

# Ancient Greek Historians in the Digital Age

Monica Berti

Universität Leipzig, Leipzig, Sachsen

ORCID: 0000-0003-2779-4229

**Abstract:** This paper presents results of ongoing digital projects on ancient Greek historians. The research question is the analysis of the language used by ancient sources to refer to historians and cite their works with a particular reference to lost historians (the so-called fragmentary authors). If a lot of scholarship has been devoted to collect fragments of many different genres and try to reconstruct the texts from which they were taken, less effort has been spent on collecting data pertaining to the language used by ancient authors to refer to them and their works. The paper discusses the use of Computational Linguistics techniques and Named Entity Recognition to extract and annotate information about ancient Greek historians and their works from the sources where they are preserved. Moreover, the paper describes a new catalog of ancient Greek authors and works based on the extraction and annotation of references to them in ancient sources.

Keywords: Digital Philology, Digital Classics, Text Reuse, Linguistic Annotation, Named Entity Recognition

## 1. Introduction\*

Within the domain of Digital History, the digitization and analysis of primary sources play a crucial role. As far as ancient Greek and Latin historical works are concerned, digital libraries have increased accessibility to primary and secondary sources.<sup>1</sup> Among the requests of the community of digital historians, there are certainly the needs for more digitized texts, for more digital editions of the same source, and for more structured data to extract information.<sup>2</sup>

When looking for primary sources in ancient Greek and Latin, we have to consider that today digital projects offer many different levels of accessibility, in terms of open access and data usability. One of the goals of the *Open Greek and Latin* project developed at Leipzig University is to offer access to at least one open access edition with structured data in XML format according to Linked Open Data (LOD) principles.<sup>3</sup> Access is provided by the *Perseus Catalog*, by its collection of sources in the *Perseus Digital Library*, and by the *Scaife Viewer*.<sup>4</sup>

---

\* This paper has been written as part of a project supported by the Deutsche Forschungsgemeinschaft (nr. 434173983).

<sup>1</sup> Gregory Crane et al., *Cyberinfrastructure for Classical Philology*, in: DHQ 3.1 (2009); Alison Babeu, *Rome Wasn't Digitized in a Day. Building a Cyberinfrastructure for Digital Classicists*, Washington DC 2011.

<sup>2</sup> Jonathan Blaney et al., *Doing Digital History. A Beginner's Guide to Working with Texts as Data*, Manchester 2021; Adam Crymble, *Technology and the Historian. Transformations in the Digital Age*, Urbana 2021; Karoline Döring et al., *Digital History. Konzepte, Methoden und Kritiken Digitaler Geschichtswissenschaft*, Berlin 2022; Shawn Graham et al., *Exploring Big Historical Data. The Historian's Macroscopic*, 2nd edition, London 2022 (1st edition 2015). Salvatore Spina, *The Digital Age of Historians*, in: *AIDAinformazioni* (2021), pp. 103–120.

<sup>3</sup> On this project see the papers edited in Monica Berti (ed.), *Digital Classical Philology. Ancient Greek and Latin in the Digital Revolution*, Berlin 2019.

<sup>4</sup> *Open Greek and Latin* (OGL): <https://opengreekanlatin.org/>; *Perseus Catalog*: <https://catalog.perseus.org/>; *PerseusDL*: <https://github.com/PerseusDL>; *Scaife Viewer*: <https://scaife.perseus.org/>. Accessed May 2023.

This paper presents experiments and results of ongoing projects about ancient Greek historiography with a focus on the so-called fragmentary historians, by which scholarship means authors whose works are now lost in their original form and preserved through quotations and reuses in later sources.<sup>5</sup>

Within this domain, I have been developing a new model for digital editions of historical fragmentary texts, addressing concrete problems that scholars have to deal with when analyzing textual fragments of historians in a digital environment.<sup>6</sup> In order to preserve characteristics of printed editions of fragmentary texts that are now digitized and in order to find a new digital model that is not dependent on the printed one, I have implemented two different projects: The first is the *Digital Fragmenta Historicorum Graecorum*, which is the digital enriched version of the *Fragmenta Historicorum Graecorum* published by Karl Müller in the 19th century; the second project is the *Digital Athenaeus*, whose goal is the extraction of data from the *Deipnosophists* of Athenaeus of Naucratis to provide an inventory of authors and works cited in the text and to implement a data model for identifying, analyzing, and citing uniquely its instances of text reuse.

In the next two paragraphs I will summarize the main characteristics of these two projects. In the third paragraph I will present recent developments in the use of web annotations and Named Entity Recognition for extracting bibliographic references to ancient historians.

