editions of ancient source texts. For example, by selecting the word Ἀτθίδος in F 2 of Hellanicus (FGrHist 323a), it is possible to get not only the morphological analysis of the word with links to the entry in different lexica (LSJ, MiddleLS, and DGE), but also statistics about the use and the distribution of the word in the TLG corpus: A summary of the use of the lemma (fig. 2.17), its distribution by century, its relative distribution by century (fig. 2.18), its highest use by author, and its relative distribution by author. The TLG provides also *work statistics* and in this case, for example, it is possible to visualize results for the FGrHist fragments of Hellanicus reproduced in the TLG (τlg0539.002) (fig. 2.19).²⁸

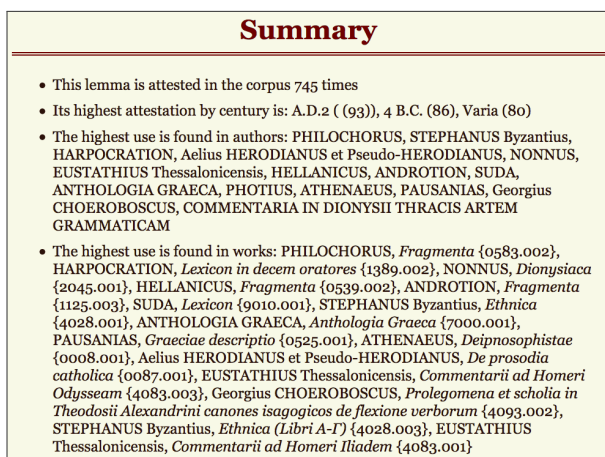


Figure 2.17. TLG: statistics for the lemma Ἀτθίδος, -ίδος, ἡ (summary)

These examples, even if limited to historians, show how fragmentary authors and works are accessible through the TLG. The same dependency on the printed editorial practice is evident if we examine fragmentary authors belonging to other literary genres in the TLG. Moreover, one of the downsides of this structure of the TLG is the fact that the corpus has duplicates of texts. This phenomenon is intrinsic to scholarship of fragmentary literature, because “a collection of frag-

27 To our knowledge, titles (*tituli*) of fragmentary works are present in the TLG if they are the unique evidence of a fragmentary work: cf. p. 19. As for fragmentary historians, an example is the TLG author Promathidas (τlg2300) with one *titulus* (Ἡμίταμβος: τlg2300.003) from Lloyd-Jones/Parsons (1983) 345.

28 It is also possible to get statistics for the entire corpus of Hellanicus or only for the FGrHist *testimonia* (τlg0539.001) and for the *fragmentum* published by Mette (1978) (τlg0539.003).

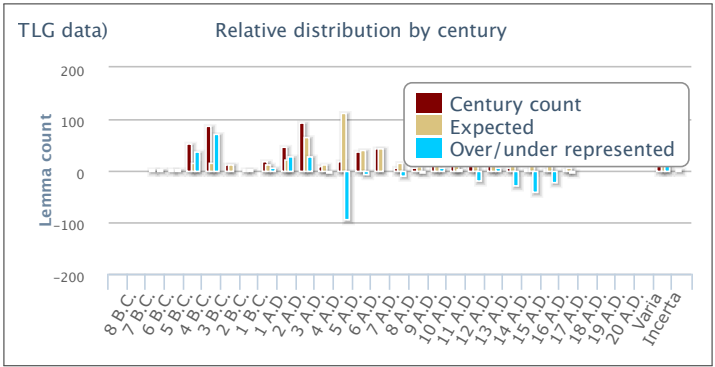


Figure 2.18. TLG: statistics for the lemma ἄτις, -ίδος, ἦ (relative distribution by century)

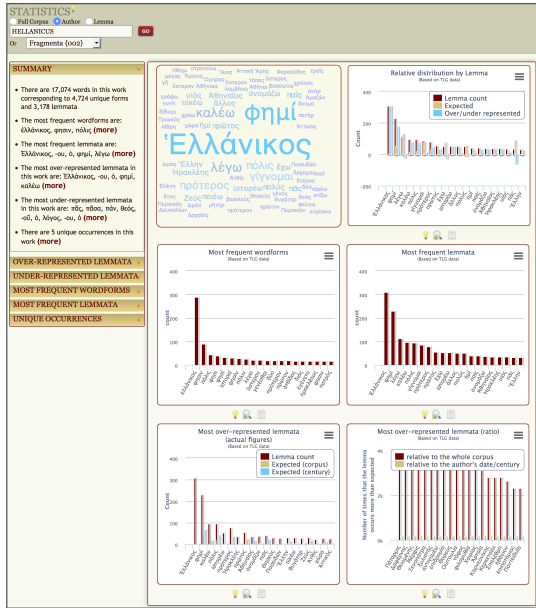


Figure 2.19. TLG: work statistics of Hellanicus' fragmenta (τ190539.002)

ments simply duplicates the information contained in other books in a good library. All the texts it brings are usually available elsewhere.” In a philological world dominated by printed books, this condition is inevitable in order to collect knowledge and “concentrate information otherwise widely disseminated.”²⁹ In a digital world, where resources are more easily accessible and linkable, the inheritance of printed editorial methods by digital libraries is problematic, because the digital duplication of texts generates distorted results.³⁰

The screenshot shows the TLG (Thesaurus Linguae Graecae) interface. At the top, there are search filters for 'Volume-Jacoby#F' (3b,323a,F), 'fragment' (1*), and 'line' (1). Below this, the search results for the N-gram 'ἄττικῶν, δεδῆλωκεν, ἴστρος' are displayed. The results list several sources, including 'HELLANICUS, Fragmenta. (0539.002)', 'Lycurg.Fr. (0034.002)', 'Ister.Fr. (1450.004)', and 'Harp. Lexicon in decem oratores (1389.002)'. Each result includes a brief description and a link to the full text. The detailed view of 'HELLANICUS, Fragmenta. (0539.002)' is shown below the results, displaying the original Greek text and its English translation.

Figure 2.20. TLG: N-grams for ἄττικῶν, δεδῆλωκεν, and ἴστρος

Going back to the examples mentioned above, if we take into consideration F 2 of Hellanicus (FGrHist 323a), we can see how the context of the fragment is repeated several times in the TLG (fig. 2.20). If we activate the TLG n-gram functionality for the string δεδῆλωκεν ἴστρος ἐν γ' τῶν Ἀττικῶν, we can see that the text is repeated four times in the TLG: as Hellanicus FGrHist 4 F 39 (τlḡ0539.002) and 323a F 2 (τlḡ0539.002), as Ister FHG I fr. 7 (τlḡ1450.004), and as Harpocr. Lex. s.v. Παναθήναια (τlḡ1389.002). The text is also repeated under Androton FHG I fr.

29 Most (1997) vii.
 30 Berti/Romanello et al. (2009).

1 (tlg1125.003), but ends before the quoted string.³¹ In all these cases, the only citable evidence is the text of the *Lexicon* of Harpocration, which is the surviving text that reuses the lost passages of Hellanicus, Ister, and Androtion and which is therefore repeated multiple times in the editions of these fragmentary authors. If this situation has the advantage of allowing users to visualize the same text in different editions, the problem is that from a computational point of view these repetitions generate wrong results when querying the TLG corpus. They also produce the wrong impression of the existence of fragmentary texts that, as a matter of fact, don't exist any more but are only preserved through quotations and reuses in other texts.

Fragmentary texts come not only from quotations and text reuses, but also from material fragments like papyri, inscriptions, and excerpts in manuscripts. In the example of Hellanicus mentioned above, there are six fragments preserved on papyrus and one testimony from an inscription (fig. 2.13).³² In this case the texts of the fragments are reproduced following the content and the layout of the FGrHist and of Mette (1978), but there are no links to external resources.³³ Another example is the *Hellenica Oxyrhynchia*, which is a historiographical work preserved only on physical fragments. Also in this case the text is reproduced in the TLG following its reference printed editions (tlg0558).³⁴

2.1.2 Jacoby Online (JO)

Jacoby Online (JO) is a project specifically aimed at ancient Greek fragmentary historians.³⁵ It is maintained by the Dutch publisher Brill and is part of a big scholarly enterprise whose goal is to continue and update the editorial work of *Die Fragmente der griechischen Historiker* (FGrHist) began by Felix Jacoby in the Twenties of last century (cf. pp. 35 ff.). The project is also the result of an ongoing effort to make more user-friendly and accessible the volumes of the FGrHist,

31 The complete text of the lexical entry is published only under the *Lexicon* of Harpocration, while in the other cases the text is partially cut in the same way as it is published in the FHG and in the FGrHist.

32 FGrHist 4 T 30, FF 19b, 68, 124b, 133bis, 189, and 201bis.

33 The text of some of these papyri is available through *Trismegistos* and other digital resources: P.Oxy. VIII 1084 (<http://arks.princeton.edu/ark:/88435/sx61dp87h>); P.Oxy. X 1241 (www.trismegistos.org/text/63428); P.Oxy. XXVI 2442 (www.trismegistos.org/text/62564); PSI X 1173 (www.trismegistos.org/text/61611). The texts of IG II/III² 2363 is available through PHI *Greek Inscriptions* (<https://inscriptions.packhum.org/text/4599>). On these and other digital resources for material fragmentary texts, see section 2.1.4.

34 Bartoletti (1959); FGrHist 66; Mette (1978) 11–12. Absent from the TLG is the text of the *Marmor Parium*: see p. 8.

35 <http://referenceworks.brillonline.com/cluster/JacobyOnline>

which is a notoriously difficult tool to consult, especially for students of Classics. The first steps of this effort were the publication of complete indices to FGrHist I–III by Pierre Bonnechère and the production of a Windows compatible CD-ROM version of the fifteen volumes of Jacoby’s FGrHist I–III and of Bonnechère’s indices.³⁶ The CD-ROM was welcomed as an “excellent tool” and as “a miracle of 20th-century scholarship with a miracle of 21st-century technology.”³⁷ For the first time, users had the possibility to search numerical and alphabetical lists of the 856 FGrHist authors and to be immediately brought with one click to the relevant part of the collection concerning the requested historian.³⁸ Those who have spent many hours in the library to consult the printed version of the FGrHist – looking for information about authors scattered in the fifteen volumes of the collection – greatly enjoyed the advantages of a single CD-ROM with search and hypertextual functionalities, and the possibility to visualize notes of the critical apparatus in dialogue babbles appearing on lines with textual problems and variants.³⁹ The CD-ROM version was based on the layout of the printed edition (“page-based”) and the aim was to produce, as far as possible, an exact representation of the printed volumes.⁴⁰

The CD-ROM version of the FGrHist is now superseded by the online edition which is part of the *Jacoby Online* project. The current online version is under revision and a new interface is going to be launched as part of a collaboration with the company *Eldarion*, that has also developed the *Scaife Viewer* for the *Perseus Digital Library*: <https://scaife.perseus.org>. An overview of the new version of the *Jacoby Online* will be described in this section after a description of the current version, which has been used by many scholars in the last ten years.⁴¹

36 Bonnechère (1999) and Jacoby (2005). For reviews of the CD-ROM, see Marincola (2005), Worthington (2005), Walter (2005), and Cornell (2006).

37 Marincola (2005) and Cornell (2006) 186.

38 The home page and the booklet accompanying the CD-ROM contained a detailed history of Jacoby and his work written by Mortimer Chambers, which is now available as part of the online version of the *Jacoby Online*: http://dx.doi.org/10.1163/1873-5363_boj_abiografia_jacoby.

39 Cornell (2006).

40 For a detailed description of the CD-ROM, see Marincola (2005), who also points at its limits, such as the price (€ 1,500) and the fact that pieces of information on the margins of the FGrHist pages were not interactive and that the *addenda* and *corrigenda* sections had to be manually searched.

41 I’m very grateful to Ernest Suyver and Mirjam Elbers for giving me access to the demo version of the new *Jacoby Online* and in general for the opportunity to work with them as a contributor and copy editor of the *Jacoby Online* project: see section 2.1.3. This collaboration now also includes the connection between the *Jacoby Online* and the *Digital Fragmenta Historicorum Graecorum* (DFHG) project, that I have been implementing and that is described in chapter 4.

In order to be accessed and consulted, the *Jacoby Online* project requires a subscription and includes five sections:⁴²

1. *Die Fragmente der griechischen Historiker* Part I–III. This is the online version of the first three parts of the FGrHist originally published by Felix Jacoby. It gives access to alphabetical and numerical lists of fragmentary authors, to the “commentary on cities and peoples,” and to the *addenda*. In the “Prelims,” it is possible to consult an introduction to the life and the work of Felix Jacoby, notes for readers (abbreviations, *corrigenda*, and indexes), and prefaces. Texts of fragments are reproduced as they appear in the FGrHist, together with commentaries and notes. Introductions, *testimonia*, *fragmenta*, commentaries, notes, and selected *addenda* to every author are published together in the same web page. A menu on the upper right part of the page contains links to each testimony and fragment. Introductions to authors contain also note numbers, but without links to the actual notes that seem not to have been included in the online collection. Figure 2.21 shows the example of FGrHist 323a F 2 in the *Jacoby Online* that can be compared with the printed page at figure 2.15. References to corresponding FHG and FGrHist fragments have been removed, as well as the notes on the margins of the printed pages. Another difference with the printed edition is the addition of links to the corresponding BNJ fragments (see below). Critical notes to the text of the fragments are reproduced in footnotes at the bottom of the web page and have numbers following the numerical sequence of the notes to the commentary. Those parts of the fragments that Jacoby considered to be direct quotations are reproduced with orange coloured letters and not with spaced-out letters as in the FGrHist. Search and index functions are shared with other parts of the *Jacoby Online*.⁴³

2. *Die Fragmente der griechischen Historiker* Part IV. This is the online version of the FGrHist Continued on ancient Greek biography and antiquarian literature that Felix Jacoby was never able to publish. According to the plan, this section will consist of 27 book volumes, some of which have been already published in a printed format.⁴⁴ Before the final printed publication, fragmentary authors are progressively published online following the editorial guidelines of the *Jacoby Online* project. This means that, after each introduction to authors with interactive footnotes, there is a “brief encyclopaedia-style entry” with chronological, literary, and geographical metadata. A similar entry is at the beginning of each testimony and each fragment with information on the witnesses, their

42 As of 2021, the “online subscription price” is € 1,1761 with an “annual update fee” of € 1,285. The “institutional outright purchase price” is € 20,332.

43 http://dx.doi.org/10.1163/1873-5363_bnj_ahelp

44 Bollansée/Schepens et al. (1998); Radicke (1999); Schepens/Bollansée (1999); Verhasselt (2018); Brusuelas/Obbink et al. (2019); Zaccaria (2021). The editorial plan is available at http://dx.doi.org/10.1163/1873-5363_jciv_fulltextxml_aaboutiv.

FGrH 323a F 2

BNJ
 Harpokr. s. v. *Παναθήναια*
 · Δημοσθένης Φιλίππειος (4, 35), διττὰ Παναθήναια ἦγετο Ἀθήνησι, τὰ μὲν καθ' ἑκάστον ἐνιαυτὸν, τὰ δὲ διὰ πενταετηρίδος, ἅπερ καὶ μεγάλα ἐκάλεον· Ἰσοκράτης Παναθηναϊκῶι (12,⁵⁹ 17) ... ἦγαγε δὲ τὴν ἐορτὴν πρῶτος Ἐριχθόνιος ὁ Ἡραϊστῶι,⁶⁰ καθὰ φησιν Ἑλλάδικός τε καὶ Ἀνθροστῶν (324 F 2), ἐκάτερον ἐν Ἄττιβίος, πρὸ τούτου δὲ Ἀθήναια ἐκαλεῖτο, ὡς δεδήλωκεν Ἰστρος ἐν γ' τῶν Ἀττικῶν (334 F 4).

Commentary F 2

Marm. Par. A 10[ἀφ' οὗ Ἐργχ]θόνιος Παναθηναίους τοῖς πρῶτοις γενομένοις ἄρμα ἔξευξε καὶ τὸν ἀγῶνα ἐπέκεισε καὶ Ἀθηναίους [ὠν]ίμ[ασε]; Philochoros 328 F 8-9. The Athidographers agree in connecting the Panathenaia with Erichthonios. Earlier tradition is lacking; Herodotos, who incidentally mentions the festival in his account of the Peisistratids⁶¹, had no reason for going into its previous history. Neither the silence of tradition nor the omission of Erichthonios in Herodotos (presumably he did not yet distinguish him from Erechtheus⁶²) justifies the idea of Niese⁶³ that H. was the first to establish the tradition as a 'typically democratic narrative which at the same time detracted from the glory of the Peisistratids'. As far as we can judge, the datings back of historical institutions to mythical times are a great deal older than the beginning of Athidography. The tradition about the Panathenaia is treated in detail on Istros 334 F 4.

Figure 2.21. *Jacoby Online*: Hellenicus, FGrHist 323a F 2

chronology, language, and literary genre. Every testimony and fragment is accompanied by an English translation and a full commentary (see figure 2.22 to be compared with the printed page at figure 1.12).⁴⁵ Critical notes are expressed in footnotes and there are interactive links to fragments of other parts of the FGrHist and the BNJ. Each author section has a bibliography at the end of the web page. Unlike the FGrHist, which was the work of one scholar, FGrHist IV is the result of a team of researchers working on different authors under the direction of Stefan Schorn and an editorial board. Search and index functions are shared with other sections of the *Jacoby Online*.

3. *Die Fragmente der griechischen Historiker* Part V. This is the part of the FGrHist Continued that concerns geography. The plan is to publish testimonies and fragments of 96 Greek historians with the collaboration of a team of scholars under the direction of Hans-Joachim Gehrke and Felix Maier.⁴⁶ The online editorial layout is the same of FGrHist IV and search and index functions are shared with other sections of the *Jacoby Online*.

4. *Brill's New Jacoby* (BNJ). The BNJ is described by its editor-in-chief Ian Worthington in the home page of the project: "*Brill's New Jacoby* is a fully-revised and enlarged edition of Jacoby's *Die Fragmente der griechischen Historiker* I–III, providing new texts of the ancient historians in many instances as well as several new historians and many new fragments of existing historians that were either unknown to Jacoby or excluded by him. Especially important is that for the first time ever commentaries are provided on the final 248 historians in FGrHist I–

45 Given that this is an ongoing project, it is possible to see variants and changes in the editorial treatment of authors and fragments.

46 http://dx.doi.org/10.1163/1873-5363_jcv_a

About This Historian

Historian:	Stesimbrotos of Thasos
Jacoby number:	1002
Attested works:	<i>On Themistokles, Thukydidēs and Perikles</i> : T 1, F 1, F 10a
Historian's date:	5th century BC, c. 470-425 BC
Historical focus:	IV. Antiquarian History and Biography () A. Biography Pre-Hellenistic Period
Place of origin:	Thasos
Textual base:	

Testimonia and Fragments

FGrHist 1002 T 1

Source:	Plutarch (Ploutarchos), <i>Life of Kimon</i> 4.5
Work mentioned:	
Source date:	1st century AD, c. 46 - 120 AD and century AD
Source origin:	Delphi
Source language:	Greek
Source genre:	Biography - To 500 - Library of Congress History, Ancient - Library of Congress Politics and government - Library of Congress
Fragment subject:	Biography - 500 - Library of Congress
Textual base:	

Plutarch (Ploutarchos), *Life of Kimon* 4.5

Στεσίμβροτος ἁ Θάσιος ἱστορῶν, κατὰ τοὺς αὐτοὺς ἀπὸ τοῦ (sc. Περικλέους) γράβουτος γράμματος καὶ ἑρασίου ἀστῆν, ἐν τῷ ἑξαγράμμῳ ἐπιτὶ Θημοκλέους καὶ Θεουκλίδου καὶ Περικλέους (cf. *Thukydidēs and Perikles*, F 10a).

As recorded by Stesimbrotos of Thasos, a contemporary of his (sc. Perikles), who had seen him, in his book entitled *On Themistokles*.

Figure 2.22. *Jacoby Online*: Stesimbrotos of Thasos, FGrHist 1002

BNJ 334 T 1

Source:	Suda, Lexikon, Ἴστρος
Work mentioned:	
Source date:	10th century AD
Source language:	Greek
Fragment subject:	biography-to 500 - Library of Congress criticism - Library of Congress
Textual base:	Jacoby

Suda, Lexikon, Ἴστρος

Ἴστρος Μενάνδρου ἢ Ἴστρου Κορυθαίης ἢ Κυρηναιανὸν καὶ Μακεδόνα, συγγραφεύς, Καλλιμάχου δοῦλος καὶ γράβουτος. Ἰερμίππος δὲ αὐτὸν φησὶ Πάριον ἐν τῷ β' βιβλίῳ τῆς ἐπιτομῆς. Ἰερμίππος δὲ αὐτὸν φησὶ Πάριον ἐν τῷ β' βιβλίῳ τῆς ἐπιτομῆς. Ἰερμίππος δὲ αὐτὸν φησὶ Πάριον ἐν τῷ β' βιβλίῳ τῆς ἐπιτομῆς. Ἰερμίππος δὲ αὐτὸν φησὶ Πάριον ἐν τῷ β' βιβλίῳ τῆς ἐπιτομῆς.

Istros, son of Menander, son of Istros, a Cyrenaecian and a Macedonian, a prosewriter, a slave and pupil of Kallimachos. Hermippos says in book two of his work about *Slaves Eminent in Learning* that he came from Paphos. He wrote many works both in prose and in verse.

Commentary

The Hermippos mentioned in T 1 is not Istros' colleague at Alexandria, Hermippos of Smyrna. Rather, he is Hermippos of Berytos, a slave by birth who flourished in the time of Trajan and Hadrian as a scholar, writer, and disciple of the historian Hieronymus Philon of Byblus (S. Fornaro, *Heremias Philo**, *BNP* 6 (Leiden, 2005), 189-201). Hermippos of Berytos' works included *Interpreting Dreams* (five books), *On the Number Seven*, and *About Slaves Eminent in Learning* (*FHG* 3, 35; *FGrH Continued* 106 T 1 and T 2). It is the latter work, of course, which includes the reference to our Istros. No doubt, Hermippos of Berytos had a great deal of empathy with Istros of Paphos inasmuch as he considered Istros, like himself, to be both a slave and an eminent scholar.

The two genitives at the beginning of the entry in the *Suda* are problematic, and scholars have tried to understand their meaning. They may refer to the names of the father and grandfather of Istros; alternatively, Ἴστρου is a mistake for Ἴστρουαῖος or Ἴστρουαῖος with a reference to the job of the father Menander. The two forms may also be variants of the name of the father of Istros (son of Menander or son of Istros) or the result of a fusion of two different entries on Istros (son of Menander (Ἴστρος Μενάνδρου) and Istros son of Istros (Ἴστρος Ἴστρου). For these possibilities and bibliography see M. Berti (ed.), *Istro di Callimacheo 1, Testimonianze e frammenti su Istros e sull'Atene* (Tivoli 2009), 2-5.

On the relationship between Istros and Kallimachos, on Istros' role as *synagogueus*, and on the meaning of the adjectives Κορυθαίης (Cyrenaecian) and Μακεδόνας (Macedonian), which may be considered indirect references to the place of origin of his master Kallimachos, see Berti, *Istro di Callimacheo* 1, 4.

Figure 2.23. *Jacoby Online*: Istros, BNJ 334 T 1

III, which Jacoby was unable to prepare before his death. In addition, and also for the first time, *Brill's New Jacoby* presents facing English translations of all the *testimonia* and fragments, new, critical commentaries on all the testimony and fragments, and a brief encyclopedia-style entry about each historian's life and works, with a select bibliography.⁴⁷ Figure 2.23 shows the example of Istros the Callimachean (BNJ 334 T 1). Editors of the BNJ are provided with guidelines and a template to be filled in with metadata and data about fragmentary authors and their works. In order to be consistent, the project offers also lists of subjects for authors, testimonies, and fragments, and special tags are used by copy editors for hyperlinks and anchors of named entities, bibliographic elements, testimonies' and fragments' numbers.⁴⁸ Each entry has a final *Biographical essay* on the fragmentary author and a bibliography.⁴⁹ The BNJ keeps the numbering system of Jacoby and new authors are inserted in the appropriate section with the same number as the preceding author followed by A or B in order to distinguish them.⁵⁰ Following the principle of offering a more user-friendly version of the FGrHist, the BNJ provides not only new commentaries and English translations, but also expanded references to source texts and to bibliographic entries, chronological and literary information, and links and metadata to help readers contextualize fragments and witnesses. Search and index functions are shared with other sections of the *Jacoby Online*.

5. *Brill's New Jacoby*, Second Edition (BNJ2). This part is a "revised and enlarged edition of *Brill's New Jacoby* (BNJ). New additions include an apparatus criticus and a discussion of the provenance of each fragment where relevant, as well as revised commentaries on the ancient historians in BNJ and updated bibliographies, all of which set BNJ2 significantly apart from the previous edition." The online editorial layout is the same of BNJ and search and index functions are shared with other sections of the *Jacoby Online*.

As mentioned before, the current version of the *Jacoby Online* is going to be substituted by a new version in the near future. This new version is presented in a webpage entitled *Documentation for Jacoby Online*, which is maintained by *Brill Scholarly Editions* and published with *GitBook*.⁵¹ Given that this documentation is public, my aim is not to repeat it here, but to summarize two main characteristics of the new *Jacoby Online* there were also discussed as part of a seminar

47 See Worthington (2005) and Lenfant (2009).

48 Lists of subjects grouped under categories are available on the website of the project and can be used to fill in a term in the search category *Subject Keyword* in the Advanced Search: http://dx.doi.org/10.1163/1873-5363_bnj_asubjects.

49 As for FGrHist IV and V, the BNJ is an ongoing project and it is possible to find inconsistencies, errors, and technical problems: see http://dx.doi.org/10.1163/1873-5363_bnj_aprel_iminaries.

50 Worthington (2005).

51 <https://brillpublishers.gitlab.io/documentation-jo/>

Table 2.1. *New Jacoby Online: CITE URNs*

Object	URN	Explanation
textgroup	urn:cite:greekLit:fgrh.1	this is the group of fragments
fragment	urn:cite:greekLit:fgrh.1.F1	this is the first fragment (of the type Fragment)
BNJ fragment	urn:cite:greekLit:fgrh.1.F1.jo-grc2	this is the second edition of this fragment
passage	urn:cite:greekLit:fgrh.1.F1@oi[1]-veloioi[1]	this URN references the phrase <i>οἱ γὰρ Ἑλλήνων λόγοι πολλοὶ τε καὶ γέλοιοι</i> in the fragment
translation of BNJ fragment	urn:cite:greekLit:fgrh.1.F1.jo-eng2	BNJ2 would be jo-eng3
commentary on BNJ fragment	urn:cite:greekLit:fgrh.1.F1.jo-eng5	BNJ2 would be jo-eng3. Hekataios belongs to the German commentaries, as it sits in FGHist I
commentary on textgroup	urn:cite:greekLit:fgrh.1.jo-eng5	This is the commentary in BNJ. As it happens, Jacoby himself wrote no commentary on the entire textgroup, only on its fragments. Unfortunately, the number 5 is confusing, as this is not the fifth commentary. But is done for the sake of consistency. The number always denotes BNJ.

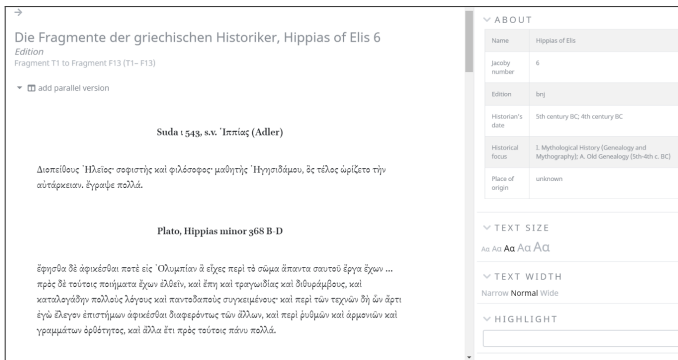


Figure 2.24. New *Jacoby Online* demo (early 2020)

organized at Brill in Leiden on January 22, 2020 with the participation of Stefan Schorn, James Tauber, Ian Worthington, Cecily Robinson, Mirjam Elbers, Ernest Suyver and myself. The first important characteristic in order to combine the needs of the contributors and the needs of the *Jacoby Online* is the substitution of Brill XML files, that were used in the past, with BPT files as a basic format to produce TEI XML files for the publication on *Brill Scholarly Editions* (fig. 2.24).⁵² New extended Guidelines for authors have been written about entry structure, publication statement, the historian, *testimonia*, fragments, biographical essay and bibliography.

The second important aspect of the new *Jacoby Online* is the adoption of the *CITE Architecture* to produce uniform and stable identifiers of the following seven JO objects:⁵³ 1) *textgroup*: a group of fragments, united by origin, theme, and/or (most common) authorship; 2) *fragment*: a textual remnant of an otherwise lost work (“Fragment is a child of textgroup. We could have a collection level = fgrh”); 3) *historian*: author of an (in the case of JO) lost historiographical work (“Historian is metadata about (a version of) a textgroup”); 4) *work*: (in the case of JO) lost historiographical work (“Work is metadata about (a version of) a fragment”); 5) *source*: text (itself a work) containing a citation of or reference to a lost work or its author; 6) *entry*: a textgroup as analyzed by modern scholars (“Entry is metadata about (a version of) a textgroup”); 7) *edition*: edition of the source texts (“Edition is metadata about (a version of) a source”). *Jacoby Online* adopts both CITE and CTS URNs to cite texts, papyri, inscriptions and fragments. The

52 BPT stands for *Brill Plain Text* and the language of these files is Markdown with some Brill-specific extensions. BPT supports the inclusion of additional mark-up, such as Leiden+ for epigraphical texts (http://papyri.info/docs/leiden_plus), YAML for metadata and references, and HTML. Figure 2.24 shows an example of the demo of the new *Jacoby Online* with fragments of Hippias of Elis and the *widget* for the metadata about the historian.

53 On the *CITE Architecture*, see section 3.2.

syntax of each URN includes fgrh as the textgroup identifier and jo as the version identifier. Table 2.1 shows examples for Hekataios of Miletos (FGrHist 1).

Brill's New Jacoby

Istros (334)
(44,99 words)

This entry was prepared by Steve Jackson and Monica Berti and published on 1 October 2015.

About this Historian

Historian:	Istros
Jacoby number:	334
Attested works:	
Historian's date:	3rd century BC
Historical focus:	III. History of Cities and Peoples (Horography and Ethnography) B. Authors on Single Cities and Regions XI. Athens
Place of origin:	unknown

BNJ 334 T 1

Source:	<i>Suda</i> , 1, 796, s.v. Ἰστρῶς (A. Adler (ed.), <i>Suda</i> Lexicon (Leipzig 1928-38))
Work mentioned:	
Source date:	10th century AD
Source language:	Greek
Fragment subject:	biography-to 500 - Library of Congress criticism - Library of Congress

Article Table Of Contents

- T 1 : Suda, Lexikon, Ἰστρῶς
- T 2 : Athenaios, Deipnosophists, 6, 103, 272b
- T 3 : Plutarch (Ploutarchos), Greek Questions, 43, 301d
- T 4 : Scholia, s694
- T 5 : Plutarch (Ploutarchos), On the Pythian Oracle, 19, 493e
- T 6 : Athenaios, Deipnosophists, 9, 38, 387f
- F 1 : Photios, Lexikon (ed. Naber, Reitzenstein), - Ττραῖδα γῆν Ν.
- F 2a : Harpokration Harpokration, Lexicon on Ten Attic Orators, Ααμῆαῖς
- F 2b : Harpokration Harpokration, Lexicon on Ten Attic Orators, Ααμῆαῖς (ex epitoma)
- F 3 : Harpokration Harpokration, Lexicon on Ten Attic Orators, Θεοῖων

Figure 2.25. Istros, BNJ 334

2.1.3 Printed and Digital Fragments: Istros the Callimachean

The edition of the fragments of Istros the Callimachean is an example of a work shifting from printed to digital characteristics. I originally conceived this work in the form of a printed book as part of the Italian series *I Frammenti degli Storici Greci*. The first volume containing the fragments on Athens and Attica was published in 2009 and its characteristics have been described in section 1.4.⁵⁴ I published a second edition with the *testimonia* and all the seventy-seven fragments of Istros the Callimachean in 2015 for the *Brill's New Jacoby*, as the result of an invitation by Ian Worthington to complete the work originally begun by the late Steve Jackson (fig. 2.25).⁵⁵ This edition presents fragments according to the traditional model of printed editions, but with the addition of metadata and hyperlinks that will be further expanded in the revision of the fragments for the second edition of BNJ and as part of the new version of the *Jacoby Online*.

54 Berti (2009b); Berti (2009a); Berti (2013b).

55 BNJ 334: http://dx.doi.org/10.1163/1873-5363_bnj_a334.

FHG_author = "Ister"						
FHG	Volume: FHG 1	Author: Ister	Pages: 418-427	Paper Edition	DFHG	urn:cite:lofts:fhg.1.ister
FGrHist		Author: Istros der Kallimacheer	Number: 334		Jacoby OnLine	
BNJ 1		Author: Istros	Number: 334		Jacoby OnLine	
Perseus Catalog		Author: Ister Cyrenaesus			Perseus Catalog Entry	
FHG	Volume: FHG 1	Author: Ister	Pages: 418-427	Paper Edition	DFHG	urn:cite:lofts:fhg.1.ister
FGrHist		Author: Istros	Number: 1768		Jacoby OnLine	
Perseus Catalog		Author: Ister Cyrenaesus			Perseus Catalog Entry	

Figure 2.26. DFHG: Ister (concordance of editions)

As it will be extensively described in chapters 4 and 5, one of the goals of new born-digital editions of fragmentary texts is to expand and connect resources. As far as Istros is concerned, the first step was accomplished by producing the digital version of its fragments published by Karl Müller in the *Fragmenta Historicorum Graecorum*, which is fundamental to understand the edition of Felix Jacoby in *Die Fragmente der griechischen Historiker*. Figure 2.26 shows the concordance among the entries of Istros in different resources as part of the *Müller-Jacoby Table of Concordance* of the *Digital Fragmenta Historicorum Graecorum* project (see section 4.4.3). Future work will conceive the edition of the fragments of Istros the Callimachean not as the extraction of chunks of text (*fragmenta*), but as the annotation of pieces of information concerning him and his works in the context of surviving sources.⁵⁶ The lack of digital versions of all the sources that preserve testimonies and text reuses of Istros is still a limit for the accomplishment of this task, but preliminary results are now available through the *Digital Athenaeus* project (see chapter 5).

Figure 2.27 shows the lemma Ἴστρος and its inflected forms in the *Named Entities Concordance* of the *Deipnosophists*.⁵⁷ The concordance highlights the name of Istros (**red**) (and the homonymous river) and other named entities (**blue-green**) pertaining to him, such as other authors, *ethnica*, and work titles that are cited in the immediate context. The extraction and annotation of Named Entities pertaining to text reuses of lost authors is the beginning of a new philological practice that will enable scholars to produce new digital and dynamic editions of fragmentary authors and works within their context of transmission.

56 For the description of this model, see section 3.1 and Berti (2019c).

57 For a detailed description of this resource, see section 5.6.2.

Digital Athenaeus

G. Kaibel: Athenaei Naucraticae Dipnosophistarum libri 15

Named Entities Concordance

Lemma: Ἰστρος [Istros]
lemma in LOGEION

Inflected form: Ἰστρον [Istros] - search in Kaibel Q
inflected form in TLG
Named Entity Class: LOC

6. 25	[αὐτοῦς διόκειεν ἐπὶ τοῖς περὶ τὸν Ἰστρον τόπους ... ἀφ' ἧς καὶ τῆν]
7. 88	[σκεναζόμενος, παραπίστιος ὅν τῷ κατὰ τὸν Ἰστρον γινόμενῳ γλάνιδι. φέρεται δ' ὁ Νεῖλος]

Named Entity Class: PER

9. 38 [οἰκίσματα ἀποπέμφω, ὡς Πολέμῳ ὁ περιηγητὴς Ἰστρον τὸν Καλλιμάχῳ συγγραφέα εἰς τὸν ὀνόμαον]

Inflected form: Ἰστρος [Istros] - search in Kaibel Q
inflected form in TLG
Named Entity Class: PER

3. 6	[μὲν, νῆ τὸν Δία, πάνν φέρεται. Ἰστρος δ' ἐν τοῖς Ἀττικαῖς οὐδ' ἐξάγεσθαι]
------	---	--	---

Named Entity Class: LOC

3. 88 [οὗτος ἐδέξατο ἄντακαον, ὃν τρέφει μέγας Ἰστρος Σκῶτατον ἡμίτηρον ἠδονήν. καὶ τὸν Μενδίσαν]

Named Entity Class: PER

6. 103	[ὁ Ἐπιτίμαχος (οὗτος δ' αὐτὸν καλεῖ Ἰστρος ὁ Καλλιμάχος ἐν ταῖς αὐτὸν]
8. 35	[μῖαν ἡμέραν μὴ δεδύνησαι ἐνεγκεῖν ὄψοφαγίαν; Ἰστρος δέ φησι Χαμῖλον τὸν ποιητὴν παρ']
11. 55	[ἐπ' ἄλλοσθην οἶμον ἔβαινε πόδα. ΚΟΝΩΝΕΙΟΣ. Ἰστρος ὁ Καλλιμάχος ἐν πρώτῳ Ἱππολεμειδῶς τῆς]
13. 4	[Ἐλένην ἄρπασας ἐξῆς καὶ Ἀριάδην ἤρπασεν. Ἰστρος γοῶν ἐν τῇ τεσσαρεσκαιδεκάτῃ τῶν Ἀττικῶν]
14. 63	[ἐπιδαψιλεύειν ἐν αὐτῇ τῷ φυτῶν, φησὶν Ἰστρος ἐν τοῖς Ἀργολικοῖς ὅτι δέ τας]

Figure 2.27. *Digital Athenaeus*: NEs concordance (Istros)

2.1.4 Digital Collections of Physical Fragmentary Texts

The expression *fragmentary texts* refers not only to quotations and text reuses, but also to physical fragments that bear textual evidence, which includes many examples of literary texts. As we have seen before, the TLG and the *Jacoby Online* collect also texts of ancient Greek authors preserved on material fragments. Scholars interested in getting more information about this type of sources have now at their disposal many digital projects and resources devoted to collecting data about physical fragments (inscriptions, papyri, manuscripts, etc.). A complete description of these resources is beyond the scope of this volume, but, in order to be up to date with them, I refer to the *Digital Classicist Wiki*, which is a hub for collecting guidelines, suggestions, and catalogs of digital projects concerning the Graeco-Roman world.⁵⁸ In this section, I limit my presentation to the most important digital reference tools that can be used in a scholarly work about ancient Greek fragmentary authors and works.

58 See <https://wiki.digitalclassicist.org> and also Babeu (2011). On papyrological resources, see Reggiani (2017) and Reggiani (2018). On epigraphical resources, see now <http://epigraphy.info/> and De Santis/Rossi (2019).