## 2. The Digital Fragmenta Historicorum Graecorum (DFHG)

The *Digital Fragmenta Historicorum Graecorum* (DFHG) is a project born from the digitization of the five volumes of the *Fragmenta Historicorum Graecorum* (FHG) printed by the publisher Ambroise Firmin Didot between 1841 and 1873 to produce the first big collection of Greek fragmentary historians. The project is available online at <https://www.dfhg-project.org/> and has also been extensively presented and described in recent publications.<sup>7</sup>

The reason for this project is twofold: 1) Transfer the collection of the FHG from a printed to a digital environment in order to preserve it, and 2) extract data from the FHG to expand it and use it for new questions about fragmentary historiography. These reasons are very important today: If one of our main goals is to produce newborn digital editions and data, we also have the responsibility to preserve sources edited in the past and the rich editorial work accomplished by scholars when collecting and interpreting ancient texts.

The collection of the FHG has been superseded by the *Fragmente der griechischen Historiker* (FGrHist) edited by Felix Jacoby, which is still ongoing in the *continuatio* of the *Brill's New Jacoby* (BNJ) of the *Jacoby Online* project.<sup>8</sup> The reason for working on the older edition of the FHG depends not only on the fact that this edition is now out of copyright and big enough for computational analyses, but also on its importance for understanding more recent editions of Greek historical fragments including the FGrHist. Moreover, fragmentary historiography offers many examples of reuse of prose texts, the complexities of which are shared by other genres of fragmentary literature and are useful today in the field of text reuse applied to historical languages.<sup>9</sup>

---

<sup>5</sup> Glenn W. Most (ed.), *Collecting Fragments. Fragmente sammeln*, Göttingen 1997.

<sup>6</sup> Monica Berti, *Historical Fragmentary Texts in the Digital Age*, in: Berti, *Digital Classical Philology*, pp. 257–276; Monica Berti, *Digital Editions of Historical Fragmentary Texts*, Heidelberg 2021. See also Monica Berti / Briget Almas / Gregory R. Crane, *The Leipzig Open Fragmentary Texts Series (LOFTS)*, in: Neil Bernstein / Neil Coffee (eds.), *Digital Methods and Classical Studies*, DHQ Themed Issue 10.2, 2016; Monica Berti, *Annotating Text Reuse within the Context: the Leipzig Open Fragmentary Texts Series (LOFTS)*, in: Ute Tischer / Ursula Gärtner / Alexandra Forst (Hrsg.), *Text, Kontext, Kontextualisierung. Moderne Kontextkonzepte und antike Literatur*, Hildesheim 2018, pp. 223–234.

<sup>7</sup> Berti, *Historical Fragmentary Texts*, pp. 261–269; Berti, *Digital Editions*, pp. 127–303.

<sup>8</sup> See <https://scholarlyeditions.brill.com/bnjo/> (accessed May 2023).

<sup>9</sup> Berti, *Digital Editions*, pp. 134–135. For differences and connections between the DFHG and the *Jacoby Online* see the Müller-Jacoby Table of Concordance: <https://www.dfhg-project.org/Mueller-Jacoby-Concordance/>. On digital data and digital editions of ancient Greek fragmentary authors and works with a focus on historical texts, see <https://www.dfhg-project.org/Fragmentary-Texts/>. Resources accessed in May 2023.

The collection of the FHG includes 636 fragmentary historians dated from the 6th century BC through the 7th century CE and the DFHG project has produced their complete digitization.<sup>10</sup> This is not the place for a description of the project, but, for the purpose of this publication the following are the most valuable data from the digitization of the FHG: 1) References to witness sources that preserve fragments of the historians and that have been fully collected and converted into a digital catalog with maps, charts, and timelines.<sup>11</sup> The Witnesses Catalog additionally offers automatic text reuse detection of fragmentary works in their witnesses.<sup>12</sup> 2) The language of the fragments which can be analyzed thanks to the integration of internal and external resources such as textual collections, authority lists, indices, dictionaries, lexica and gazetteers.<sup>13</sup> 3) Translations of fragments into Latin that have been used for experimenting with translation alignment for historical languages.<sup>14</sup>

### 3. The Digital Athenaeus

The *Digital Athenaeus* project is not a new critical edition of the *Deipnosophists* of Athenaeus of Naucratis, but the critical selection, extraction, and analysis of data from the text of the Naucratis (ed. Kaibel).<sup>15</sup> The project is available at <https://www.digitalatheneus.org/> with descriptions in recent publications.<sup>16</sup>

The main goal of the project is to analyze bibliographic references provided by Athenaeus to cite authors and works. The *Deipnosophists* preserves a very rich collection of references to 895 authors and to 1,976 works, most of which are now lost. If scholarship has already addressed questions about sources and bibliographic tools used by Athenaeus for his work, a complete analysis of the language of the author to refer to all these sources is still lacking.<sup>17</sup> Therefore, the *Digital Athenaeus* project has been annotating the text of the *Deipnosophists* and extracting data about authors and works in the context in which they are preserved. This work offers also a new model where fragmentary texts become contextualized annotations of elements about authors and works and not extracted and decontextualized chunks of texts, as it happens with fragments in a printed environment.<sup>18</sup>

The reference text of the *Digital Athenaeus* is the Teubner edition of Georg Kaibel and various tools have been created before producing annotations. In particular, issues related to text citations were addressed, since the *Deipnosophists*' passages can be cited according to the Casaubon and Kaibel systems, but only the Kaibel one – by books and paragraphs – fits canonical citations in a digital environment.<sup>19</sup> Also, experiments for drawing data from traditional indexes to align it with the Greek

---

<sup>10</sup> The catalog of FHG authors is available at <https://www.dfhg-project.org/Fragmentary-Authors-Catalog/> (accessed May 2023). A map and a chart allow to visualize their places of origin and intellectual activity and their distribution in the collection: See Berti, Digital Editions, pp. 194–209.

<sup>11</sup> See <https://www.dfhg-project.org/Witnesses-Catalog/> (accessed May 2023).

<sup>12</sup> Berti, Digital Editions, pp. 245–259.

<sup>13</sup> Berti, Digital Editions, pp. 165–169.

<sup>14</sup> See <https://ugarit.ialigner.com/dfhg/> (accessed May 2023) and Tariq Yousef et al., Automatic Translation Alignment for Ancient Greek and Latin, in: Rachele Sprugnoli / Marco Passarotti (eds.), Proceedings of the LREC 2022 Second Workshop on Language Technologies for Historical and Ancient Languages LT4HALA 2022, European Language Resources Association (ELRA) 2022, pp. 101–107.

<sup>15</sup> A new traditional print edition of the *Deipnosophists* is now almost completed by S. Douglas Olson for De Gruyter, but it is not yet open access: see <https://www.degruyter.com/serial/deipn-b/html> (accessed May 2023). For a review of it, see Monica Berti, Review of S. Douglas Olson, ed., Athenaeus Naucratis. *Deipnosophistae*. Vol. III.A-B, LIBRI VIII-XI. Berlin-Boston: De Gruyter 2020. S. Douglas Olson, ed., Athenaeus Naucratis. *Deipnosophistae*. Vol. IV.A-B, LIBRI XII-XV. Berlin-Boston: De Gruyter, 2019, in: *Exemplaria Classica* 26 (2022), pp. 311–313.

<sup>16</sup> Berti, *Historical Fragmentary Texts*, pp. 270–274; Berti, Digital Editions, pp. 305–414.

<sup>17</sup> David Braund / John Wilkins (eds.), *Athenaeus and His World. Reading Greek Culture in the Roman Empire*, Exeter 2000; Christian Jacob, *The Web of Athenaeus*, Washington DC 2013.

<sup>18</sup> Berti, *Annotating Text Reuse within the Context*, pp. 223–234.

<sup>19</sup> Monica Berti et al., Documenting Homeric Text-Reuse in the *Deipnosophistae* of Athenaeus of Naucratis, in: Gabriel Bodard / Yanne Broux / Ségolène Tarte (eds.), *Digital Approaches and the Ancient World*, BICS Themed Issue 59.2, London 2016, pp. 121–139; Berti, Digital Editions, pp. 312–322. An online converter and

text of the *Deipnosophists* and for text reuse detection have been done to deepen the analysis of the text and provide inventories of authors and works cited by Athenaeus.<sup>20</sup> Results obtained through this work are foundational for what is described in the following paragraph, which is about the creation of catalogs of authors and works based on linguistic annotations in ancient sources.