The *PHI Greek Inscriptions* is an electronic database produced by the *Packard Humanities Institute*.⁵⁹ It offers a comprehensive collection of searchable Greek inscriptions arranged by ancient regions and modern scholarly collections. The reconstructed text of the inscriptions is reproduced according to the main *corpora* and reference printed editions, but without the critical apparatus and the commentaries. This resource allows users to find the text of inscriptions that have been classified also as *testimonia* or *fragmenta* of fragmentary authors. Examples are IG II/III² 2363, IG XII 5, 444, and IG XIV 1293. IG II/III² 2363 is a 2nd-1st century BC inscription from Piraeus with a catalog of mostly Attic writers including Hellanicus (col. 2, 4 = FGrHist (BNJ) 4 T 30).⁶⁰ IG XII 5, 444 is a 3rd century BC inscription with the text of the *Marmor Parium*, which is a historiographical fragmentary work attested only on stone (= Jacoby (1904) and FGrHist (BNJ) 239).⁶¹ IG XIV 1293 is a marble plaque (*Tabula Albana*) with inscriptions of uncertain provenance and date that preserve an anonymous history of Heracles (FGrHist (BNJ) 40 F 1).⁶² The complete texts of these inscriptions are available through the *PHI Greek Inscriptions* website (fig. 2.28).⁶³ The resource provides links to other publications within the PHI database, but not to external resources. Each inscription has a unique reference number, which is also embedded in a stable URL.

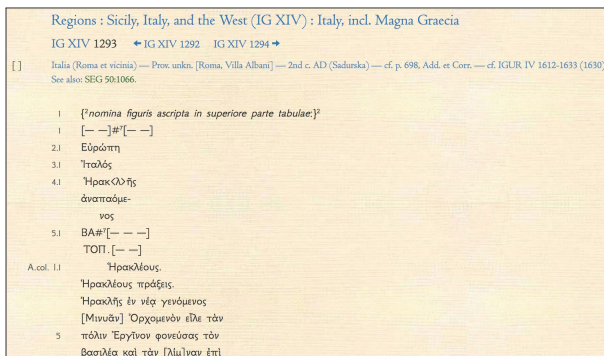


Figure 2.28. *PHI Greek Inscriptions*: IG XIV 1293

Trismegistos (TM) is an interdisciplinary portal of papyrological and epigraphical resources formerly focused on Egypt and the Nile valley (800 BC–800 CE)

59 See Iversen (2007) and, for a review of the online project, Gawlinski (2017). On other projects of the *Packard Humanities Institute* for Classical sources, see p. 43.

60 Blum (1991) 186 and 191.

61 See section 4.5 (*Digital Marmor Parium* project).

62 Sadurska (1964) 83–85.

63 See <https://inscriptions.packhum.org/text/4599>, <https://epigraphy.packhum.org/text/77668>, and <https://inscriptions.packhum.org/text/141279>.

Trismegistos Authors About

Hellanicus of Lesbos

TM Author id: 358 (Hellanicus of Lesbos)

more info: [Wikipedia](#), [Pinakes](#), [BNJ \(\\$\)](#), [BNJ \(\\$\)](#), [BNJ \(\\$\)](#), [BNJ \(\\$\)](#), [BNJ \(\\$\)](#), [FGRHist 1-3 \(\\$\)](#), [FGRHist 1-3 \(\\$\)](#), [FGRHist 1-3 \(\\$\)](#), [FGRHist 1-3 \(\\$\)](#), [FGRHist 1-3 \(\\$\)](#), [FGRHist 1-3 \(\\$\)](#), [FGRHist 1-3 \(\\$\)](#), [Perseus Catalog](#), [TLG Canon \(register\)](#)

lived AD 495 - 405 names: Hellanikos of Lesbos ethnic: of Lesbos; of Mytilene language: Greek genre: chronology, geography, history biblio: FGRHist 4 & 323a & 601a & 608a & 645a & 687a

4 attestations of works by this author in TM ?
 1 direct attestations ([filter](#)), 3 other ([filter](#))

Stable URI (with TM Author ID):
www.trismegistos.org/author/358

Figure 2.29. *Trismegistos Authors*: Hellanicus of Lesbos

and now expanding to the Ancient World in general.⁶⁴ *Trismegistos* offers many resources and gives the possibility to obtain metadata concerning also physical fragments that have been attributed to fragmentary authors and works. An interesting service is the *Authors* database for searching ancient author names and work titles. The aim of the resource is to collect information about all authors who wrote between 800 BC and 800 CE including also “authors attested only as fragments in other works.” As for now, the resource is based on the *Leuven Database of Ancient Books* (LDAB) and *Trismegistos* editors warn users about limits, errors, and possible lack of data.

Going back to the examples mentioned in the previous pages, *Trismegistos* has an entry about Hellanicus of Lesbos with a stable URI: <https://www.trismegistos.org/author/358>. The resource provides metadata (chronology, onomastics, ethnic, genre, language and bibliography) and links to different types of external resources such as *Wikipedia*, the manuscript collection of *Pinakes*, the *Perseus Catalog*, the *TLG Canon*, and *Jacoby Online* (fig. 2.29). As far as Hellanicus’ works are concerned, *Trismegistos* collects three papyri differentiating them between direct attestations and quotations (fig. 2.30): P.Oxy. VIII 1084 (*Atlantis*), P.Oxy. XXVI

64 For a detailed history and description of the project, see Reggiani (2017) 56–73, and Depauw (2018). *Trismegistos* was an open resource that, starting from January 1, 2020, requires a subscription to access all search interfaces and visualisations due to shortcuts in funding: 1) unlimited access through a subscribed institution (€ 990,91 excl. VAT per year), 2) institutional access for one concurrent user through a login (€ 299 per year excl. VAT), and 3) personal single user access through a login (€ 199 per year incl. VAT).

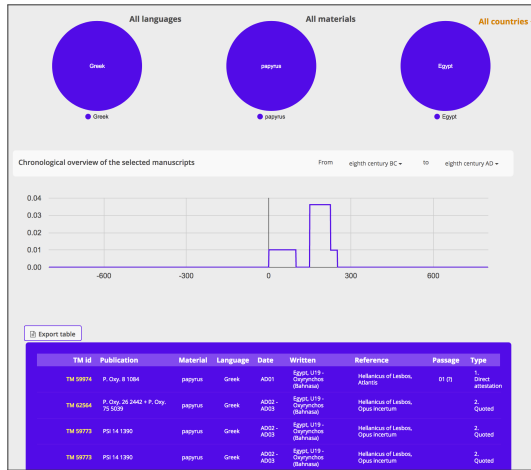


Figure 2.30. *Trismegistos Authors: Hellenicus of Lesbos (works)*

TM Texts
TM Home
About
Contact

stable url (with TM number):
www.trismegistos.org/text/59974

P. Oxy. 8 1084 (Hunt, Arthur S.)

= FGvHist 4 [Hellenikos] F 19 b (Jacoby, Felix)

Trismegistos nr: 59974

Publication: P. Oxy. 8 1084 (Hunt, Arthur S.; 1911)

Inventory: Princeton, University Library AM 4096

Other inventory nrs: formerly Princeton, University Library CC 0174.6.1084

Related inv. inf.:

Material: papyrus

Material form:

Reuse type:

Reuse detail:

Reuse note:

Language/script: Greek

Language detail:

Provenance: Egypt, U19 - Oxyrynchos (Babnasa) [found]
 Egypt, U19 - Oxyrynchos (Babnasa) (?) [written]

Provincia: Aegyptus

Date: AD 1 - 99 (cf. P. Oxy. 79 5199 introd.)

Seal:

Note:

Attested Authors: Hellenicus of Lesbos, Atlantis 01 (?) (1. Direct attestation)

Related resources: PN-APIS (Princeton) - one (TM) to one (APIS)

The above information is largely based on partner projects. Click the icon(s) for more information.

Figure 2.31. *Trismegistos: P.Oxy. VIII 1084*

2442 + P.Oxy. LXXV 5039 (*opus incertum*), and PSI XIV 1390 (*opus incertum*).⁶⁵ P.Oxy. VIII 1084 is a papyrus dated between the 1st and the 2nd century CE, whose text has been attributed to the *Atlantis* of Hellanicus of Lesbos (FGrHist (BNJ) 4 F 19b). *Trismegistos* offers a detailed description of the papyrus including the attribution to Hellanicus (*direct attestation*), a reference to the FGrHist, and a link to *Papyri.info* for other metadata and pictures (fig. 2.31).⁶⁶ P.Oxy. XXVI 2442 is constituted by several fragments of papyrus dated to the 2nd and 3rd centuries CE with fragments and *scholia* to Pindar that mention the name of Hellanicus (fr. 29, 1–8 = Mette (1978) 7, fr. 133bis = BNJ 4 F 101a). P.Oxy. LXXV 5039 is associated to the previous papyrus because it belongs to the same set of rolls, but doesn't contain the name of Hellanicus. *Trismegistos* has a page with metadata about both papyri and includes the reference to Hellanicus specifying that this is a reference to his name (*quoted*) and not one of his fragments (*direct attestation*), but in the bibliography doesn't refer to the BNJ.⁶⁷ PSI XIV 1390 is constituted by three fragments dated to the 2nd century CE and contains a *scholion* to Euphorion that mentions the name of Hellanicus (FGrHist 4 F 197bis = BNJ 4 F 197a). *Trismegistos* has a page on the papyrus with metadata, a reference to the fact that Hellanicus is quoted in the text but this is not one of his direct attestations, and a link to the database of the *Papiri della Società Italiana* with further information and pictures.⁶⁸

The TM *Authors* database allows also to search work titles. An example is the *Hellenica Oxyrhynchia*. In this case *Trismegistos* collects – as *direct attestations* under the heading *Anonymus of the Hellenica Oxyrhynchia, Hellenica Oxyrhynchia* – the fragments from the collections of Oxyrhynchos, Florence, and

65 It is not clear why PSI XIV 1390 is repeated twice, but *Trismegistos* editors warn about possible duplicates still present in the database. *Trismegistos* text types are related to the *Leuven Database of Ancient Books* (LDAB) metadata. As of now, there are four text types in *Trismegistos*: 1) *Direct attestation* (13,445 of the 15,101 attestations of authors in texts): this means that the text preserves the work of author X; 2) *Quoted* (882 attestations): this means that in the text a work of author X is quoted or referred to; 4) *Commented upon* (352 attestations): this means that a work of author X is the subject of a commentary; 5) *Epitomised* (422 attestations): this means that a work of author X is summarised. In the past there was also 3) *Translated*, but now there is a separate entry in works for each translation. I'm very grateful to Mark Depauw for this information about the current state of text types in *Trismegistos*.

66 See <https://www.trismegistos.org/text/59974> and <http://papyri.info/apis/princeton.apis.p2>
1. *Papyri.info* aggregates material and metadata from the *Advanced Papyrological Information System* (APIS), *The Duke Databank of Documentary Papyri* (DDbDP), the *Heidelberger Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens* (HGV), the *Bibliographie Papyrologique* (BP), *Trismegistos*, and also *The Arabic Papyrological Database* (APD). On the development of the project as part of *Integrating Digital Papyrology* and on its search and editing functionalities, see Reggiani (2017) 222 ff.

67 <https://www.trismegistos.org/text/62564>

68 <https://www.trismegistos.org/text/59773>

Cairo.⁶⁹ Given that *Trismegistos* collects data from other resources and the *Authors* database is a work in progress, there are papyri associated with the name of Hellanicus that are still missing or that are not yet imported in the author section, and data from *Jacoby Online* seems not to have been completely ingested.⁷⁰ In spite of that, *Trismegistos* is producing a fundamental resource for connecting and aggregating databases and metadata about the ancient world through stable identifiers and following recommendations and best practices of the *Linked Open Data* (LOD) initiative.⁷¹ Considering the huge amount of data to be collected and inserted into a complex database structure, the project is a model for establishing a collaborative environment and an integrated network of scholars on the ancient world.⁷² As far as fragmentary texts are concerned, this resource is very promising not only for publishing comprehensive digital data about physical fragments of literary texts, but also for aggregating catalog data about fragmentary authors and works (cf. sections 4.4.1 and 4.4.2).

I.Sicily (Inscriptiones Siciliae) is a project directed by Jonathan Prag at the University of Oxford for making freely available online the complete corpus of inscriptions from ancient Sicily in all languages from the 7th century BC through late antiquity.⁷³ Inscriptions are encoded in XML according to the TEI Epi-Doc schema and are stored in a database that can be filtered and searched in many different ways.⁷⁴ Each document has a unique identifier, as for example ISic000298.⁷⁵ Identifiers are cross-referenced with other collections like *Trismegistos* (TM), *PHI Greek Inscriptions*, and the *Epigraphic Database Roma* (EDR).

The goal of the project is to provide a new edition of every inscription with images, a commentary, and an up to date bibliography maintained in a separate public Zotero group library: <https://www.zotero.org/groups/382445>. An example for our interests in ancient Greek fragmentary historiography is represented

69 <https://www.trismegistos.org/authorwork/2177>

70 An example is P.Oxy. X 1241 that is part of *Trismegistos* and bibliographical metadata includes a reference to FGrHist 4 F 189, but the papyrus is not yet part of the *Authors* database: <https://www.trismegistos.org/text/63428>. Another resource connected to *Trismegistos* is the *Digital Corpus of Literary Papyri* (DCLP), which is building on tools and data of the *Integrating Digital Papyrology* project and *Papyri.info* to establish a database of literary papyri: <http://www.litpap.info>. *Trismegistos* numbers allow to obtain information concerning literary papyrological resources about fragmentary historians. For example, through TM 59974, it is possible to browse the DCLP and visualize the page on P.Oxy. VIII 1084, which is the papyrus with a fragment of the *Atlantis* of Hellanicus (see above in the text): <http://litpap.info/dclp/59974>.

71 Depauw/Gheldof (2014); Gheldof (2016); Reggiani (2017) 56 ff.; Depauw (2018). On LOD for data about the ancient world, see Elliott/Heath et al. (2014) and Cayless (2019).

72 Reggiani (2017) 56 ff.

73 Prag/Chartrand (2018).

74 Filters in the *I.Sicily* database include *id*, *date*, *place*, *material*, *object*, *inscription type*, *execution type*, *language*, *museum*, *status*, *other identifiers* (*Trismegistos*, EDR, and PHI).

75 <http://sicily.classics.ox.ac.uk/inscription/ISic000298>

by five painted fragments dated between the 3rd and the 2nd century BC from Tauromenium, which preserve traces of entries possibly belonging to a library catalogue. The text contains bio-bibliographic information about ancient authors who wrote in Greek: the fragmentary historians Callisthenes of Olynthus (BNJ 124), Philistos of Syracuse (BNJ 556), and Quintus Fabius Pictor (BNJ 809), an author from Elea whose name is lost and the fragmentary philosopher Anaximander.⁷⁶ The edition of the fragments in *I.Sicily* is still incomplete without images, a physical and epigraphic description, a critical apparatus and a commentary, but already includes bibliographic records, the current geo-location, and the date of the autopsy of the document.⁷⁷ The Greek text is based on the edition provided by PHI and is published in three versions: *interpreted*, *diplomatic*, and downloadable TEI EpiDoc XML. The text has a corresponding identifier in *Trismegistos* (TM 494031), whose entry includes bibliographic records but still misses further meta-data about the fragments and their linguistic content.⁷⁸

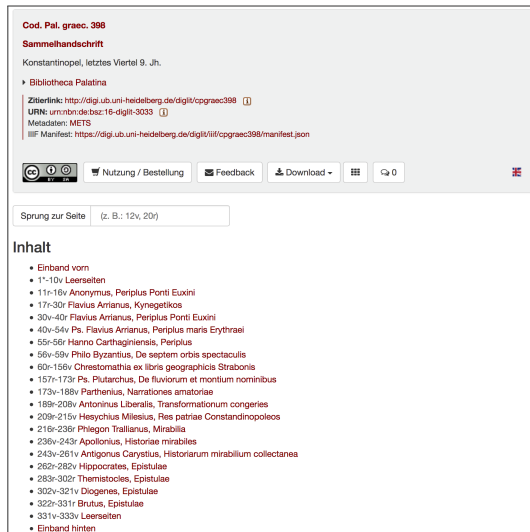


Figure 2.32. Bibliotheca Palatina digital: Codex Palatinus Graecus 398

The last resource is *Pinakes (Textes et manuscrits grecs)*, which is a French database for collecting catalog data about manuscripts of ancient Greek texts up to the end of the 16th century (excluding papyri). When the project was launched in 2008, the online collection counted 200,000 records concerning the manuscript tradi-

76 For recent and new readings of the fragments, see Battistoni (2006) and Matijašić (2018) 80–81.

77 <http://sicily.classics.ox.ac.uk/inscription/ISic000613>

78 <https://www.trismegistos.org/text/494031>

tion of 13,000 works from 40,000 manuscripts preserved in 1,300 libraries. The resource is a very good starting point for obtaining information about manuscripts and about authors and works preserved by them.

An example is *Codex Palatinus Graecus 398*, which is a manuscript of the 9th century from Constantinople that is part of the *Bibliotheca Palatina* of Heidelberg. The manuscript collects texts of sixteen authors including Phlegon of Tralles and Hesychius of Miletus, who are part of the *Fragmenta Historicorum Graecorum* and *Die Fragmente der griechischen Historiker*.⁷⁹ *Pinakes* has an entry about the manuscript with a detailed description and bibliography, and with the list of authors and texts preserved by it which are part of a general *Pinakes* catalog of ancient authors and works transmitted through manuscript tradition: <http://pinakes.irht.cnrs.fr/notices/cote/32479>. *Pinakes* offers also a link to the page of the *Bibliotheca Palatina digital* project with a complete description of the manuscript, high resolution images of each page, and a stable identifier expressed as a URN (<urn:nbn:de:bsz:16-diglit-3033>) (fig. 2.32).⁸⁰

2.2 Textual Fragments as Hypertexts

Editions of fragmentary texts are collections of excerpts from many different sources and are therefore representations of hypertexts.⁸¹ Figure 2.33 shows a lost text of Istros the Callimachean quoted by Athenaeus of Naucratis that has been extracted from the context of the *Deipnosophists* (on the right) and reproduced in a printed collection of fragments of Istros (on the left).⁸²

As discussed in the previous sections, this is a characteristic of the print culture that has been inherited by the first generation of digital libraries, which have been digitizing both source texts and collections of textual fragments derived from them (p. 55).

79 FHG III 602–624 = FGrHist (BNJ) 257; FHG IV 143–177 = FGrHist (BNJ) 390.

80 The permalink is <http://digi.ub.uni-heidelberg.de/diglit/cpgraec398>. The link includes other links for visualizing the pages of the manuscript with the works of authors preserved on it. An example is the *Mirabilia* of Phlegon of Tralles: <http://digi.ub.uni-heidelberg.de/diglit/cpgraec398/0435>. *Bibliotheca Palatina digital* includes also a *Creative Commons* licence (BY-SA 3.0 DE), an XML METS file with metadata of the manuscript, and a IIF Manifest JSON id.

81 On the definition of hypertext in computing and literary studies, see Landow (2006). On the impact of hypertext in Classical scholarship, see Crane (1987).