#### 4. Annotations and the Catalog

A significant component of the DFHG and the *Digital Athenaeus* is the possibility to analyze the language used by extant sources to refer to lost historians and cite their works. If a lot of scholarship has been devoted to collect fragments of many different literary genres and try to reconstruct the texts from which they were taken, less effort has been spent on collecting data pertaining to the language used by ancient authors to refer to other authors and works.<sup>21</sup>

This research question has brought me to develop a project for creating a catalog of ancient Greek authors and works based on the extraction and annotation of references to them in ancient sources. Given that most elements of these references are represented by proper names, I have been making use of Linguistic Annotation and Named Entity Recognition techniques to extract and annotate linguistic data related to names of Greek authors and to descriptions of their works.<sup>22</sup>

Three main reasons lay behind this project: 1) The lack of digital data in ancient Greek. A significant number of linguistic annotations is already available, but Named Entities (i.e., proper names) are still largely missing and, within this category, authors and works are under-represented. Moreover, if names of authors and titles of works have entered traditional indexes and catalogs in Latin or in modern languages, the original linguistic information related to them is lost and hidden in ancient sources that have to be separately consulted to analyze their textual characteristics.<sup>23</sup> 2) The second reason is that this work benefits from Computational Linguistics techniques, but it is also a contribution to the creation of more linguistic data about the ancient world and therefore to the implementation of computational technologies based on challenging cases like text reuses in ancient Greek. 3) Finally, the catalog is not only limited to fragmentary historians, but to all authors of ancient Greek literature, enabling us to analyze references to historians in a bigger context for addressing well known issues about modern categories used to classify ancient literary texts.<sup>24</sup>

In its first phase the catalog is based on data extracted from the *Deipnosophists* of Athenaeus and from the *Lexicon of the Ten Orators* of Valerius Harpocration.<sup>25</sup> These two texts are different, being the first a scholarly miscellany and the second a lexicon mainly dealing with Athenian legal customs. In spite of that, both authors share many citations deriving from the same cultural and librarian environment of Alexandria in Egypt.<sup>26</sup> Further steps will include the lexicon of the *Suda* and other

---

related tools have been created to use both systems: [https://www.digitalathenaeus.org/tools/Casaubon-Kaibel\\_converter/](https://www.digitalathenaeus.org/tools/Casaubon-Kaibel_converter/) (accessed May 2023).

<sup>20</sup> Berti, Digital Editions, pp. 322–398.

<sup>21</sup> Berti, Digital Editions, pp. 7–39.

<sup>22</sup> Monica Berti, Named Entity Annotation for Ancient Greek with INCEpTION, in: Kiril Simov / Maria Eskevich (eds.), Proceedings of CLARIN Annual Conference 2019, Leipzig 2019, pp. 1–4.

<sup>23</sup> For example, the reference to the lost treatise by Theopompus on the money plundered from Delphi at Ath., *Deipn.* 12.532d (περὶ τῶν ἐκ Δελφῶν συληθέντων Χρημάτων) is usually hidden behind the label *fragmentum*.

<sup>24</sup> Monica Berti, Digital Canons and Catalogs of Fragmentary Literature, in: Theophanis Tsiampokalos et al. (ed.), Fragmente einer fragmentierten Welt. Gespräche zur Problematisierung eines traditionellen Begriffes in der gegenwärtigen klassisch-philologischen Forschung, Berlin forthcoming 2023; Monica Berti, Digital Practice for Studying Indirect Transmission of Classical Authors and Works, in: Virginia Mastellari / Federico Favi (eds.), Treasuries of Literature: Anthologies, Lexica, Scholia, and the Indirect Tradition of Classical Texts in the Greek World, Berlin forthcoming 2023.

<sup>25</sup> Data from the work of Harpocration (ed. Dindorf) is part of the *Linked Ancient Greek and Latin* (LGL) project (<https://www.lagl.org/>): See <https://www.lagl.org/tools/harpocration/> (accessed May 2023).

<sup>26</sup> Giuseppe Zecchini, Harpocration and Athenaeus. Historiographical Relationships, in: Braund / Wilkins, Athenaeus and His World, pp. 153–160. Based on extraction and annotation of data, the *Lexicon* of Harpocration has citations of 192 authors and 156 works.

lexical and encyclopedic texts, which are significant for the preservation of ancient bibliographic data including ancient Greek historians.

Hellanicus - urn:cts:greekLit:tlg0539
Submit

**Hellanicus** ( **Hellanicus of Mytilene** )

**Historicus**

**Lesbius Lesbos**

urn:cts:greekLit:tlg0539

**Works**

urn:cts:greekLit:tlg0539.ath001 - Aegyptiaca  
 urn:cts:greekLit:tlg0539.ath002 - Journey Inland to Ammon's Shrine  
 urn:cts:greekLit:tlg0539.ath003 - Deucalionia  
 urn:cts:greekLit:tlg0539.ath004 - Names of Ethnic Groups  
 urn:cts:greekLit:tlg0539.ath005 - Victories at the Carneia  
 urn:cts:greekLit:tlg0539.ath006 - Foundations  
 urn:cts:greekLit:tlg0539.ath007 - Phoronis

1 . 61 [ [Ἑλλάνικος](#) δὲ φησιν ἐν τῇ Πλινθίνῃ πόλει Αἰγύπτου πρώτη εὐρεθῆναι τὴν ἀμπελον. ]  
 9 . 80 [ τὸν δὲ τῷ κερνήτῳ ῥάναντα παῖδα διδόντα κατὰ χειρὸς Ἡρακλεῖ ὕδωρ, ὃν ἀπέκτεινεν ὁ Ἡρακλεῖς κονδύλω, [Ἑλλάνικος](#) μὲν ἐν ταῖς [ἱστορίαις](#) Ἀρχίαν φησὶ καλεῖσθαι. ]  
 10 . 9 [ [Ἑλλάνικος](#) δ' ἐν α' [Ἀευκαλιωνείας](#) Ἐρυσίχθονά φησι τὸν Μυρμιδόνοσ, ὅτι ἦν ἀπληστος βορᾶς, Αἴθωνα κληθῆναι. ]  
 10 . 67 [ [Ἑλλάνικος](#) δ' ἐν [Κτίσεσι](#) καὶ ἐκ ῥιζῶν, φησὶ κατασκευάζεται τὸ βρῦτον γράφων ὧδε: ]  
 11 . 6 [ οἶδα δὲ καὶ [Ἑλλάνικον](#) ἐν [Ἐθνῶν Ὀνομασίαις](#) λέγοντα ὅτι Λιβύων τῶν Νομάδιων τινὲς οὐδὲν ἄλλο κέκτηνται ἢ κύλικα καὶ μάχαιραν καὶ ὑδρίαν, καὶ ὅτι οἰκίας ἔχουσιν ἐξ ἀνθρῆκου πεποιημένας μικράς ὅσον σκιάς ἔνεκα, ἃς καὶ περιφέρουσιν ὅπου ἂν πορεύωνται. ]  
 11 . 40 [ [Ἑλλάνικος](#) ἐν [Αἰγυπτιακοῖς](#) οὕτως γράφει· ]  
 14 . 37 [ τὰ Κάρνεια πρῶτος πάντων [Τέρπανδρος](#) νικᾷ, ὡς [Ἑλλάνικος](#) ἱστορεῖ ἐν τε τοῖς ἐμμέτροις [Καρνεονίκαις](#) κἂν τοῖς καταλογάδην. ]  
 14 . 66 [ φοῖνικα δὲ τὸν καρπὸν καὶ [Ἑλλάνικος](#) κέκληκεν ἐν τῇ [εἰς Ἀμμωνος Ἀναβάσει](#), εἰ γνήσιον τὸ σύγγραμμα, καὶ [Φόρμος](#) ὁ κωμικὸς ἐν [Ἀταλάνταις](#). ]  
 15 . 25 [ περὶ δὲ τῶν ἐν Αἰγύπτῳ αἰεὶ ἀνθούτων στεφάνων [Ἑλλάνικος](#) ἐν τοῖς [Αἰγυπτιακοῖς](#) οὕτως γράφει· ]  
 15 . 25 [ ὁ δὲ προσηρμένος [Ἑλλάνικος](#) καὶ Ἀμασιν Αἰγύπτου βασιλεύσας, ἰδιώτην ὄντα καὶ τῶν τυχόντων κατὰ τὸν πρῶτον βίον, διὰ στεφάνου δωρεάν, ὃν ἐπεμψεν ἀνθέων πλεξάμενος τῇ ὥρᾳ περικαλλεστάτων γενέθλια ἐπιτελοῦντι Πατάρμιδι τῷ τῆς Αἰγύπτου τότε βασιλεύοντι. ]

Fig. 1: Hellanicus of Mytilene in the Catalog of the Digital Athenaeus.