82 Berti (2009b) 99.

In this case, the model of the printed edition generates a static hypertext that in a digital enviroment can be converted into a hyperlink from the fragment to the passage of the *Deipnosophists* in order to help readers contextualize the reuse of the lost text of Istros.⁸³

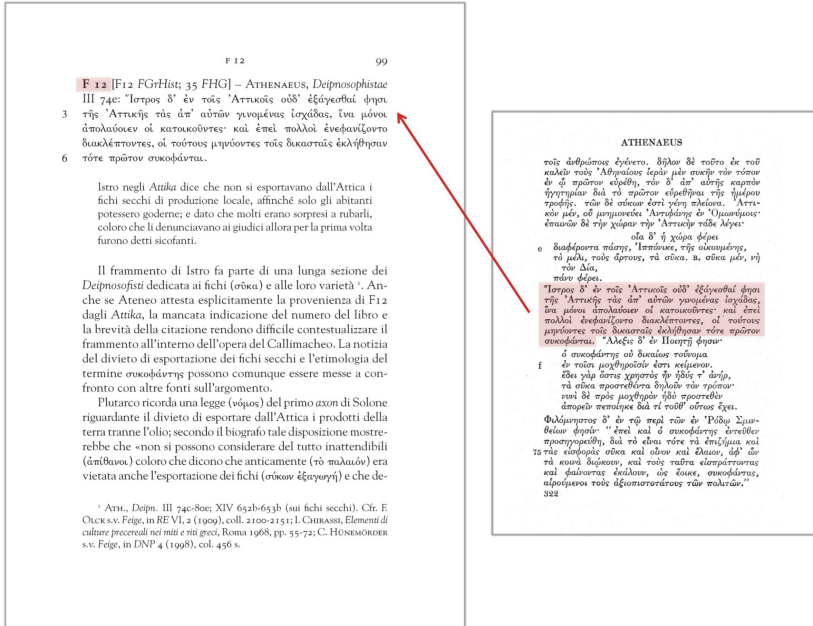


Figure 2.33. Excerpting fragments: Istros F 12 Bertl = *Deipn.* 3.74e

Printed collections of fragmentary texts contain many other hypertextual elements, as visible in figure 2.34.⁸⁴ The number of the fragment (F 1) corresponds to numbers of fragments in other collections (F 1 FG rHist; 1–2 FHG), where the same and other source passages have been excerpted, edited, commented and

83 On the relationship between context and text reuse, see section 2.3. Cf. also Landow (2006) 55: “Hypertext, which is a fundamentally intertextual system, has the capacity to emphasize intertextuality in a way that pagebound text in books cannot.”

84 Bertl (2009b) 43.

classified to reconstruct the lost text of Istros.⁸⁵ These correspondences are static hyperlinks to other editions that have to be consulted for analyzing different interpretations of text reuses of the same lost text.

F 1 [F1 FG^rHist; 1-2 FHG] – PHOTIUS [T 591] s.v. Τιτανίδα γῆν· οἱ μὲν τὴν πᾶσαν· οἱ δὲ τὴν Ἀττικὴν· ἀπὸ Τιτηνίου ἐνὸς τῶν Τιτάνων ἀρχαιοτέρου οἰκίσαντος περὶ Μαραθῶνα· ὅς μόνος οὐκ ἐστράτευσεν ἐπὶ τοὺς θεούς, ὡς Φιλόχορος ἐν Τετραπόλει. Ἰστρος δ' ἐν Ἀττικῶν * * Τιτάνας βοᾶν· ἐβοήθουν γὰρ τοῖς ἀνθρώποις ἐπακούοντες, ὡς Νίκανδρος ἐν Ἀἰτωλικῶν· ἐνομιζόντο δὲ τῶν Πριαπωδῶν θεῶν εἶναι.

Cfr. Suda [T 677] s.v. Τιτανίδα γῆν ([T 686] s.v. Τιτηνίδα γῆν) et Apostol. XVI 69 (s.v. Τιτανίδα παροικεῖς) 4 Φιλόχορος ἐν Τετραπόλει : FG^rHist 328 F74 6 Νίκανδρος ἐν Ἀἰτωλικῶν : FG^rHist 271-272 F4

1 Τιτανίδα : Τιτηνίδα Suda (A^oFV^o, cfr. [T 686]) 1-2 Τιτανίδα ~ πᾶσαν : Τιτανίδα παροικεῖς· ἐπὶ τῶν φιλοθέων Apostol. 2 πᾶσαν : πᾶσαν γῆν Apostol. Ἀττικὴν : Ἀττικὴν φασὶν Apostol. Τιτηνίου Suda, Apostol. : Τιτηνίου Phot., Τιτάνου (Τιτανίου V) Et. M. s.v. Τιτανίδα γῆν, πῶν κατασχόντων Hesych. [T 974] s.v. Τιτάνης γῆ, Τιτάκου Wilamowitz 3 ἀρχαιοτέρου : τὸ ἀρχαῖον vel ἀρχαιότερον vel [ἀρχ.] «ut huic irrepserit Τιτάνων ἀρχαιότεροι articulus explicacione carens ex Aristoph. Av. 469» Dobree περὶ : παρὰ Suda (FV) Μαραθῶνα : Μαραθῶρα Suda (A) 3-7 ὅς ~ εἶναι om. Suda (F) 5 Ἰστρος δ' ἐν : καὶ Ἰστρος ἐν Apostol. α' : πρώτη Apostol. * * Jacoby βοᾶν : βοᾶν Suda 5-7 Τιτάνας ~ εἶναι om. Apostol.

Figure 2.34. Istros F 1 Berti

In this example the *Lexicon* of Photius is the source text that quotes Istros the Callimachean (Photius [T 591] s.v. Τιτανίδα γῆν). This is a pure hypertext because the entry of the lexicographer, that has been extracted and reproduced in the collection of the fragments of Istros, points to the entire lexicon of Photius and its different editions. In figure 2.34, after the Greek text of the fragment, there is a section that collects *loci paralleli*, which are other sources that preserve a similar text reuse or discuss the same topic, and references to other lost authors who are mentioned by Photius in the same context where appears the quotation of Istros.⁸⁶ All these parallel sources and editions are hypertextual elements de-

85 FG^rHist 334 F 1 and Istros F 1 Berti publish the entry of Photius' *Lexicon* ([T 591] s.v. Τιτανίδα γῆν) as the main source of the fragment of Istros (in Berti the entry is complete, while Jacoby prints only the first part of it). Müller in FHG I 418, fr. 1 publishes the same entry of the *Suda* (s.v. Τιτανίδα γῆν) as the source text of the lost fragment of Istros and cites Photius' entry (s.v. Τιτανίδα γῆν) in the commentary to the fragment. Müller also publishes a passage of the *Collectio paroemiarum* of Apostolius (XVIII 77) as the source text of fragment 2 of the *Attika* of Istros, while Jacoby and Berti cite Apostolius (XVI 69) as a *locus parallelus* of fragment 1. The two different citations of Apostolius depend on the use of different editions (Jacoby and Berti used the edition by Ernst Ludwig von Leutsch, while Müller used the edition by Daniel Heinsius).

86 *Suda* [T 677] s.v. Τιτανίδα γῆν (cf. [T 686] s.v. Τιτηνίδα γῆν); Apostol. XVI 69 (s.v. Τιτανίδα παροικεῖς); Philoch., FG^rHist (BNJ) 328 F 74; Nicander, FG^rHist (BNJ) 271-272 F 4. The TLG is adding some of these hyperlinks in its collection. For example, in the TLG entry of Photius (t1g4040) there is a link to the fragment of Philochorus as published in the FG^rHist.

rived from the analysis of the fragment of Istros. The last section of figure 2.34 is the *apparatus criticus*, that contains a critical summary of the historical tradition of the lexical entry of Photius and that generates another group of possible hyperlinks to sources, manuscripts, and philological conjectures.⁸⁷ Beyond these elements that pertain to a single fragment, a printed edition of fragmentary texts includes other hypertexts and potential hyperlinks in the commentaries, in the footnotes, and in other sections at the end of the volume. Figure 2.35 is a screenshot from the project *demo.fragmentarytexts.org* that summarizes these elements and describes them in separate web pages: 1) editing and commenting text reuse, 2) concordance tables, 3a) indexes of sources, 3b) indexes of names, and 4) bibliography.⁸⁸

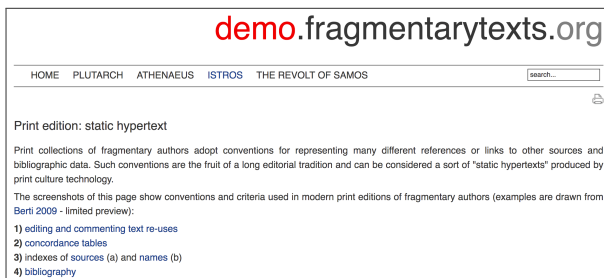


Figure 2.35. Istros: print edition (static hypertext)

According to the definitions presented in chapter 1 and if I exclude physical fragments of ancient texts, textual fragments can be described as quotations and reuses of other texts that generate a complex multisequential and non-linear network of hypertexts. As we have seen before, the first natural hypertext is between the extracted fragment (e.g., Istros F 12 Berti) and its source text (Athen., *Deipn.* 3.6 = 74e). Other kinds of hypertexts are produced by parallel sources (*loci paralleli*). I have mentioned the example of Istros F 1 Berti and I can also analyze Istros F 4 Berti.⁸⁹ In this case the main source who quotes Istros the Callimachean is

87 For example: hyperlinks to the texts of the *Etymologicum Magnum* (s.v. Τιτανίδα γῆν) and of the *Lexicon* of Hesychius ([T 974] s.v. Τιτανις γῆ), to different readings in different manuscripts of the *Suda*, and to conjectures by Ulrich von Wilamowitz-Moellendorff and Peter Paul Dobree.

88 See <http://demo.fragmentarytexts.org/en/istros.html> from which are also taken the screenshots reproduced in the following pages. Examples are from Berti (2009b).

89 Berti (2009b) 59–64.

Harpocration in his *Lexicon of the Ten Attic Orators* (s.v. Παναθήναια).⁹⁰ The text of Harpocration includes references to still extant sources, who are Demosthenes and Isocrates, and to three lost authors, who are Hellanicus, Androtion, and Istros. On the other hand, the tradition shows that the text of Harpocration was reused by the author of the *Suda* ([Π 152] s.v. Παναθήναια) and probably also by Photius in his *Lexicon* ([Π 376] s.v. Παναθήναια) (fig. 2.36).

The text of Harpocration has been reproduced in an abridged form by Photius in the <i>Lexicon</i> (Π 376) and by the <i>Suda</i> (Π 152):		
Harpocration, s.v. Παναθήναια	Photius, <i>Lexicon</i> (Π 376)	<i>Suda</i> (Π 152)
Δημοσθένης Φιλίππικός, διττά Παναθήναια ἤγετο Ἀθήνησι, τὰ μὲν καθ' ἕκαστον ἑνιαυτὸν, τὰ δὲ διὰ πεντετηρίδος, ἄνω καὶ μεγάλα ἐκάλουν. Ἰσοκράτης Παναθηναϊκῶ φησι μικρὸν δὲ πρὸ τῶν μεγάλων Παναθηναίων ἤγαγε δὲ τὴν ἑορτὴν πρῶτος Ἐριχθόνιος ὁ Ἡφαίστου, καθὰ φησὶν Ἑλληνικός τε καὶ Ἀνδροτίων, ἐκάτερος ἐν Ἀθηίδος, πρὸ τούτου δὲ Ἀθήναια ἐκαλεῖτο, ὡς δεδήλωκεν Ἰστρος ἐν γ τῶν Ἀττικῶν.	Παναθήναια: ἀγὼν πεντετηρικός Ἀθήνησιν.	Παναθήναια: διττά Παναθήναια ἤγετο Ἀθήνησι, τὰ μὲν καθ' ἕκαστον ἑνιαυτὸν, τὰ δὲ διὰ πεντετηρίδος, ἀ καὶ μεγάλα ἐκάλουν. ἤγαγε δὲ τὴν ἑορτὴν πρῶτος Ἐριχθόνιος ὁ Ἡφαίστου. τὰ δὲ Παναθήναια πρότερον Ἀθήναια ἐκαλοῦντο.

Figure 2.36. Istros F 4 Berti: source alignment

<p>F 22a [F22 FG^rHist; 21 FHG] – Schol. in Sophoclis Oedipum Coloneisim 1053: προσέειπεν Εὐμόλπουδ' ἱερέειαι τί δήποτε οἱ Εὐμόλπειδαι τῶν τελεπῶν ἐξάρχουσι, ξένου ὄντες; εἶποι δ' ἂν τις ὅτι ἀξιοῦσιν ἔνοι πρῶτον Εὐμόλπον μῆσαι τὸν Δημόστῃ τῆς Τριπολιέμου τὰ ἐν Ἐλευσίῃ μυστήρια καὶ οὐ τὸν Θράκα καὶ τοῦτο ἱστορεῖν Ἰστρον ἐν τῷ ἱ περὶ τῶν Ἀττικῶν, Ἀκαστόδορος δὲ πέμπτον ἀπὸ τοῦ πρώτου Εὐμόλπου εἶναι τὸν τὰς τελεπῶν καταδείξαντα γράφει οὕτως: « κατοικήσει δὲ τὴν Ἐλευσίαν ἱστοροῦσι πρῶτον μὲν τοὺς αὐτόχθονας, εἰτα Θράκας τοὺς μετὰ Εὐμόλπου παραγενομένους πρὸς βοήθειαν εἰς τὸν κατ' Ἐρεχθίδος πόλεμον. τινὲς δὲ φασὶ καὶ τὸν Εὐμόλπον εὐρεῖν τὴν μῆσιν τὴν συντελουμένην κατ' ἑνιαυτὸν ἐν Ἐλευσίῃ. Δῆμητρι καὶ Κόρη », Ἄνδρον μὲν οὖν γράφει οὐ τὸν πρῶτον Εὐμόλπον εὐρεῖν « τὴν μῆσιν, ἀλλ' ἀπὸ τούτου Εὐμόλπον πέμπτον γεγόνота: Εὐμόλπου γὰρ γενέσθαι Κίρκια, τοῦ δὲ Εὐμόλπον, τοῦ δὲ Ἀντίφθμον, τοῦ δὲ Μουσαίου τὸν ποιητῆν, τοῦ δὲ Εὐμόλπον τὸν καταδείξαντα τὴν μῆσιν καὶ πρῶτον ἱεροφάντην γεγόνота.</p> <p>6 Ἀκαστόδορος : FHG II, p. 464 13 Ἄνδρον : FG^rHist 10 F13</p>	<p>F 22b [20 FHG] – Schol. in Lycophronis Alexandram 1328: Εὐμόλπος γὰρ οἶχ ὁ Θράξ κατὰ Ἰστρον, ἀλλ' ὁ θεὸς τὰ μυστήρια ἐκέλευσε ξένους μὴ *μικεῖσθαι ἴ. ἄλλήτως δὲ τοῦ Ἡρακλέους ἐν Ἐλευσίῃ καὶ θέλοντος* μικεῖσθαι τὸν μὲν τοῦ Εὐμόλπου ἴ νόμον φυλάττοντες, θέλοντες δὲ καὶ τὸν κοινὸν ἐφεγγέτην Ἡρακλέα θεραπεύουσι οἱ Ἐλευσίονα ἐπ' αὐτῷ τὰ μικρὰ ἐποιήσαντο μυστήρια. οἱ δὲ μοῦσεμον μωροῖν ἐπέτρεφοντο.</p>
--	--

Figure 2.37. Istros F 22 Berti a and b

90 “Panathenaia: Demosthenes in the *Philippics* (4.35). The Panathenaia held at Athens was two-fold, one festival being held annually, and the other celebrated every five years, the latter also called the Great Panathenaia. Isocrates in the *Panathenaichus* (12.17) says ‘a short time before the Great Panathenaia.’ The first to conduct the festival was Erichthonius, son of Hephaistos, according to the reports of Hellanicus (FHG I 54, fr. 65 = FG^rHist (BNJ) 4 F 39 = FG^rHist (BNJ) 323a F 2 = Ambaglio 1980 F 162) and Adrotion (FHG I 371, fr. 1 = FG^rHist (BNJ) 324 F 2 = Harding 1994 F 2), both in the first book of the *Atthis*. Before his time the festival was called the Athenaia, as Istros makes clear in the third book of his *Attika* (FHG I 419, fr. 7 = FG^rHist (BNJ) 334 F 4 = Berti 2009 F 4).” See pp. 11 and 52 ff. for the Greek text of this entry, a commentary, and its treatment in the TLG.

Another interesting example is Istros F 22 Berti.⁹¹ In this case we have two different reuses of the same lost text of Istros preserved by two sources: the *scholion* to Sophocles' *Oedipus at Colonus* 1053 (Berti F 22a) and the *scholion* to Lycophron's *Alexandra* 1328 (Berti F 22b). Editors have classified the two sources in different ways. I consider them as two parallel sources of the same fragment of the *Atakta* of Istros, given that their texts are different but they both explicitly mention Istros about the same topic (fig. 2.37).⁹² Jacoby prints only the text of the *scholion* to Sophocles and adds the reference to the *scholion* to Lycophron in parentheses as a parallel text (FGrHist 334 F 22). Müller publishes the two sources as two different fragments of Istros' *Attika* (FHG I 421, fr. 20–21).

<p>Athenaei <i>Naucraltitae Dipnosophistarum Libri XV</i>, rec. G. Kaibel. Vol. I. Lipsiae 1887</p>	<p>Thucydides. <i>History of the Peloponnesian War II</i> (Books III-IV), ed. C.F. Smith. Cambridge, Ma 1958</p>
<p>Ath. <i>Deipn.</i> 5.15 (189c) ἔτι δὲ αὐλός μὲν τὸ ὄργανον, ὅτι διέρχεται τὸ πνεῦμα, καὶ πᾶν τὸ διατεταμένον εἰς εὐθύτητα σχῆμα αὐλὸν καλοῦμεν ὡπερ τὸ στάδιον καὶ τὸν κρουόν τού αἵματος αὐτίκα δ' αὐλὸς ἀνά ρίνας παχύς ῥῆθε, καὶ τὴν περικεφαλαίαν ὄταν ἐκ τοῦ μέσου πρὸς ὀρθὸν ἀνατείνῃ αὐλῶπιν. λέγονται δὲ Ἀθήνησι καὶ ἱεροὶ τινες αὐλῶνες, ὧν μὲν ἡται Φιλόχορος ἐν τῇ ἐνάτῃ (FHG I 409 fr. 147 = FGrH 328 F 68), καλοῦσι δ' ὀρεινικῶς τοὺς αὐλῶνας, ὡπερ Θεουκιδίης ἐν τῇ δ' (4.103.1) καὶ πάντες οἱ καταλογάδην συγγραφεῖς, οἱ δὲ ποιηταὶ θηλυκῶς.</p>	<p>Thuc. 4.103 (1) Ἐπὶ ταύτην οὖν ὁ Βρασιδᾶς ἄρας ἐξ Ἀρνῶν τῆς Χαλκιδικῆς ἐπορεύετο τῷ στρατῷ. καὶ ἀφικόμενος περὶ δεῖλην ἐπὶ τὸν Αὐλῶνα καὶ Βορμίσκον, ἧ ἡ Βόλβη λίμνη ἐξήσταν ἐς θάλασσαν, καὶ δειπνοποισάμενος ἐχώρει τὴν νύκτα. (...)</p>

Figure 2.38. Athen., *Deipn.* 5.189c = Thuc. 4.103.1

The last case I can mention is when an extant source text quotes or alludes to another extant source text, as for example Athenaeus (*Deipn.* 5.189c) who quotes Thucydides (4.103.1). This is a clear example of a hypertext that can generate word alignments to see the differences between the two texts and how reliable is Athenaeus in his reference (fig. 2.38).⁹³

Fragmentary literature has the power to generate a huge amount of possible hypertexts beyond the examples presented in these pages. For our purposes, I can classify them into two main groups: 1) hypertexts produced by extant sources that preserve quotations and reuses of other texts, and 2) hypertexts produced by critical editions of fragmentary texts that point to other sources, editions, commentaries and reference tools. Now that source editions from which fragments are extracted are becoming available in a digital form, it is possible to create a new model of editions that are truly hypertextual and that include not only excerpts but also links to scholarly sources from which those excerpts are drawn. Building a digital corpus of fragmentary authors means addressing the problem of encoding and representing both the text and the structure of a fragment.⁹⁴ It is widely

91 Berti (2009b) 142–151.

92 See also BNJ 334 F 22ab.

93 See p. 10.

94 The following pages collect reflections published in Berti (2015b).

accepted that a digital representation of the internal and external characteristics of a text consists not simply of a mere reproductive and mechanical process, but of an interpretative act.⁹⁵ Accordingly, encoding fragments is first of all the result of interpreting them, developing a language appropriate for representing every element of their textual features, thus creating meta-information through an accurate and elaborate semantic markup. Editing fragments, therefore, signifies producing meta-editions that are different from printed ones because they consist not only of isolated quotations but also of pointers to the original contexts from which the fragments have been extracted. While editors should be able to define the precise chunks of text that they feel relevant and annotate these texts in various ways (e.g., distinguishing what they consider to be paraphrase from direct quotation), such fragments should also be dynamically linked to their original contexts and to up-to-date contextualizing information.