Figure 1 shows the entry of the fragmentary historian Hellanicus of Lesbos in the first web version of the Catalog, which is currently available at <https://www.digitalathenaeus.org/tools/Catalog/>. This entry and the data behind it reveal the language used by Athenaeus to refer to seven works of the Lesbian historian that are now lost. These expressions are usually hidden behind the label *fragmenta* in printed editions and catalogs, including those produced in the first generations of digital libraries, like the *TLG Canon* and the *Perseus Catalog*.<sup>27</sup>

In order to generate these annotations, the text of the *Deipnosophists* (ed. Kaibel) has been tokenized and lemmatized, and lemmata have been used to query external authority lists for a first level of disambiguation of personal and place names.<sup>28</sup> Author names have been identified in the group of personal names and homonyms have been disambiguated.<sup>29</sup> Pre-annotated data has been uploaded to the web-based platform INCEPTION for Named Entity relation and linking, and for completing the annotation of titles/descriptions of works.<sup>30</sup>

Figure 2 shows the example of the annotations of paragraph 37 in book 14 of the *Deipnosophists* of Athenaeus, where we find a citation of the work of Hellanicus on the victories at the Carneia (*Καρνεονίκαις*). Authors and works are identified with CTS URNs according to the CITE Architecture, which ingests numbers of the *TLG Canon* to identify authors and works.<sup>31</sup> In the *TLG Canon* Hellanicus

<sup>27</sup> Cf. Hellanicus in the *Perseus Catalog*: <https://catalog.perseus.org/catalog/urn:cite:perseus:author.701> (accessed May 2023). On the *TLG Canon* see <https://stephanus.tlg.uci.edu/> (accessed May 2023) with Maria Pantelia, *Thesaurus Linguae Graecae: A Bibliographic Guide to the Canon of Greek Authors and Works*, Berkeley 2022. On the *Perseus Catalog*, see Alison Babeu, *The Perseus Catalog: of FRBR, Finding Aids, Linked Data, and Open Greek and Latin*, in: Berti, *Digital Classical Philology*, pp. 53–72.

<sup>28</sup> For a detailed description see Berti, *Named Entity Annotation for Ancient Greek*.

<sup>29</sup> For example to distinguish three different authors named Aristophanes: the comic poet (tlg0019), the Boeotian historian (tlg1196), and the Byzantine grammarian (tlg0644).

<sup>30</sup> Berti, *Named Entity Annotation for Ancient Greek*. For further information see the bibliography at note 24. On the web-based platform INCEPTION see <https://inception-project.github.io/> (accessed May 2023) with bibliography.

<sup>31</sup> On the CITE Architecture and the *Canonical Text Services*, see Christopher W. Blackwell / Neel Smith, *The CITE Architecture: a Conceptual and Practical Overview*, in: Berti, *Digital Classical Philology*, pp. 73–94.

Lesbius is tlg0539 and in the annotated Catalog the occurrences of his name are identified with urn:cts:greekLit:tlg0539.<sup>32</sup> Metadata of the TLG is also included in the annotated Catalog of the *Deipnosophists*. In this case we have the generic epithet *historicus* and the geographic epithet *Lesbius*, whose form Lesbos is used to track the place name in the *Pleiades* gazetteer for a future visualization on a map.<sup>33</sup>

The work of Hellanicus cited in the passage of the *Deipnosophists* is identified according to the CITE Architecture. The form *Καρνεονίκαις* is lemmatized, annotated as Named Entity (ORGderiv because it derives from the name of the festival of the *Κάρνεια*), and identified with the CTS URN urn:cts:greekLit:tlg0539.ath005. This string means that this is the fifth work of Hellanicus cited in the *Deipnosophists*. The element *ath* in the CTS URN refers to the fact that this is the form of the title of the work read by Athenaeus or found by Athenaeus in one of his sources. The TLG *Canon* has three identifiers for the works of Hellanicus: tlg0539.001 (*Testimonia*), tlg0539.002 (*Fragmenta*), and tlg0539.003 (*Fragmenta*). These identifiers refer to modern editions of the testimonies and the fragments of Hellanicus by Felix Jacoby and Hans Joachim Mette, but not to Hellanicus' lost works. An annotated catalog allows to point to the actual reference to the work of Hellanicus as it was cited in ancient sources and identify it in a persistent and unique way.