On a broader level, the goal of a digital edition of fragments is to represent multiple transtextual relationships as they are defined in literary criticism, which include *intertextuality* (the presence of a text inside another text, such as quotations, allusions, and plagiarism), *paratextuality* (i.e., all those elements which are not part of the text, like titles, subtitles, prefaces, notes, etc.), *metatextuality* (critical relations among texts, such as commentaries and critical texts), *architextuality* (the entire set of categories from which emerges each text), and *hypertextuality* (i.e., the derivation of a text from a preexisting hypotext through a process of transformation or imitation).⁹⁶ Designing a digital edition of fragmentary texts also means finding digital paradigms and solutions to express information about printed critical editions and their editorial and conventional features. Working on a digital edition means converting traditional tools and resources used by scholars such as canonical references, tables of concordances, and indices into machine actionable contents (cf. chapters 4 and 5).

In order to show some of the complex transtextual relations produced by quotations and text reuses, I consider an example constituted by a series of fragmentary references embedded in a long section of the *Life of Theseus* by Plutarch, which pertains to the unification of Attica and the beginning of democracy, the annexation of the territory of Megara to Attica, the institution of the Isthmian games, and the war against the Amazons.⁹⁷

95 Fiormonte (2003) 163–172; Apollon/Bélisle et al. (2014); Pierazzo (2015).

96 Genette (1982), part. 7–17; Landow (2006). On these categories applied to the domain of fragmentary literature, see Berti (2012) and Berti (2013a) with bibliography.

97 Plut., *Thes.* 24–28. Citation references are based on the edition of Perrin (1914) 50–66. For a visualization of these chapters with alignments of the Greek and the English texts and with annotations of text reuses, see <http://demo.fragmentarytexts.org/en/plutarch.html>.

In these chapters Plutarch mentions many different sources: 1) three oracles;⁹⁸ 2) the text of an inscription;⁹⁹ 3) surviving authors, such as Aristotle, Homer, Plutarch himself, and Pindar;¹⁰⁰ 4) a series of fragmentary historians, such as Hellanicus, Andron of Halicarnassus, Philochorus, Pherecydes, Herodorus, Bion, Menecrates, Clidemus, and the author of the *Theseid*.¹⁰¹ Beside these sources, Plutarch adds also generic references to other unnamed authors as witnesses of his account.¹⁰²

<p>26 (1) Εἰς δὲ τὸν πόντον ἐπλευσε τὸν Εὐξείνου, ὡς μὲν Φιλόχορος καὶ τινες ἄλλοι λέγουσι, μεθ' Ἡρακλέους ἐπὶ τὰς Ἀμαζόνας αὐστρατεύσας, καὶ γέρας Ἀντιόπην ἔλαβεν· οἱ δὲ πλείους, ὧν ἐστὶ καὶ Φερεκύδης καὶ Ἑλλάνικος καὶ Ἡρόδοτος, ὕστερόν φασιν Ἡρακλέους ἰδιοστολον πλεῦσαι τὸν Θησαῖα καὶ τὴν Ἀμαζόνα λαβεῖν αἰχμάλωτον, πιθανώτερα λέγοντες· οὐδεὶς γὰρ ἄλλος ἰστόρηται τῶν μετ' αὐτοῦ στρατεύσαντων Ἀμαζόνα λαβεῖν αἰχμάλωτον. (2) Βίων (FHG II 19 fr. 1) = FGrH 14 F 2 = FGrH 332 F 2) δὲ καὶ ταύτην παρακρούσασμενον οἴχεσθαι λαβόντα· φύσει γὰρ οὐσας τὰς Ἀμαζόνας φιλόφρονος οὔτε φυγεῖν τὸν Θησαῖα προσβάλλοντα τῇ χώρῃ, ἀλλὰ καὶ ξένια πέμψειν· τὸν δὲ τὴν κομίζουσαν ἐμβήναι παρακαλεῖν εἰς τὸ πλοῖον· ἐμβάσης δὲ ἀναχθῆναι. Μενεκράτης δὲ τις, ἱστορίαν περὶ Νικαίας τῆς ἐν Βιθυνίᾳ πόλεως ἐκδεδικώς, Θησαῖα φησὶ τὴν Ἀντιόπην ἔχοντα διατρίψαι περὶ τούτους τοὺς τόπους. (3) τυγχάνειν δὲ</p>	<p>26 (1) He also made a voyage into the Euxine Sea, as Philochorus and sundry others say, on a campaign with Heracles against the Amazons, and received Antiope as a reward of his valour; but the majority of writers, including Pherecydes, Hellanicus, and Herodorus, say that Theseus made this voyage on his own account, after the time of Heracles, and took the Amazon captive; and this is the more probable story. For it is not recorded that any one else among those who shared his expedition took an Amazon captive. (2) And Bion says that even this Amazon he took and carried off by means of a stratagem. The Amazons, he says, were naturally friendly to men, and did not fly from Theseus when he touched upon their coasts, but actually sent him presents, and he invited the one who brought them to come on board his ship; she came on board, and he put out to sea. And a certain Menecrates, who published a history of the Bithynian city of Nicaea, says that Theseus, with Antiope on board</p>
--	--

Figure 2.39. Bion, FHG II 19, fr. 1

The text of Plutarch has been split by Karl Müller and Felix Jacoby into extracts scattered and repeated in the sections of their collections of Greek historical fragments corresponding to the authors mentioned by the biographer.¹⁰³ Accord-

98 Two oracles from Delphi (*Thes.* 24.5 = Parke-Wormell II 154; *Thes.* 26.4 = Parke-Wormell II 411); one oracle of the Sibyl (*Thes.* 24.5 = Hendess 23).

99 The pillar on the Isthmus (*Thes.* 25.3). At 27.2 and 27.4, without quoting the text, Plutarch mentions also the graves of those who fell in battle and the pillar by the sanctuary of Olympian Earth.

100 Aristotle (*Thes.* 25.2 = *Ath. Pol.* 41.2; FHG II 105, fr. 2 = F 384 Rose³); Homer (*Thes.* 25.2 = *Iliad* 2.547); Plutarch himself (*Thes.* 27.6 = *Dem.* 19.2); Pindar (*Thes.* 28.2 = F 176 Sn.-Mae).

101 Hellanicus (*Thes.* 25.5 = FHG I 55, fr. 76 = FGrHist (BNJ) 4 F 165 = FGrHist (BNJ) 323a F 15; *Thes.* 26.1 = FHG I 55, fr. 76 = FGrHist (BNJ) 4 F 166 = FGrHist (BNJ) 323a F 16a; *Thes.* 27.2 = FGrHist (BNJ) 4 F 167a = FGrHist (BNJ) 323a F 17a); Andron (*Thes.* 25.5 = FHG II 351, fr. 13 = FGrHist (BNJ) 10 F 6); Philochorus (*Thes.* 26.1 = FHG I 392, fr. 49 = FGrHist (BNJ) 328 F 110); Pherecydes (*Thes.* 26.1 = FGrHist (BNJ) 3 F 151); Herodorus (*Thes.* 26.1 = FHG II 32, fr. 16 = FGrHist (BNJ) 31 F 25a); Bion (*Thes.* 26.2 = FHG II 19, fr. 1 = FGrHist (BNJ) 14 F 2 = FGrHist (BNJ) 332 F 2); Menecrates (*Thes.* 26.2 = FHG II 345, fr. 8 = FGrHist (BNJ) 701 F 1); Clidemus (*Thes.* 27.3 = FHG I 360, fr. 6 = FGrHist (BNJ) 323 F 18); the author of the *Theseid* (*Thes.* 28.1 = EGF 217 Kinkel).

102 *Thes.* 25.1 (φασί); 25.3 (φασί); 25.4 (ἔνιοι δὲ φασιν); 26.1 (καὶ τινες ἄλλοι λέγουσι [...] οἱ δὲ πλείους [...] οὐδεὶς γὰρ ἄλλος ἰστόρηται); 27.2 (μαρτυρεῖται καὶ τοῖς ὀνόμασι τῶν τόπων καὶ ταῖς θήκαις τῶν πεσόντων); 27.4 (ἔνιοι δὲ φασιν); 27.5 (μαρτύριόν ἐστιν); 27.6 (λέγεται δὲ καὶ [...] φαίνονται δέ); 28.2 (παρὰ τῶν ἱστορικῶν τοῖς τραγικοῖς). On “unnamed and named quotations” in ancient sources, see Berti (2012) 456–458, and Berti (2013a) 275–276.

103 See n. 101.


<p>26 (1) Εἰς δὲ τὸν πόντον ἔπλευσε τὸν Εὐξείνιον, ὡς μὲν Φιλόχορος καὶ τινες ἄλλοι λέγουσι, μεθ' Ἡρακλέους ἐπὶ τὰς Ἀμαζόνας συστρατεύσας, καὶ γέρας Ἀντιόπην ἔλαβεν· οἱ δὲ πλείους, ὧν ἐστὶ καὶ Φερεκύδης καὶ Ἑλλάνικος καὶ Ἡρόδοτος, ὑστερὸν φασὶν Ἡρακλέους ἰδιοστολὸν πλεῦσαι τὸν Θησέα καὶ τὴν Ἀμαζόνα λαβεῖν αἰχμάλωτον, πιθανώτερα λέγοντες· οὐδεὶς γὰρ ἄλλος ἰστόρηται τῶν μετ' αὐτοῦ στρατευσάντων Ἀμαζόνα λαβεῖν αἰχμάλωτον. (2) Βίων (FHG II 19 fr. 1  = FGrH 14 F 2 = FGrH 332 F 2) δὲ καὶ ταύτην παρακρουσάμενον οἴχεσθαι λαβόντα· φύσει γὰρ οὐσας τὰς Ἀμαζόνας φιλάνθρωπος οὔτε φυγεῖν τὸν Θησέα προσβάλλοντα τῇ χώρᾳ, ἀλλὰ καὶ ξένια πέμπειν· τὸν δὲ τὴν κομιζούσαν ἐμβῆναι παρακαλεῖν εἰς τὸ πλοῖον· ἐμβάσης δὲ ἀναχθῆναι. Μενεκράτης, δὲ τις, ἱστορίαν περὶ Νικαίας τῆς ἐν Βιθυνίᾳ πόλεως ἐκδεδικώς, Θησέα φησὶ τὴν Ἀντιόπην ἔχοντα διατρίβειν περὶ τούτους τοὺς τόπους. (3) τυγχάνειν δὲ</p>	<p>26 (1) He also made a voyage into the Euxine Sea, as Philochorus and sundry others say, on a campaign with Heracles against the Amazons, and received Antiope as a reward of his valour; but the majority of writers, including Pherecydes, Hellanicus, and Herodorus, say that Theseus made this voyage on his own account, after the time of Heracles, and took the Amazon captive; and this is the more probable story. For it is not recorded that any one else among those who shared his expedition took an Amazon captive. (2) And Bion says that even this Amazon he took and carried off by means of a stratagem. The Amazons, he says, were naturally friendly to men, and did not fly from Theseus when he touched upon their coasts, but actually sent him presents, and he invited the one who brought them to come on board his ship; she came on board, and he put out to sea. And a certain Menecrates, who published a history of the Bythinian city of Nicaea, says that Theseus, with Antiope on board</p>
---	--

Figure 2.40. Bion, FGrHist 14 F 2 = 332 F 2

ingly, the result of the printed representation of these fragments is that the same text of the *Life of Theseus* is not only broken off in many excerpts, but also repeated as many times as are the authors quoted in it.¹⁰⁴ Moreover, given that it is not possible to clearly identify the boundaries of the quotations preserved by Plutarch, editors have adopted different criteria for extracting them, and the same fragment may have different lengths and divisions from one edition to another.¹⁰⁵ Digital technologies allow scholars to go beyond these limits because standards, protocols, and tools now available permit to generate a model that can express the hypertextual and hermeneutical nature of fragmentary texts, providing an interconnected corpus of primary and secondary sources of fragments that also includes critical apparatuses, commentaries, translations, and modern bibliography on ancient texts. The first requirement for building a digital collection of fragmentary texts is to make the semantic contents of printed critical editions machine readable, defining a general architecture for representing at least the following main hypertextual elements that pertain to the domain of historical fragmentary texts.¹⁰⁶

1) *Quotation as machine actionable link*. The passage of the *Life of Theseus* should be linked to the whole context of still extant sources and to editions of lost

104 On this problem for digital libraries, see p. 57.

105 Fig. 2.39 shows in blue the portion of text extracted by Müller and printed in the FHG. Fig. 2.40 shows in red a different portion of text for the same fragment extracted by Jacoby and printed in the FGrHist. Different cut, copy, and paste methods used for the same fragment in different editions are noticeable in the case of Philochorus (FHG I 392, fr. 49 = FGrHist 328 F 110) and Clidemus (FHG I 360, fr. 6 = FGrHist 323 F 18). There is also an example where the same fragment of Hellanicus has two different lengths within the same collection: FGrHist 4 F 167a and FGrHist 323a F 17a. Finally Hellanicus, FHG I 55, fr. 76 partially corresponds to four different fragments in Jacoby (FGrHist 4 F 165 = FGrHist 323a F 15 and FGrHist 4 F 166 = FGrHist 323a F 16a). For a digital and dynamic visualization of these differences, see <http://demo.fragmentarytexts.org/en/plutarch.html>.

106 Berti/Romanello et al. (2009); Romanello/Boschetti et al. (2009); Romanello (2011).

authors cited by Plutarch.¹⁰⁷ On the other hand, editions of fragments should be linked to the whole text of the *Life of Theseus*. This is the first function for a proper representation of fragmentary texts to see each fragment directly within its context of transmission and avoid the misleading idea of an independent material existence of fragmentary texts, which derives from typographical representation of excerpts that are actually the result of modern reconstructions of lost works.¹⁰⁸

2) *Start and end of a fragment*. The next step is to provide a mechanism for marking the beginning and the end of a fragment in its context according to the choices of different editors. The result is that a scholar, while reading the excerpt inside its source of transmission, is able to visualize simultaneously how different editors have extracted different portions of text from the same context in order to generate a fragment. As we will see in chapter 3, the ultimate goal of a new born-digital edition of fragmentary texts is to go beyond the problem of defining borders of text reuses and to produce what we could call *borderless fragments*. This expression means that the result is not a chunk of text extracted from the context, but a collection of annotations that mark up different elements pertaining to a text reuse within its context of transmission.¹⁰⁹ This function has another important advantage in a digital library because it eliminates the problem of the repetition of the same text inside a collection, as it happens for example in the TLG (see pp. 57 ff.)

3) *Numbering and ordering fragmentary authors and fragments*. Numbering and ordering fragmentary authors and their fragments may vary in a significant way from one edition to another. These differences depend on the choices of the editor, who can decide to date and classify authors and order fragments according to different internal and external characteristics of the fragments themselves and of their sources.¹¹⁰ Differences may also be the result of different fragmentations of the same text or of the need to add new authors and texts to a collection of fragments. My model provides the possibility of encoding this kind of information, which is usually registered in the table of concordances of a printed edition. Aligning multiple references to the same textual object can help readers visualize different numberings and orderings of fragments in different editions, and the model also permits to include new data if new editions are added.¹¹¹

107 E.g., Homer and Aristotle, and Hellanicus and Philochorus in the FHG and the FGrHist.

108 On the role of the context, see section 2.3.

109 I have coined the expression *borderless fragment* from the concept of “borderless electronic text” described by Landow (2006) 110–118.

110 In the FHG, Greek fragmentary historians are arranged chronologically, while in the FGrHist they have a sequential number and are organized by genres. Fragments are grouped by works inside both collections: see section 1.4.

111 See section 3.2 on the use of the *CITE Architecture* for this purpose. See also section 4.4.3 for the digital table of concordance between the FHG and the FGrHist.

4) *Representing information about fragmentary authors and works.* The sources that transmit fragments may include many elements that reveal the presence of the textual reuse, such as the name of a fragmentary author, the title or the description of a fragmentary work, and other references to a fragmentary work passage as for example the book number. Attributing a fragment to an author and a work can be a difficult task, because there are homonymous authors and also because managing titles of ancient works can be quite challenging.¹¹² Witnesses do not always cite work titles, and in ancient times titles were not fixed and definitive as nowadays because they could be referred to with variants and in the form of a description of the work content. The result is that different editors may attribute the same fragment to different authors and works.¹¹³ The goal is to develop a comprehensive catalog with unique identifiers for every fragmentary author and work that will include multiple expressions of the same author and work and where each entry will have associated metadata, providing scholars with a sort of canon that simultaneously includes all available information on fragmentary authors and works, with pointers to primary and secondary sources (cf. section 5.6). This function can help enhance one of the “theoretical questions” suggested by Glenn Most when collecting fragments, which is the relationship between fragmentary authors and the “shifting boundaries of canon formation over time.”¹¹⁴

5) *Classifying fragments.* Fragmentary authors and works are classifiable according to multiple criteria that range from internal to external factors. The first classification is based on literary genres and subgenres that have led scholars of printed editions to generate very complex categories for arranging authors and texts within their collections (cf. sections 1.2 and 1.4). Another traditional way of classifying fragments is distinguishing them between *testimonia* (i.e., fragments providing biographical and bibliographical information about fragmentary authors) and *fragmenta* (i.e., actual text reuses of lost works). The printed representation of these categories has many limitations because it is impossible to draw a demarcation line among many different genres of fragmentary authors and works that can be inserted in different overlapping categories. The result again is that the same fragment is often repeated in many different sections corresponding to different categories.¹¹⁵ A digital collection in which every fragment

112 See the example of Crates of Athens and Crates of Mallus, who are both considered possible authors of a work on Attic glosses attributed by ancient sources to a not further specified Crates: Broggiato (2000). On titles in Greek literature see Castelli (2020).

113 See Harding (2008) 1 on the different ways in which ancient authors refer to the works of the Athidographers. See Berti (2009b) 6–8 on the different forms of the title of the work on Athens of Istros the Callimachean.

114 Most (1997) vi. On the Alexandrian canon and the “canons” of ancient Greek historiography, see Nicolai (2013) and Matijašić (2018).

115 Cf. Berti (2013a) 271–272.

is preserved in its original context and represented with multiple pieces of meta-data can express the complexity of modern classifications, while not scattering and repeating the same excerpt many different times. In this way, it is possible to avoid the strictness of printed categories, allowing scholars to compare a fragment with many other excerpts and visualizing its belonging to different categories in a more dynamic and simultaneous way.