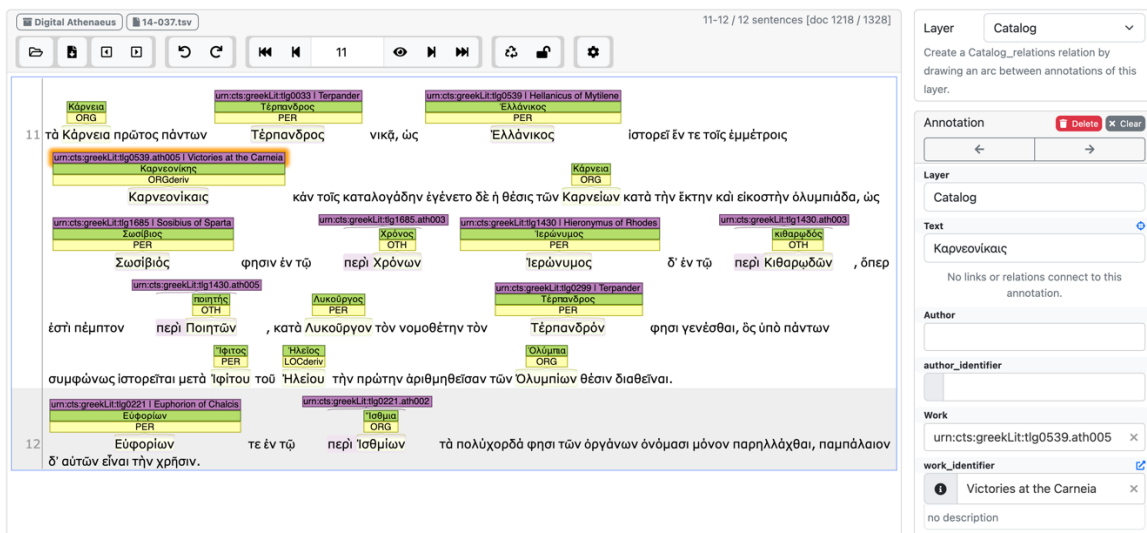


Fig. 2: Annotation of Ath. *Deipn.* 14.37 (635e) in INCEpTION.

The context itself is also interesting. In this case, only the form *Καρνεονίκαις* has been annotated, even if in the text we have other linguistic elements related to it because Athenaeus informs us that the work existed both in a metrical and a prose version (ὡς Ἑλλάνικος ἱστορεῖ ἐν τε τοῖς ἐμμέτροις *Καρνεονίκαις* κἀν τοῖς καταλογάδην).<sup>34</sup> In this phase, the Catalog includes only annotations of explicit forms of names of authors and work titles basically corresponding to their proper names, but of course further annotations can be added in the future with more identifiers and metadata.

More example and data could be added, but this is beyond the scope and extension of this paper. As for now, annotations of names of authors and descriptions/titles of works highlight a language which is usually hidden in traditional print resources. The work behind the catalog and the contextual language of these annotations allow to pose new questions about the language used by ancient sources to cite historians, about standards and formats for their linguistic annotation, about issues related to Named Entity Recognition and to annotation for ancient Greek and in particular for historiography, about the

<sup>32</sup> For the entry in the *Perseus Catalog*, see note 27.

<sup>33</sup> On the epithetes of the TLG *Canon*, see Pantelia, *Thesaurus Linguae Graecae*, pp. xxiii–xxviii. See also Berti, *Digital Editions*, pp. 18–26. The canonical URI of Lesbos in the *Pleiades* gazetteer is <https://pleiades.stoa.org/places/550696> (accessed May 2023). For the use of *Pleiades* URIs to visualize maps of authors and their witnesses in the DFHG project, see Berti, *Digital Editions*, p. 207–209 and 234–235.

<sup>34</sup> Cf. FGrHist (BNJ) 4 F 85a.

relation between linguistic data and external authority lists, about data export and interchange according to Linked Open Data (LOD) principles, and finally about visualizations of annotations in their context.