2.3 Cover-Text: From Fragments to Text Reuses

When 16th century humanists began to collect fragments of textual sources, the main interest was in revealing and publishing the best traces of the most important authors of Classical antiquity. Later scholarship established philological and rigorous methods to find every possible evidence about lost authors, focusing the attention on the concept of *textual fragment* and therefore producing big collections of fragmentary authors and works, upon which we still depend for our knowledge of otherwise unknown literary figures of ancient times.¹¹⁶

Recent scholarship developed during the 20th and the 21st century has been moving the attention from the *fragment* to the *context* that preserves it. In an important paper concerning fragmentary historiography, Guido Schepens has coined the term *cover-text* to explain and define the complexity of extracting “fragments” from their source of transmission: “[...] the methodological key-problem the student of (historical) fragments has to face is invariably a problem of context: either there is no context for giving (some) meaning to a detached quotation, or only a drastically reduced context [...], or there is another context: the one of the work written by a later author in which the ‘fragment’ (how inappropriate is the term!) supposedly survives in some form. The latter is the way the great majority of fragments of historical works have survived, a fact which entails important consequences as to method. Of course, the context of the later work must not always entail a distortion of the original meaning of a fragment, but it often does. The student of historical fragments should be aware of the fact that his basic working material — the texts quoted with the author’s name — consists for the greater part of references that are made with a special purpose, mostly in a critical or polemical spirit. We know that ancient historians, when they wanted to take advantage of what their predecessors had written, usually preferred an anonymous reference to one by name. As a rule they only cited their precursor’s name when they disagreed or wanted to show off their better knowledge. This tendency carries two important methodological implications.

116 See section 1.2.

First, the reference by name always needs to be examined critically before we can think of using it as evidence for reconstructing the contents of lost works. [...] In view of the paramount importance of the analysis of the (con)text of the later works in which the ‘fragments’ survive, one could perhaps think of calling these works *cover-texts*. Apart from being a convenient short-hand, the notion ‘cover-text’ conveys – I believe, better than the phrases commonly used (‘sources of fragments’ or expressions like the ‘citing’ or ‘quoting’ later authors) – the consequential and multiple functions these texts perform in the process of transmitting a fragment. [...] the word ‘cover’ has the triple meaning of: to conceal, protect or enclose something. These are all activities which the later authors perform (or can perform) when transmitting a precursor text: they, first of all, *preserve* (= protect from being lost) texts drawn from works that are no longer extant; very often, too, they more or less *conceal* the precursor text (for characteristics such as the original wording and style of the precursor text are no longer discernible; often also fragments seem to ‘hide’ in the cover-text, so that one can only guess where a paraphrase begins or where a quotation ends); and, last but not least, the cover-text *encloses* the precursor text: it is inserted or enveloped in a new context, which may impose interpretations that differ considerably from the original writer’s understanding of his text. [...] Much work on Greek historiography still fails sufficiently to take into account the full implications of the fact that in many cases we are dependent on cover-texts. [...] Second, any study of fragments needs, if possible, to be supplemented, though under strictly limited conditions, by an examination of the indirect tradition: such an investigation must always take the named fragments as its starting-point, lest it end up in the speculative, circular arguments of unwarranted *Quellenforschung*.¹¹⁷

Guido Schepens points at two fundamental components of modern philological methods for dealing with fragmentary texts: 1) the role of the context that transmits information about lost texts by citing and quoting them in many different ways, and 2) the necessity of a careful examination of the indirect tradition of lost texts, which means a comparison between the context of the fragment and other sources. Schepens doesn’t use the expressions *text reuse* and *textual alignment*, which are now key terms of many projects in the digital and computational humanities for exploring and developing techniques of text reuse detection and intertextual services. The goal of these services is to semi-automatically identify and represent relations and reuses of texts that include phenomena such as quotations, allusions, paraphrases and plagiarism.

In the three following sections, I introduce new projects that have been applying text mining techniques to historical sources for text reuse detection and intertextual alignment. The application of these techniques to historical texts is

117 Schepens (1997) 166–167. Cf. also Grafton (1997) 143 and Gorman/Gorman (2014) ch. 3.

still at the beginning and definitely needs more data and further developments. Nevertheless, it is very interesting to see how recommendations expressed by traditional philologists like Guido Schepens and experiments performed by digital philologists are converging into the idea of focusing the attention on primary sources and of carefully exploring them as precious contexts of transmission of further information about the ancient world.

2.3.1 Text Reuse Detection

In the last ten years many experiments have been carried out for applying text reuse detection techniques to many different kinds of textual and electronic resources.¹¹⁸ Experiments and projects are also currently in progress for applying these techniques to historical documents.¹¹⁹ In this case, the detection is performed for text reuses of still surviving sources where it is possible to compare the reuse with the original text from which the reuse itself derives. As we have seen in section 1.3, most of what was written in Classical antiquity has been lost and now we rely on reuses of a lost textual heritage. The development of technologies for detecting reuses of lost texts has still to come and, as we will see in the next chapters, it still requires the creation of more digital resources and the preparation of training data.¹²⁰

The Proteus Project. This is a project developed at the Center for Intelligent Information Retrieval at the University of Massachusetts Amherst for building and evaluating research infrastructure for scanned books.¹²¹ The goal of the

118 Barrón-Cedeño/Basile et al. (2010); Sánchez-Vega/Villaseñor-Pineda et al. (2010); Trillini/Quassdorf (2010); Smith/Manmatha et al. (2011); Alzahrani/Salim et al. (2012); Smith/Cordell et al. (2013); Ganascia/Glaudes et al. (2014); Smith/Cordell et al. (2014); Colavizza/Infelise et al. (2015). On text reuse detection from the web, see Potthast/Hagen et al. (2013) and Hagen/Potthast et al. (2017). Martin Potthast has been also implementing *Picapica*, which is a text reuse search engine for comparing a text to *Wikipedia* in ten modern languages (English, German, Spanish, French, Italian, Dutch, Swedish, Portuguese, Catalan and Basque): <http://www.picapica.org>.

119 Lee (2007); Bamman/Crane (2008b); Bamman/Crane (2009); Büchler/Geßner/Eckart et al. (2010); Büchler/Geßner/Heyer et al. (2010); Büchler/Crane et al. (2012); Büchler/Crane et al. (2013); Büchler (2013); Büchler/Geßner et al. (2013); Büchler/Burns et al. (2014); Gorman/Gorman (2016); Pöckelmann/Dähne et al. (2020). Text reuse detection is also now part of KITAB (*Knowledge, Information Technology, and the Arabic Book*), which is a project for studying the formation and development of the written Arabic tradition with digital methods: <http://kitab-project.org>.

120 I'm very grateful to the German Federal Ministry of Education and Research (BMBF) for supporting the *Historical Text Reuse Data Workshop* that I organized at the University of Leipzig on July 12–13, 2017 and that offered many fruitful discussions on text reuse of historical sources.

121 <http://books.cs.umass.edu/mellon>

project (*Proteus Books*) is to work with unstructured scanned book collections, as for example the *Internet Archive*, and help scholars in the humanities navigate and use them in an easier way. The project has five components: 1) language identification, 2) duplicate detection, 3) duplicate alignment, 4) entity extraction, and 5) quotation detection. The project has identified the language of 3,628,227 OCR'd books from metadata of the *Internet Archive*, individuating also language identification differences and errors.¹²² *Proteus* has then acquired the canonical text of 803 English works and of 401 Latin works from the *Perseus Digital Library*, in order to compare them with English and Latin OCR'd books from the *Internet Archive* and find full and partial duplicates of the canonical works. After performing duplicate detection, the OCR'd text of duplicates have been aligned with the text of canonical works to identify corresponding portions of the works. *Proteus* has also performed Named Entity Recognition on 1,072,356 books from the *Internet Archive* to identify people, places, organizations and things, and visualize them in a JSON format. In addition to NER, the project has been working on finding matching quotations to see all occurrences of quotations of canonical works in OCR'd books from the *Internet Archive*, including commentaries and not only copies of canonical works. An example is the *Germania* by Tacitus, which is identified as urn:cts:latinLit:phi1351.phi002 in the *Perseus Catalog*. *Proteus* allows scholars to read sections of the work and visualize portions of text that have been identified as quotations in OCR'd volumes from the *Internet Archive*. The system provides an alignment between the canonical text of works and the OCR'd output with links to page images of the relevant OCR'd book (fig. 2.41).¹²³

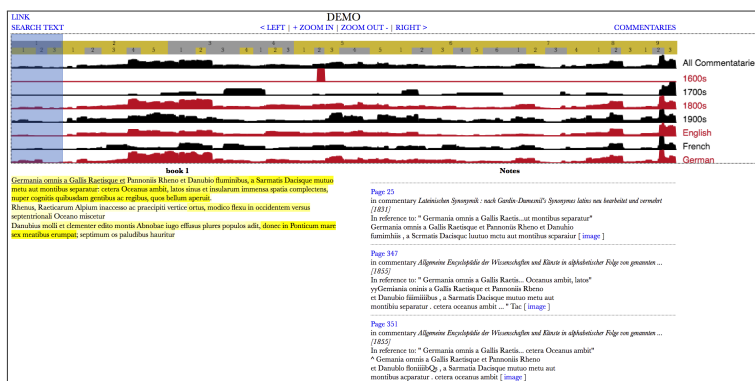


Figure 2.41. The Proteus Project: quotations of Tacitus' *Germania* in OCR'd books

122 *Proteus* uses ten languages: English, French, German, Spanish, Italian, Latin, Portuguese, Dutch, Danish and Swedish.

123 <http://books.cs.umass.edu/mellon/quotes/hb/urn:cts:latinLit:phi1351.phi002>

eTRAP (*Electronic Text Reuse Acquisition Project*). This is a text reuse project also with a focus on historical languages.¹²⁴ The project has been developing the TRACER machine, which is a command line engine for text reuse detection written in Java. The goal is to semi-automatically detect text reuse between two or multiple texts in the same language. The project provides guidelines for preparing *corpora* that can be analyzed with TRACER and the machine is continuously improved thanks to the feedback gathered by tutorials and workshops organized by eTRAP at international conferences and events.¹²⁵ The project has been developing two related research works: the *Digital Breadcrumbs of Brothers Grimm* and the *Tracing Authorship In Noise* (TrAIN) for detecting traces of the Brothers Grimm's tales and authorship attribution.

TRAViz (*Text Reuse Alignment Visualization*). eTRAP makes also use of TRAViz, which is a JavaScript library that “generates visualizations for Text Variant Graphs that show the variations between different editions of texts.”¹²⁶ As stated on the website of the project, TRAViz supports the collation task by providing methods to align various editions of a text, visualize the alignment, improve the readability for Text Variant Graphs compared to other approaches, and interact with the graph to discover how individual editions disseminate. The project provides examples with different English and German translations of the *Bible*.

Text reuse detection experiments have been also carried out as part of the project eAQUA (*Extraktion von strukturiertem Wissen aus Antiken Quellen für die Altertumswissenschaft*), which was developed at the University of Leipzig for the application of text mining methods and techniques to ancient Greek and Latin sources.¹²⁷ The first phase of the project (2008–2011) included 8 sub-projects: 1) *Projekt Atthidographen*, 2) *Projekt Platon*, 3) *Projekt Metrik*, 4) *Projekt Camena*, 5) *Projekt Inschriften*, 6) *Projekt Papyri*, 7) *Projekt Fehlererkennung*, and 8) *Projekt Mental Maps*.¹²⁸ The second phase of the project (2011–2013) further applied text mining techniques to specific research questions arising from sources related to the Atthidographers (co-occurrences) and Plato (quotations and text reuses). eAQUA has an online portal where it is possible to read information about the project and access demo versions of the tools *Kookkurrenz-Analyse* and *Zitationen*. Both tools analyze data from different *corpora* that are free or protected by copyright. Co-occurrences are searchable in the free *corpora* of the *Codex Sinaiticus*, the *Deutsches Textarchiv* (DTA), *Epiduke* (Duke Databank of Documentary Papyri), *Herodot*, *PHI Latin Texts* (PHI 5) and the *Perseus Digital Library* (Greek,

124 <https://www.etrapp.eu>

125 Böhler (2013); Böhler/Burns et al. (2014).

126 See <http://www.traviz.vizcovery.org>, Jänicke/Geißner et al. (2014), and Yousef/Janicke (2021).

127 <http://www.eaqua.net>

128 Schubert/Heyer (2010); Schubert (2011).

Latin, and Renaissance Shakespeare).¹²⁹ Limited to accounts belonging to the project is the access to the *Bibliotheca Teubneriana Latina* (BTL), the *Patrologia Latina* (PL/ML) and the *Thesaurus Linguae Graecae* (TLG CD-ROM E). The tool *Zitationen* offers access to the free *corpora* of the *PHI Latin Texts* (PHI 5) and the *Perseus Digital Library* (Greek and Latin), and through a limited account to the *corpus* of the *Thesaurus Linguae Graecae* (TLG CD-ROM E).

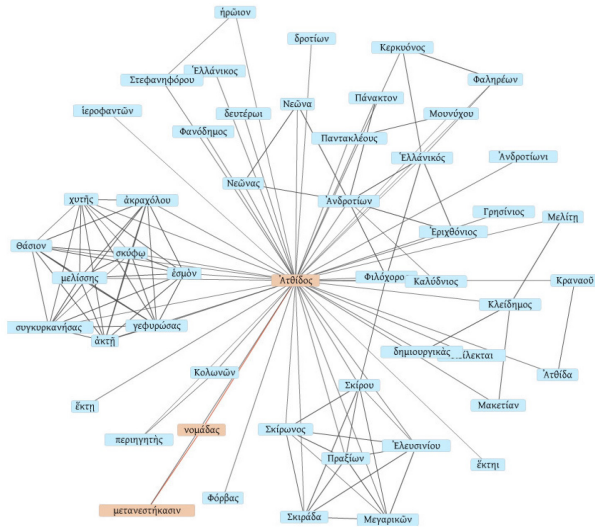


Figure 2.42. eAQUA: graph of Ἀτθίδος

As far as lost authors are concerned, eAQUA has published experimental results on the *Atthidographers*, which is an expression alluding to a literary genre that includes a group of ancient Greek local historians who wrote about Athens and Attica but whose works are now lost.¹³⁰

An example is the use of the term Ἀτθίδος, which is the genitive of the adjective Ἀτθίς used as a title of histories of Attica written by the Atthidographers: ἡ Ἀτθίς (συγγραφὴ) (*Atthis* and plural *Atthides*).¹³¹ eAQUA has analyzed co-occurrences of Ἀτθίδος in the TLG (CD-ROM E) and has produced visualizations through graphs, revealing in this way interesting connections about the use of this work title in ancient Greek literature.¹³² The online tool *Kookkurrenz-Analyse* allows to search and visualize lists of co-occurrences of ancient Greek words, including Ἀτθίδος on which are based the results presented by Schubert

129 On the use of *Epiduke* through eAQUA, see Reggiani (2017) 186–187.

130 Jacoby (1949); Berti (2009b) (*Introduzione*); Bearzot/Landucci (2010); Schubert (2010a).

131 Jacoby (1949) 80; Harding (1994) 1; Berti (2009b) 7–8.

132 See fig. 2.42; Bünte (2010); Schubert (2011) 38–44.

eAQUA: Kooccurrenz-Analyse ?

Corpus auswählen: **Thesaurus Linguae Graecae (TLG)** ?

Virtuelle Tastatur: Keine Griechisch ?

Wort-Suche: **αἰθίδος** ? oder Wort-ID: **19175** ?

Trefferanzahl: **500** ?

Sortierung: häufig selten ?

Stoppwörter anzeigen: ja nein ?

Start

TLG: Αἰθίδος [19175] - Häufigkeit: 276 ?	
Wörter mit derselben normalisierter Form:	Αἰθίδος (19175); Ἄιθιδος (854980); αιθίδος (1081066);
Wörter mit ähnlichem Zusammenhang:	κατηρίμκα [743211]; συγκυρκανήσας [492033]; σκόφη [307369]; ἀκραχόλου [37832]; Ἀνδρότιον [20787]; Φιλόκορος [9786]; χυτῆς [147037]; Ἐλλάνικος [8231]; Πράξιον [399130]; Σκίραθα [39918]; Σκίρωνος [195268]; Μεγαρικῶν [144633]; Ἐλευσινίου [131400]; ἀσπῆ [55210]; Φορτίης [15995]; Φελοπιμος [4891]; Σκίρου [154413]; Δευκιάδης [139980]; Αἰτίων [115823]; ἱστορεῖ [2074]; FHG [4256]; Ἐφορος [7974]; ἐπιόν [80383]; Πανταλάου [296413]; Ἰουαρκῶν [13528]; Ἡράδοτος [3341]; γλυκὸν [13069]; κομῆτος [14366]; ποιμα [17978]; μιλῶσις [32731]; Ἐλλάνος [63393]; ὄραον [5662]; νεφουριος [79647]; Μουνοῦκα [226979]; Ἰπποβουτῆδος [54572]; Εὐδαρος [13888]; κελήσθαι [12178]; Φερεκούσις [11238]; Κλειδιώσις [63894]; Περηνήσιος [179264]; Χρονικὸν [69832]; τριώσιον [16637]; FGH [1346]; ἱστοριῶν [10503]; ἔκτα [19705]; Ἀθηνῶν [5602]; Απολλοδώρου [5447]; Θουκυδιάδης [3144]; F [1642]; Αἰθίδη [5473]; Παρκίῶν [51863]; μάντεως [16957];
Signifikante Kooccurrenzen:	Φιλόκορος (60); Ἀνδρότιον (53); ἐν (165); Ἐλλάνικος (26); Κλειδιώσις (12); ἔτις (111); F (17); ἔκτι (12); β' (22); Στεφανόφορος (6); Ἐλλάνος (7); ἀκραχόλου (9); Πανταλάου (9); ἀσπίς (3); Κραναοῦ (6); πολλαὶς (10); Μουνοῦκα (5); συγκυρκανήσας (4); Φανόδημος (6); Σκίρου (5); χυτῆς (5); Πράξιον (4); Μακτιῶν (4); Σκίραθα (4); Ἐλευσινίου (5); Νεώνα (4); FHG (9); α' (15); Ποσειδάων (9); μιλῶσις (6); B (13); ὤς (62); ἐπιόν (5); νεφουριος (5); ὄραον (5); Σκίρωνος (4); ἀσπῆ (3); Δευτέρου (5); μετανεατικῶν (5); Ἐρριῶν (3); Καλλόνιος (3); Μεγαρικῶν (4); τρίτω (9); καθά (8); νέας (7); Ἐκατόσις (7); πρώται (10); φησὶν (17); Φορβας (4); ἔκτω (6); Κολωνίαν (3); Νεώνα (3); Ἐραχθῶνος (4); Ἡράδοτος (7); ὄρατιον (3); Φαλίμερον (3); μάντεως (5); ἱστοριῶν (3); κούρατος (5); διαλεκτοῦ (5); ἰ; πόλις (13); φησὶν (14); κελήσθαι (5); ἰονομοῖσι (9); Ἀττικῆς (6); Ἀνδρότιον (3); διαλεκταὶ (4); ἱεροφαντῶν (3); ἰδῆμα (5); Κερκυκῶνος (3); Ἡφαίστου (5); Πάνκτων (3); Βιθίδη (6); φιλοκβαριατικῶν (2); Πανδῆνος (2); Καλυδικῆς (2); Βουλίης (2); δειλόφουα (2); κατεπιθεῖσθαι (2); κατηρίμκα (2); Μυκαθῆς (2); Ἀλικαρνασσῶς (2); Ἀνδάν (2); Ἀρίστη (2); Ἀρκεσιανῶν (2); Διμήτρος (5); Μελιτή (3); γ' (10); ἀρκοντος (6); ἔκτι (3); α' (37); ἰγάγε (5); περιηγητής (3); Ὀμομοργικῶς (3); ποιμα (4); δευτέρω (7); τιμῶν (4); FGHist (4); διαδικασίας (2); Ἐπιήρων (2); μεταδόθη (2); Ἀμυργιῶν (2); Ἀμυργεῖς (2); Ἀριστοκρέων (2); ἀνομοζομένη (4); Αἰθίδη (3); τῆι (8); Ἀττικῶν (4); νομάδας (3); γλυκὸν (4); τετάρτω (5); Ἐκλείπας (2); Μακτιῶν (2); ἔθονον (6); ὁμονομοῖ (2); Βυθόλου (2); γυνακοκομοῖ (2); Ἐρικμιοῖσιαν (2); φουραῶν (5); τῆος (9); ἀπεκκερατονήθη (2); Καρναῖου (2); σιγνεγομένης (2); Κερκιοῖσιν (2); Δερεκ (2); ἦς (11); δωδεκῆθι (3); ΝῆΞ (33); ἐξ (2); ἐδωλῶσιν (2); ποιητή (2); ἰγνηγομένησιν (3); Ἀλαπιεγκόνηρον (2); διασηφίσιον (2); Ἀγγραυλον (2); Μεγάρος (3); διεξήθεν (3); τρίτη (5); μάντε (6); συμπεριβουτῆς (2); κελήσθαι (2); Ἐπιπολιῶν (2); ἀρακῆς (4); μετανεατικῶς (2); μαγιέρον (3); Ἰλιον (5); ἦδη (3); Ἰβου (5); ἄστικῶν (2); πεβήρημον (2); Φιλοκόρου (2); κελῶθ (2); α' (10); θέρον (2); ἀνομάδατος (2); Δεδάρω (2); Κακωνικῶν (2); ἐνεπέταστα (2); Ἐλληνιστοῦ (3); γέροντα (4); κατεστραῖον (4); Ἀθηνῶν (4); Ἰπποβοῦτῆ (2); Μακρόνηον (4); Κίρκω (2); Ἀλφαιῶν (2); κόδδων (2); Ἀλλυροῖν (2); Πηγαί (2); Μόλις (2); Ἐπιτομήν (2); Ἰλλ (5); εὐογον (2); σποραδῶν (3); συγκαταλέγει (2); Ἀλικαρνασσῶς (2); Ἐπιόνυμος (3); Ἰν (12); Συμαρτῆς (2); δοσρα (2); Ἀλλεπιῶν (2); θράμα (3); Ἀθηναίω (6);
Network [Frequenz]	
Network [Log-Likelihood]	