## 5. Literaturverzeichnis

- Babeu, Alison, *Rome Wasn't Digitized in a Day. Building a Cyberinfrastructure for Digital Classicists*, Washington DC 2011.
- Babeu, Alison, The Perseus Catalog: of FRBR, Finding Aids, Linked Data, and Open Greek and Latin, in: Monica Berti (ed.), *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*, Berlin: De Gruyter 2019, pp. 53–72. DOI: 10.1515/9783110599572-005.
- Berti, Monica (ed.), *Digital Classical Philology. Ancient Greek and Latin in the Digital Revolution.*, Berlin 2019. DOI: 10.1515/9783110599572.
- Berti, Monica / Almas, Bridget / Crane, Gregory R., *The Leipzig Open Fragmentary Texts Series (LOFTS)*, in: Neil Bernstein / Neil Coffee (eds.), *Digital Methods and Classical Studies*, DHQ Themed Issue 10.2, 2016.
- Berti, Monica et al., Documenting Homeric Text-Reuse in the *Deipnosophistae* of Athenaeus of Naucratis, in: Gabriel Bodard / Yanne Broux / Ségolène Tarte (eds.), *Digital Approaches and the Ancient World*, BICS Themed Issue 59.2, London 2016, pp. 121–139. DOI: 10.1111/j.2041-5370.2016.12042.x.
- Berti, Monica, Annotating Text Reuse within the Context: the *Leipzig Open Fragmentary Texts Series (LOFTS)*, in: Ute Tischer / Ursula Gärtner / Alexandra Forst (eds.), *Text, Kontext, Kontextualisierung. Moderne Kontextkonzepte und antike Literatur*, Hildesheim 2018, pp. 223–234.
- Berti, Monica, Digital Canons and Catalogs of Fragmentary Literature, in: Theophanis Tsiampokalos et al. (eds.), *Fragmente einer fragmentierten Welt. Gespräche zur Problematisierung eines traditionellen Begriffes in der gegenwärtigen klassisch-philologischen Forschung*, Berlin forthcoming 2023.
- Berti, Monica, *Digital Editions of Historical Fragmentary Texts*, Heidelberg 2021. DOI: 10.11588/propylaeum.898.
- Berti, Monica, Digital Practice for Studying Indirect Transmission of Classical Authors and Works, in: Virginia Mastellari / Federico Favi (eds.), *Treasuries of Literature: Anthologies, Lexica, Scholia, and the Indirect Tradition of Classical Texts in the Greek World*, Berlin forthcoming 2023.
- Berti, Monica, Historical Fragmentary Texts in the Digital Age, in: Monica Berti (ed.), *Digital Classical Philology. Ancient Greek and Latin in the Digital Revolution*, Berlin 2019, pp. 257–276. DOI: 10.1515/9783110599572-015.
- Berti, Monica, Named Entity Annotation for Ancient Greek with INCEpTION, in: Kiril Simov / Maria Eskevich (eds.), *Proceedings of CLARIN Annual Conference 2019, Leipzig 2019*, pp. 1–4.
- Berti, Monica, Review of S. Douglas Olson, ed., *Athenaeus Naucratis. Deipnosophistae. Vol. III.A-B, LIBRI VIII-XI*. Berlin-Boston: De Gruyter 2020. S. Douglas Olson, ed., *Athenaeus Naucratis. Deipnosophistae. Vol. IV.A-B, LIBRI XII-XV*. Berlin-Boston: De Gruyter, 2019, in: *Exemplaria Classica* 26 (2022), pp. 311–313. DOI: 10.33776/ec.v26.
- Blackwell, Christopher W. / Smith, Neel, The CITE Architecture: a Conceptual and Practical Overview, in: Monica Berti (ed.), *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*, Berlin 2019, pp. 73–94. DOI: 10.1515/9783110599572-006.
- Blaney, Jonathan et al., *Doing Digital History. A Beginner's Guide to Working with Texts as Data*. Manchester 2021.
- Braund, David / Wilkins, John (eds.), *Athenaeus and His World. Reading Greek Culture in the Roman Empire*, Exeter 2000.

- Crane, Gregory et al., Cyberinfrastructure for Classical Philology, in: DHQ 3.1 (2009).
- Crymble, Adam, Technology and the Historian. Transformations in the Digital Age. Urbana 2021.
- Döring, Karoline et al. (eds.), Digital History: Konzepte, Methoden und Kritiken Digitaler Geschichtswissenschaft, Berlin 2022. DOI: 10.1515/9783110757101.
- Graham, Shawn et al., Exploring Big Historical Data. The Historian's Macroscope, 2nd edition London 2022 (1st edition 2015). DOI: 10.1142/12435.
- Jacob, Christian, The Web of Athenaeus, Washington DC 2013.
- Most, Glenn W. (ed.), Collecting Fragments. Fragmente sammeln, Göttingen 1997.
- Pantelia, Maria, Thesaurus Linguae Graecae: A Bibliographic Guide to the Canon of Greek Authors and Works, Berkeley 2022.
- Spina, Salvatore. The Digital Age of Historians, in: AIDAinformazioni (2021), pp. 103–120. DOI: 10.57574/5965090006.
- Yousef, Tariq et al., Automatic Translation Alignment for Ancient Greek and Latin, in: Rachele Sprugnoli / Marco Passarotti (eds.), Proceedings of the LREC 2022 Second Workshop on Language Technologies for Historical and Ancient Languages LT4HALA 2022, European Language Resources Association (ELRA) 2022.
- Zecchini, Giuseppe, Harpocraton and Athenaeus. Historiographical Relationships, in: David Braund / John Wilkins (eds.), Athenaeus and His World. Reading Greek Culture in the Roman Empire, Exeter 2000, pp. 153–160.