Figure 2.43. eAQUA: co-occurrences of Αἰθίδος

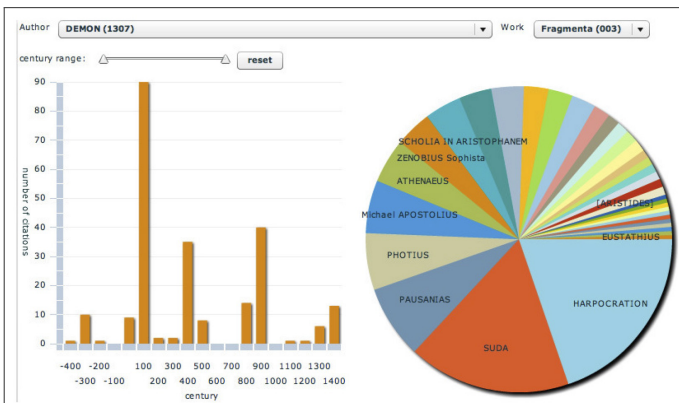


Figure 2.44. eAQUA: witnesses of Demon

(2011).¹³³ Other experiments in eAQUA have been performed using the *CitationGraph* to visualize sources preserving quotations and text reuses of the lost works of the Atthidographers.¹³⁴

Figure 2.44 shows an example of visualization of the sources that preserve text reuses of the lost work of the Atthidographer Demon. Charlotte Schubert has also experimented with text reuses of Clidemus. The *CitationGraph* doesn't reveal new fragments of the lost Atthidographer, but allows to visualize the reuses of his lost works with a perspective not dependent on the editorial selections and arrangements of scholars who have published editions of Clidemus' fragments.¹³⁵ These tools still need further developments and the ingestion of other digital sources to produce more results and try to obtain previously unexplored relations among texts.¹³⁶ In any case, they already offer a first set of experimental functions and visualization possibilities that allow scholars to go beyond the limits of traditional printed editions where the selection and the presentation of source texts of fragmentary authors strongly depend on the decisions of the editor.¹³⁷

Trismegistos (reuse of texts). A special and in some way related case of "reuse of texts" is presented by the project *Trismegistos*.¹³⁸ This case concerns physical documents and the reuse of papyri, stones and other materials as writing surfaces, which was a very common practice in antiquity. Text reuse is not the focus of *Trismegistos*, but its team has been starting to devote a section of the project with different categories for specifying if there is a relation among texts written on the same physical objects.¹³⁹ Even if strictly related to the material reuse of writing objects, this tool has the possibility to expand and reveal relations among ancient

133 See fig. 2.43; Schubert (2018).

134 Bünte (2010); Schubert (2010b).

135 Schubert (2010b) 51–54.

136 For example, the project still includes the texts of the CD-ROM version E of the TLG, which is now superseded by the online version that constantly adds new sources to the TLG *corpus*.

137 Cf. Schubert (2010b) 54. eAQUA and its *CitationGraph* were also used for a new research on the lost author Ehippus of Olynthus and Nicobule: Pfeil (2013). Other results in eAQUA have been published for detecting quotations in still extant sources, like Plutarch and Plato: Schubert (2010b); Schubert/Klank (2012); Schubert (2017); Geßner (2010). As far as the reception of Plato in antiquity is concerned, further research is now developed as part of the project *Digital Plato* at the University of Leipzig: Pöckelmann/Ritter et al. (2017); Schubert (2019); Pöckelmann/Dähne et al. (2020).

138 See https://www.trismegistos.org/tm/search_reuse.php. On *Trismegistos* see p. 69.

139 There are many different cases where texts could be written on an object because related or not related to the text already written on it. *Trismegistos* provides an experimental search engine for exploring "sets of texts connected with reuse." Interesting examples are documents that were joined in a second stage for their users' convenience and because of their complementary contents. In this cases *Trismegistos* provides specific categories for these types of connection.

texts generating further search criteria and types of reuse with the addition of new records and the collaboration with other projects.

2.3.2 Intertextual Analysis

Strictly connected to text reuse is the concept of *intertextuality*, which aims at exploring the intricate structure of meaningful relationships between texts. As for Classical sources, the theory of intertextuality was originally developed in the field of Latin literature, but is now expanding to cover other genres as for example Greek historiography.¹⁴⁰ Digital philologists have been recently explored methodologies for digitally representing intertextuality and for training and testing the machine to automatically detect intertextual matches between historical texts.¹⁴¹ Taking into consideration the complexity of historical sources and the lack of complete and fully comprehensive digital *corpora*, a lot of work has still to be done in order to prepare data and get proper results from it. Nevertheless, the tools that I present in this section are already generating interesting results that could be expanded with more data and a bigger involvement of the scholarly community.

Tesserae. *Tesserae* is a collaborative project of the Departments of Classics and Linguistics of the University at Buffalo, the Department of Computer Science and Engineering of the University of Notre Dame, and the Département des Sciences de l'Antiquité of the University of Geneva.¹⁴² The project offers a free and open web interface for exploring intertextual parallels and detecting allusions in Latin poetry by generating lists of lines that share two or more words within a single line or phrase regardless of inflectional changes.¹⁴³ *Tesserae* makes use of *corpora* from different databases, such as *The Latin Library*, the *Perseus Digi-*

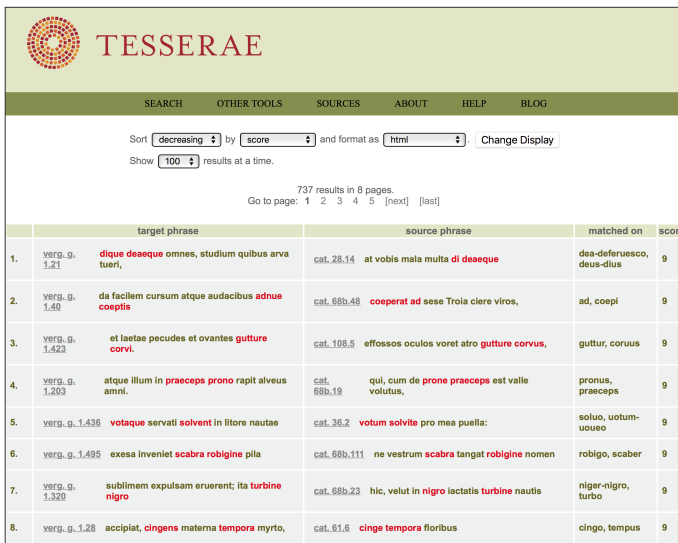
140 Berti (2012) 442–446 with bibliography; Coffee/Koenig et al. (2012) 383–384; Coffee (2018). On the application of intertextual concepts to Classical historiography, see the papers on *Allusion and Intertextuality in Classical Historiography* presented at the APA Annual Meeting in San Antonio (Jan. 8, 2011), and on *Historiography, Poetry, and the Intertext* and *Intertextual Relationships Between Poetry, Prose and Historiography* presented at the APA Annual Meeting in Seattle (Jan. 4, 2013) and at the CA Annual Conference in Reading (Apr. 6, 2013). These papers are available online on the website of *Histos. The On-line Journal of Ancient Historiography*. On *intratextuality*, which is the interaction between parts of the same text or body of texts within a single author, see Harrison/Frangoulidis et al. (2018).

141 On how hypertext is “a fundamentally intertextual system,” see Landow (2006) 55.

142 See <http://tesserae.caset.buffalo.edu> that offers the new version (5) of the project. Version 3 is currently updated to continue its functionality. Examples in this book are taken from version 3.

143 See Coffee/Koenig et al. (2013) for an evaluation of *Tesserae* search methods by comparing book 1 of Lucan's *Civil War* with Vergil's *Aeneid*. See also Forstall/Coffee et al. (2015), Coffee (2018), and Coffee (2019).

tal Library, DigilibLT (*Digital library of late-antique Latin texts*), the *Open Greek and Latin* project, *Musisque Deoque*, and the *Corpus Scriptorum Latinorum*.¹⁴⁴ As of 2021, the project allows to begin testing with Greek and English texts, and offers other experimental tools: Latin Multi-Text Search (cross-references discovered parallels against the rest of the Latin corpus), Greek Multi-Text Search (cross-references discovered parallels against the rest of the Greek corpus), LSA Search Tool (search for thematic similarities even where phrases have no words in common), Tri-gram visualizer (customizable, color-coded visualization of 3-gram concentrations), Full-text display (displays the full text of the poems with references highlighted in red), and Lucan-Vergil benchmark test (perform a search of Lucan's *Pharsalia* Book 1 against Vergil's *Aeneid*, and compares the results against a 3000-parallel benchmark set).



	target phrase	source phrase	matched on	score
1.	VERG. G. 1.21 dique deaeque omnes, studium quibus arva lueri,	cat. 28.14 at vobis mala multa di deaeque	dea-deferuesco, deus-dius	9
2.	VERG. G. 1.40 da facilem cursum atque audacibus adnue coepit	cat. 68b.48 coeperat ad sese Troia clere viros,	ad, coepi	9
3.	VERG. G. 1.423 et laetae pecudes et ovantes guttire corvi .	cat. 108.5 effossos oculos voret atro guttire corvus ,	guttur, corvus	9
4.	VERG. G. 1.203 atque illum in praeceps pronu rapit alveus amni.	cat. 68b.19 qui, cum de prone praecipis est valle volutus,	pronus, praecipis	9
5.	VERG. G. 1.436 votaque servati solvent in litore nautae	cat. 36.2 votum solvite pro mea puella:	solvio, uotum- uovio	9
6.	VERG. G. 1.495 exesa inveniet scabra robigine pila	cat. 68b.111 ne vestrum scabra tangat robigine nomen	robigo, scaber	9
7.	VERG. G. 1.320 sublimem expulsaam eruerent; ita turbine nigro	cat. 68b.23 hic, velut in nigro lactatis turbine nautis	niger-nigro, turbo	9
8.	VERG. G. 1.28 accipiat, cingens materna tempora myrto,	cat. 61.6 cinge tempora floribus	cingo, tempus	9

Figure 2.45. *Tesserae*: comparison of Catullus' *Carmina* with Vergil's *Aeneid* book 1

Figure 2.45 shows the example of a comparison between Catullus' *Carmina* and the first book of the *Aeneid* of Vergil. The search generates 737 results and in each case displays two common words between the *target text* (alluding text: Vergil) and the *source text* (alluded-to text: Catullus). Advanced search options allow users to set different parameters such as units to be compared (lines or phrases), features to be matched across texts (exact word, lemma, semantic match, lemma + semantic match, and sound), number of stop words and the stoplist basis to

144 Texts from these databases are modified by changing the markup and sometimes also the orthography, and by removing all punctuation and capitalization.

determine frequencies for the stoplist, score and frequency basis for getting rarer words closer together, and maximum distance and distance metric to exclude matching words that are too far from each other. Results allow to visualize highlighted matching words, to re-sort results with sort options, to read both target and source texts in their entire context, and to export data in CSV, TSV, and XML formats. Regarding Latin poetry, the automatic detection of parallel phrases in *Tesserae* is producing promising results. As reported by the team of the project, it “recovers approximately a third of the parallels captured by traditional commentators, and adds a third not previously recorded.”¹⁴⁵ As far as prose texts and Greek sources are concerned, the tool allows users to explore intertextual parallels and obtain many results that need to be further selected and verified in order to test the effectiveness of the algorithms.

TLG Intertextual Phrase Matching. As part of new features offered by the online version of the *Thesaurus Linguae Graecae* (TLG), there is also a tool for intertextual phrase matching. The service is based on n-gram comparison (bigrams and trigrams) for detecting sequences of two or three content words shared between two texts in order to identify common text passages.¹⁴⁶ N-grams are used in the TLG for 1) comparing two texts or two authors side by side (*N-Grams*), 2) for seeing attestations of phrases in the *corpus* (*Browse – Browse one text*), and 3) for selecting two passages and viewing their similarities (*Browse – Parallel browsing*).¹⁴⁷

A first example is a comparison between the *Historiae* of Thucydides (ed. Jones-Powell: tlg0003.001) and all texts of Athenaeus of Naucratis (tlg0008).¹⁴⁸ Figure 2.46 shows 8 results, 6 of which have as a target text the *Deipnosophists* and 2 the *epitome*. As for the *Deipnosophists*, there are matches in books 5, 9, 10 and 11. Passages in books 5 and 11 have been also detected in printed editions of the *Deipnosophists* by August Meineke, Georg Kaibel, and Douglas Olson.¹⁴⁹ Passages in books 9 and 10 are not referred to in printed editions and seem not to be relevant.¹⁵⁰ As for the *epitome*, only the passage from the summary of book 11 of the *Deipnosophists* is pertinent.¹⁵¹ Missing are a direct quotation of a passage

145 Coffee/Koenig et al. (2012) 386.

146 The comparison is based on lemmata and the order of words within n-grams is ignored. Stop-words that don't contribute to the meaning of the comparison are removed.

147 Descriptions of these tools are available on the TLG website. For a recent review of the TLG intertextual phrase matching, see Boogert (2019).

148 The texts of Athenaeus in the TLG are constituted by the *Deipnosophists* (ed. Kaibel: tlg0008.001), one fragment of *On the kings of Syria* (FGrHist 166 F 1: tlg0008.002), and the *epitome* of the *Deipnosophists* (ed. Peppink: tlg0008.003).

149 Ath., *Deipn.* 5.55 = 215f and 216a = Thuc. 4.96.5, 8; *Deipn.* 11.57 = 478f = Thuc. 7.87.2. The passage in book 11 is one, but the TLG has chopped it up into two consecutive matches because it is an extended passage.

150 Ath., *Deipn.* 9.29 = 383a = Thuc. 4.50.2; *Deipn.* 10.87 = 458a = Thuc. 5.111.4.

151 Ath., *Epit.* 2.2.58.10 Peppink = Thuc. 7.87.2.

from Thucydides and of course direct references to his name without quotations or paraphrases of his text.¹⁵²

INTERTEXTUAL PHRASE MATCHING*

Source Text		Target Text	
THUCYDIDES (0008) Historiae (001)		ATHENAEUS (0008) All	
Compare Texts			
Lines of context: 1		Results per page: 20	
Prev		Next	
1	Th.Hist. (0003.001) 4.50.2.6 λέγων· εἰ οὐκ ἔβουλόμην σαφῆς λέγειν, πᾶσαι μετὰ		Ath.Deipn. (0008.001) 9.29.37 εἰς σαφέστερον θ' ὁ βούλοι μοι λέγειν;
2	Th.Hist. (0003.001) 4.96.5.1 (5) ἔμπροσθοῦσιν, καὶ ἐνέβη Πηγώνου περιψέφμενος δύο τέλη τῶν ἰσπίων ἐκ τοῦ ἀφανοῦς περὶ τὸν λόφον, ὡς ἐπόναι		Ath.Deipn. (0008.001) 5.55.32 Πηγώνου (Thuc. IV 96) δύο τέλη περιψέφμενος τῶν ἰσπίων ἐκ τοῦ ἀφανοῦς περὶ τὸν λόφον, τότε γὰρ οἱ
3	Th.Hist. (0003.001) 4.96.8.1 (8) βουλοῖα δὲ ἐρεψόμενος ἔστησαν, καὶ μάλιστα οἱ ἰσπίες οἱ τε αὐτῶν καὶ οἱ Λοκροῦν βεβρηγόμενες ἄρα τῆς τροσῆς		Ath.Deipn. (0008.001) 5.55.36 τὸ ὄρος· βουλοῖα δ' ἐρεψόμενος ἔστησαν καὶ μάλιστα οἱ ἰσπίες οἱ τε αὐτῶν καὶ οἱ Λοκροῦν, τοιοῦτοι αὖν
4	Th.Hist. (0003.001) 5.111.4.7 κρείσσου καλοῦς προσφέρονται, πρὸς δὲ τοῖς ἥσσουσ ἡμεῖροι		Ath.Deipn. (0008.001) 10.87.10 ἀγαθὸς ἂν εἴη γὰρ φέρον καλοῦς κακά. (110)
5	Th.Hist. (0003.001) 7.87.2.5 ἀνεκτοί, καὶ μὰρ ἄμα καὶ διήγη ἐπέζοντο (εἰδοῦσιν γὰρ (5) αὐτῶν ἕκαστη ἐπὶ ὅστω μῆγας κοτύλην ἕδατος καὶ δύο		Ath.Deipn. (0008.001) 11.57.21 ὄης (7.87)· ἔδοσαν μὲν αὐτῶν ἕκαστη ἐπὶ ὅστω μῆγας κοτύλην ἕδατος καὶ δύο κοτύλας οἴνου· Ἀριστο-
6	Th.Hist. (0003.001) 7.87.2.6 αὐτῶν ἕκαστη ἐπὶ ὅστω μῆγας κοτύλην ἕδατος καὶ δύο κοτύλας οἴνου), ἄλλα τε δύο εἰδὸς ἐν τῇ τοιοῦτῃ χωρῇ		Ath.Deipn. (0008.001) 11.57.22 κοτύλην ἕδατος καὶ δύο κοτύλας οἴνου· Ἀριστο-
7	Th.Hist. (0003.001) 5.111.4.7 κρείσσου καλοῦς προσφέρονται, πρὸς δὲ τοῖς ἥσσουσ ἡμεῖροι		Ath.Deipn.Epist. (0008.003) 2.2.49.21 ὁμοίως λαμβάνει· ἀγαθὸς ἀνηγέλου· ἂν ὁ φέρον τάχαθ· ἄρα καὶ ὁ φέρον καλοῦς κακά. Οἰμητικοὶ ἀπὸ τοῦ εἰπε το εἶπε Ἀ'
8	Th.Hist. (0003.001) 7.87.2.6 αὐτῶν ἕκαστη ἐπὶ ὅστω μῆγας κοτύλην ἕδατος καὶ δύο κοτύλας οἴνου), ἄλλα τε δύο εἰδὸς ἐν τῇ τοιοῦτῃ χωρῇ		Ath.Deipn.Epist. (0008.003) 2.2.58.10 ταὶ κοτύλην ἀνομοίαια, οἷς Θεοσευδῆς κοτύλην δὲ κοτύλας οἴνου. Ἀριστοφάνης· ἀφίμων γόνινας τρεῖς κοτύλας

Figure 2.46. TLG intertextual phrase matching: comparison of Thucydides with Athenaeus

Concerning lost texts, we can try to compare the fragments of Istros the Calimachean (t1g1450) with Athenaeus of Naucratis (t1g0008).¹⁵³ In this case, we obtain 19 results (partial screenshot in fig. 2.47). Considering that passages are chopped up into consecutive matches, the actual detected fragments are four from the *Deipnosophists* and the *epitome* (fr. 14, 35, 38, 43) and they correspond to all those published in the FHG. In this case, the intertextual phrase matching is useful to align the lost text of the fragments as they were edited by Karl Müller in the FHG and the original text of the *Deipnosophists* in the edition by Kaibel and of the *epitome* in the edition by Peppink.

Through the TLG *Parallel browsing* it is also possible to select one fragment of Istros and compare its text in the edition of the FHG with the whole context of the passage of the *Deipnosophists* in the edition by Kaibel (e.g., fr. 14 in fig. 2.48). As part of the *comparing* functions, the TLG now offers also the possibility

152 Ath., *Deipn.* 1.42 = 23b = Thuc. 1.70.5 (the quoted passage is νικώμενοι ἐπ' ἐλάχιστον ἀναπίπτουσιν, which is detected in the text of the *Deipnosophists* when enabling n-grams in the *Browse one text* section of the TLG); *Deipn.* 3.73 = 108f = Thuc. 7.33.4; *Deipn.* 5.15 = 189c = Thuc. 4.103.1; *Deipn.* 5.55 = 215d = Thuc. 5.2.1.

153 Istros' fragments in the TLG are from Müller's FHG (t1g1450.004) and from Mette (1978) (t1g1450.003).

INTERTEXTUAL PHRASE MATCHING*

Source Text		Target Text	
ISTER (1450) All		ATHENAEUS (0008) All	
<input type="button" value="Compare Texts"/>			
Lines of context: 1		Results per page: 20	
Prev		Next	

1	<p>Inter Fr. (1450.004) 14.1 (14) Athen. XIII, 3 Θεραΐες Ἐλένην ἀρπάζας, ἔξις καὶ Ἀριώδην ἥρπασεν. Ἰστρος γόν ἐν τῇ τετρασερακιδεκάτῃ τῶν Ἀττικῶν καταλέγων τὰς τοῦ</p>	1	<p>Ath.Deipn. (0008.001) 13.4.12 Μῆθειαν, Θεραΐες δὲ Ἐλένην ἀρπάζας ἔξις καὶ Ἀριώδην ἥρπασεν. Ἰστρος γόν ἐν τῇ τετρασερακιδεκάτῃ τῶν</p>
2	<p>Inter Fr. (1450.004) 14.2 ἔξις καὶ Ἀριώδην ἥρπασεν. Ἰστρος γόν ἐν τῇ τετρασερακιδεκάτῃ τῶν Ἀττικῶν καταλέγων τὰς τοῦ Θεραΐος γενναίως γενναίας, φησί τὰς μὲν αὐτῶν ἔξ</p>	2	<p>Ath.Deipn. (0008.001) 13.4.13 ἥρπασεν. Ἰστρος γόν ἐν τῇ τετρασερακιδεκάτῃ τῶν Ἀττικῶν (FHG I 420) καταλέγων τὰς τοῦ Θεραΐος γεννομένης γενναίας φησὶ τὰς μὲν αὐτῶν ἔξ ἔρωτος γεγενησίου, τὰς δ' ἔξ ἀρπαγῆς, ἄλλας δ' ἔκ</p>
3	<p>Inter Fr. (1450.004) 14.4 Θηραΐος γενναίως γενναίας, φησί τὰς μὲν αὐτῶν ἔξ ἔρωτος γεγεννησίου, τὰς δ' ἔξ ἀρπαγῆς, ἄλλας δ' ἔκ νομίμων γένων ἔξ ἀρπαγῆς μὲν Ἐλένην, Ἀριώδην,</p>	3	<p>Ath.Deipn. (0008.001) 13.4.14 Ἀττικῶν καὶ τὰς Κερκεϊνοῦ καὶ Σίνδος θυγατέρας νομίμως δ' αὐτῶν γήμια Μεμβρόων τῶν Αἰώντων μητέρα. Ἡσιόδου δὲ</p>
4	<p>Inter Fr. (1450.004) 14.5 ἐξ ἔρωτος γεγεννησίου, τὰς δ' ἔξ ἀρπαγῆς, ἄλλας δ' ἔκ νομίμων γένων ἔξ ἀρπαγῆς μὲν Ἐλένην, Ἀριώδην,</p>	4	<p>Ath.Deipn. (0008.001) 13.4.16 νήσθου, τὰς δ' ἔξ ἀρπαγῆς, ἄλλας δ' ἔκ νομίμων γένων ἔξ ἀρπαγῆς μὲν Ἐλένην, Ἀριώδην, Ἰσπολίτην καὶ τὰς</p>
5	<p>Inter Fr. (1450.004) 14.6 νομίμων γένων ἔξ ἀρπαγῆς μὲν Ἐλένην, Ἀριώδην, Ἰσπολίτην καὶ τὰς Κερκεϊνοῦ καὶ Σίνδος θυγατέρας νομίμως δ' αὐτῶν γήμια Μεμβρόων τῶν Αἰώντων μητέρα.</p>	5	<p>Ath.Deipn. (0008.001) 13.4.17 ἔξ ἀρπαγῆς μὲν Ἐλένην, Ἀριώδην, Ἰσπολίτην καὶ τὰς Κερκεϊνοῦ καὶ Σίνδος θυγατέρας, νομίμως δ' αὐτῶν γήμια Μεμβρόων τῶν Αἰώντων μητέρα. Ἡσιόδου δὲ</p>
6	<p>Inter Fr. (1450.004) 14.7 Ἰσπολίτην καὶ τὰς Κερκεϊνοῦ καὶ Σίνδος θυγατέρας νομίμως δ' αὐτῶν γήμια Μεμβρόων τῶν Αἰώντων μητέρα.</p>	6	<p>Ath.Deipn. (0008.001) 13.4.18 Κερκεϊνοῦ καὶ Σίνδος θυγατέρας, νομίμως δ' αὐτῶν γήμια Μεμβρόων τῶν Αἰώντων μητέρα. Ἡσιόδου δὲ</p>
7	<p>Inter Fr. (1450.004) 35b.1 (35b) Athen. III, 74, 1c Ἰστρος δ' ἐν τοῖς Ἀττικαῖς οὐδ' ἐγγενεσίου φησὶ τῆς Ἀττικῆς τὰς ἀπ' αὐτῶν γνομίμως λεγόμενας, ἵνα μόνον ἀποκρίσθαι οὐ κατωκυρτέας</p>	7	<p>Ath.Deipn. (0008.001) 3.6.17 Ἰστρος δ' ἐν τοῖς Ἀττικαῖς (FHG I 423) οὐδ' ἐγγενεσίου φησὶ τῆς Ἀττικῆς τὰς ἀπ' αὐτῶν γνομίμως λεγόμενας, ἵνα μόνον ἀποκρίσθαι οὐ κατωκυρτέας καὶ ἐπὶ</p>

Figure 2.47. TLG intertextual phrase matching: comparison of Istros with Athenaeus

PARALLEL BROWSING*

ISTER (1450) Fragments (004)	Lines: 20 Display: Greek Links: Active Select: Highlight similarities <input type="button" value="Show Texts"/>	ATHENAEUS (0008) Deipnosophistae (001)
---------------------------------	---	---

<p>Text Structure</p> <p style="text-align: right;">Fragment 14 line 1 60</p> <p style="text-align: center;">Prev Next</p> <p style="text-align: center;"><input type="button" value="Jump 30"/></p> <p>ISTER, Fragments. (1450.004) (14) Athen. XIII, 3 Θεραΐες Ἐλένην ἀρπάζας, ἔξις καὶ Ἀριώδην ἥρπασεν. Ἰστρος γόν ἐν τῇ τετρασερακιδεκάτῃ τῶν Ἀττικῶν καταλέγων τὰς τοῦ Θεραΐος γενναίως γενναίας, φησί τὰς μὲν αὐτῶν ἔξ ἔρωτος γεγεννησίου, τὰς δ' ἔξ ἀρπαγῆς, ἄλλας δ' ἔκ νομίμων γένων ἔξ ἀρπαγῆς μὲν Ἐλένην, Ἀριώδην, Ἰσπολίτην καὶ τὰς Κερκεϊνοῦ καὶ Σίνδος θυγατέρας νομίμως δ' αὐτῶν γήμια Μεμβρόων τῶν Αἰώντων μητέρα. (15) <i>Etiam Lex.</i> Θεὰς ἢ Ἀναΐδου. Ἐταμοῦ δὲ καὶ Ἀθήρων ἢ Ἀναΐδου, καὶ ἰσθὸν ἦν αὐτῆς, ὡς Ἰστρος ἐν δ'. (16) LIBER XVI. (16) <i>Harpocra.</i>: Τραπεζοφόρος Ἀντοκόγος ἐν τῷ περὶ τῆς ἱερείας οὐ ἱεροσύνης ὄνομα ἐστὶν ἢ τραπεζοφόρος, ὅσα αὐτῇ τε καὶ ἢ κοσμοσυνδέσμοι πάντα τῆς Ἀθηνῶν ἱερείας, αὐτῆς τε ὁ ἄριστος ἐν τῷ αὐτῷ λόγῳ δεδήλωσε καὶ Ἰστρος ἐν ἑκτῇ καὶ δεκάτῃ τῶν Ἀττικῶν συναγωγῶν. (17-32) FRAGMENTA INCERTAE SEDIS. (17) <i>Schol.</i> Aristoph. <i>Lysistr.</i> 642: Τῇ Ἐρωτῇ</p> <p style="text-align: right;">Prev Next</p> <p style="text-align: right;">Fragment 14 line 1 60</p>	<p>Text Structure</p> <p style="text-align: right;">Book 13 Kallipar paragraph 4 line 12 60</p> <p style="text-align: center;">Prev Next</p> <p>ATHENAEUS, Deipnosophistae. (0008.001) Μῆθειαν, Θεραΐες δὲ Ἐλένην ἀρπάζας ἔξις καὶ Ἀριώδην ἥρπασεν. Ἰστρος γόν ἐν τῇ τετρασερακιδεκάτῃ τῶν Ἀττικῶν (FHG I 420) καταλέγων τὰς τοῦ Θεραΐος γεννομένης γενναίας φησὶ τὰς μὲν αὐτῶν ἔξ ἔρωτος γεγεννησίου, τὰς δ' ἔξ ἀρπαγῆς, ἄλλας δ' ἔκ νομίμων γένων ἔξ ἀρπαγῆς μὲν Ἐλένην, Ἀριώδην, Ἰσπολίτην καὶ τὰς Κερκεϊνοῦ καὶ Σίνδος θυγατέρας, νομίμως δ' αὐτῶν γήμια Μεμβρόων τῶν Αἰώντων μητέρα. Ἡσιόδου δὲ φησὶ τῆς Ἐλένης καὶ ἐκ Τροφίλων ἥρπασεν Ἀναΐδου, μετὰ δὲ τῆν Ἰσπολίτην Φειδῶν ἴσθον. (5) Φιλύππος δ' ὁ Μεσσηνῶν οὐκ ἐπίστρος μὲν ἐστὶ τοῦ πολέμου γενναίως, ὅσπερ Ἀχιλεὺς ὁ ἰσὶ Ἀλεξάνδρῳ καταλιθεὶς, ὅς περὶ τὸν ὄλον πόλεμον τροφίλος ἐβρόναιον περιήρητο πολλοτάτος, ὡς ἰστροὶ Διοκλεῖος ἐν τρίτῳ περὶ τοῦ Ἑλλάδος βίου (FHG I 240): ὁ δὲ Φιλύππος αἰεὶ κατὰ πόλεμον ἐράσμι, ἐν ἔρωτι γόν ἐίστοι καὶ βροῦν οὐκ ἐβροῦνεν, ὡς φησὶ Στέφανος ἐν τῷ περὶ τοῦ βίου αὐτοῦ (ib. III 161), Ἀσδό- @ 1</p> <p style="text-align: right;">Prev Next</p> <p style="text-align: right;">Book 13 Kallipar paragraph 4 line 12 60</p>
---	--

Figure 2.48. TLG parallel browsing: Istros, FHG I, fr. 14 and Ath., *Deipn.* 13.4

to compare two editions of the same text. In this case the TLG uses differences between individual word forms, beta escapes, and punctuation in order to capture finer distinctions between texts than with n-grams. The TLG provides the current available list of texts with multiple editions with the goal of expanding it over time.

2.3.3 Translation Alignment

Text reuses and intertextual parallels can be analyzed and detected also across different languages. Figure 2.49 shows a passage of Livy (30.45), who explicitly refers to Polybius (16.23) about the presence of the king Syphax in the triumphal march of Scipio to Rome. Given that the Greek text of Polybius is preserved, it is possible to compare it with its Latin reuse and generate an alignment of the corresponding words.¹⁵⁴ Many other examples are also offered in the field of fragmentary literature, where original texts are lost. Figure 2.50 shows a passage of the *Astronomica* of the Latin polymath Hyginus (2.40) mentioning the lost Greek author Istros the Callimachean about Koronis (FGrHist 334 F 66 = BNJ 334 F 66). Given that we don't have the original text, it is not possible to check the accuracy of the reference of Hyginus, except for speculating about the words that could possibly derive from Istros' work, whose name is the only evidence in the passage of Hyginus.

Machine translation tools have been developed for “automatically producing in a target language the translation of a text in a source language.” These tools are devised for translations of everyday texts written in modern languages and not for literature or poetry. Results are still not really satisfactory because translation is a very difficult task that requires a profound knowledge and comprehension of the text that has to be translated, and because machines still need a lot of training.¹⁵⁵

154 Corresponding words are red in the figure. The page is available at <http://demo.fragmentarytexts.org/en/istros/digital-edition/digital-edition-exploring-text-re-uses-across-languages.html>, where it is also available an XML output of the alignment produced with the translation alignment editor of the *Alpheios* project.

155 For an introduction to the development of automatic machine translation since the Second World War, see Poibeau (2017). For an overview of the revolutionary effects that online translation services and crowdsourced translations are producing and for their implications for human languages, cultures and society, see Cronin (2013).

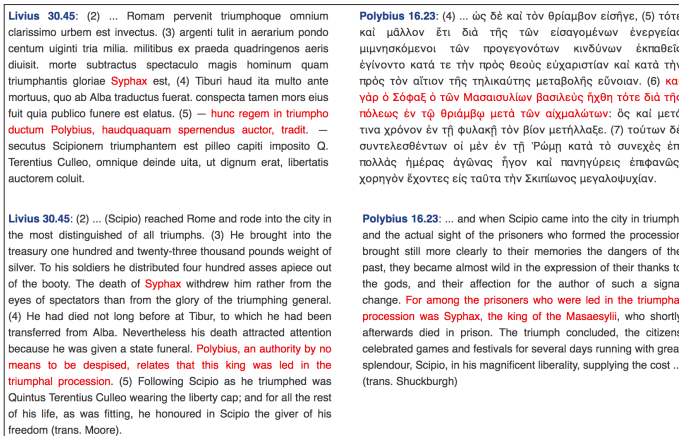


Figure 2.49. Text reuse of preserved texts across languages (Livy cites Polybius)

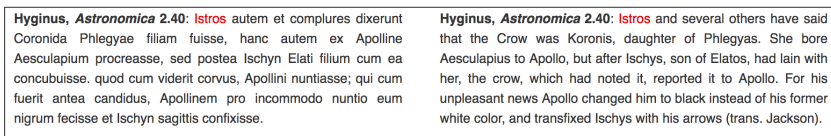


Figure 2.50. Text reuse of lost texts across languages (Hyginus cites Istros)

Translation of historical texts is an even more difficult task, because it is about texts produced in the past and problems of comprehension are much bigger than for modern and contemporary texts, and also because it is very difficult to agree on what we mean by translation and by *good* translation. This is one of the reasons why we still miss automatically generated translations of historical texts and experiments are at the very beginning.¹⁵⁶

Tesserae has been implementing a Greek-Latin search, which is available online for testing results, and is producing a *translation dictionary* for linking Greek lemmata to associated Latin terms.¹⁵⁷ A similar method has been used as part of the *Digital Fragmenta Historicorum Graecorum* (DFHG) project in order to align the Greek text of the fragments with their Latin translation provided by the editor Karl Müller. This case is different from those mentioned before, because it is not about two ancient languages, but about the translation of ancient Greek into 19th century scholarly Latin. The method was based on the use of data from the *Dynamic Lexicon*, which is a project of the *Perseus Digital Library* for creating automatic bilingual dictionaries of Greek-English and Latin-English, using source

156 Bamman/Crane (2009); Crane (2019).

157 <https://tesseraev3.caset.buffalo.edu/cross.php>

texts in Greek or Latin aligned with their English translations and using also morpho-syntactic data from Greek and Latin treebank.¹⁵⁸

Another method has been explored by combining data of the *Dynamic Lexicon* with data of the *Ancient Greek WordNet* (AGWN), which is a project of the Institute for Computational Linguistics “A. Zampolli” in Pisa for producing a lexico-semantic resource mapped on Princeton WordNet 3.0.¹⁵⁹ Both approaches have produced translation pairs that are not completely correct because they still need accuracy improvement and manual validation, but they have shown that they can be integrated in order to improve performances. One of the problems that clearly emerges from these experiments is the lack of training data and the necessity of producing it in order to expand dictionaries of historical languages that can be used for increasing machine translation results.

Having this goal in mind, translation alignment tools have been developed in the last years, such as the translation alignment editor of the *Alpheios* project and *Ugarit iAligner* of the *Open Philology* project at the University of Leipzig. The first editor is part of a set of reading and learning environments developed by *Alpheios* to support worldwide study of classical languages and literatures. The tool has been experimentally used for research and teaching initiatives and as part of the *Perseids* project at Tufts University, allowing users to manually align two texts in two different languages including Greek, Latin, Aramaic, Egyptian, Persian, and Syriac.¹⁶⁰ *Ugarit iAligner* is a tool that performs automatic syntax-based intra-language alignment and automatic alignment of different versions of a text using a modified version of the *Needleman-Wunsch Algorithm*. It includes an editor for manual alignment of up to three languages.¹⁶¹

158 Yousef/Berti (2015). First experiments on this method were carried out as part of a MSc dissertation written by Yousef (2015) under my supervision at the University of Leipzig. On the *Dynamic Lexicon*, see Bamman/Crane (2008a).

159 Berti/Bizzoni et al. (2016). On AGWN, see Bizzoni/Boschetti et al. (2014) and Boschetti/Del Gratta et al. (2016).

160 See, for example, Almas/Beaulieu (2016), Mernitz (2016), and Almas (2017). Teaching experiments have been also performed as part of the *Sunoikisis Digital Classics* program: Berti/Crane et al. (2015); Berti (2016b); Berti (2017b); Berti (2017c).

161 See <http://ialigner.com>, Yousef/Palladino (2017), and Yousef (2020). *Ugarit iAligner* is currently used at the University of Leipzig as part of the *Open Persian* project for aligning Persian poetry with modern languages and as part of the *Digital Rosetta Stone* project (section 4.6) for aligning the Hieroglyphic, Demotic, and Greek scripts of the inscription. See Berti/Jushaninowa et al. (2016), Foradi/Crane (2017), Berti/Naether/Amin et al. (2018b), Berti/Naether/Amin et al. (2018a), Berti/Naether/Bozia (2018).