

Nomina Omina: Ancient Greek and Latin Proper Names in the Age of Artificial Intelligence – Introduction

Monica Berti

Abstract: This paper is the introduction to the volume *Nomina Omina. Detecting and Preserving Ancient Greek and Latin Proper Names in the Age of Artificial Intelligence*, which collects the proceedings of a workshop held at Leipzig University in June 2024, thanks to the support of the German Research Foundation (Deutsche Forschungsgemeinschaft). This introduction outlines the goals and contributions of the workshop and the volume.

Nomina Omina – The Workshop

This volume is a collection of papers published as proceedings of the international workshop entitled *Nomina Omina. Detecting and Preserving Ancient Greek and Latin Proper Names in the Age of Artificial Intelligence*.¹ The workshop was held at Leipzig University from June 27 to 29, 2024, thanks to the generous support of the German Research Foundation (Deutsche Forschungsgemeinschaft – DFG) and as part of the project *Text-based Extraction, Analysis, and Annotation of Ancient Greek References to Authors and Works*.²

The goal of the workshop was to discuss the current state of data and research concerning proper names in ancient Greek and Latin, with a focus on the computational extraction of personal names, geographic names, and proper names related to authors and works cited in ancient sources. This kind of data is essential to studies in disciplines such as linguistics, philology, and historiography, and, in fact, traditional scholarship has consistently published repertoires, lexica, and indices to collect it. The advancement of Digital Classics is showing the importance of this kind of data not only for traditional purposes, but also for computational ones, given that proper names function as anchors in the text to structure unstructured data.³ This is one of the reasons why the so-called Named Entity Recognition (NER) technique is gaining growing interest in the Classics community, with projects devoted to extracting and annotating named entities (i.e., proper names) in ancient sources.⁴

The challenge with historical languages is due to their complexities, the fact that they are no longer spoken, and their high inflection, as in ancient Greek and Latin. As a matter of fact, data for proper names in the original language is still scarce, considering also that traditional print dictionaries, from which digital data is digitized, are not rich in proper names. Research in recent years has been focusing on this aspect to expand corpora of proper names, providing data in the original languages and not

1 The cover image was generated using Envato AI and licensed via Envato.

2 DFG Project No. 434173983. On this project, see Berti (2023b) and Berti (2024a).

3 See Berti (2019b) and Berti (2021), 398–414 with bibliography.

4 Berti (2019a), Berti (2019b), Beersmans et al. (2023), Beersmans et al. (2024), Palladino / Yousef (2024), Berti (2025b), Berti (2025c), Berti (2026a), Berti (2026b), Berti (2026c). Cf. also Ripoll Alberola et al. (2025).

just in modern contemporary ones. The goal is to enrich datasets with inflected forms and their lemmata, and annotate and represent this data according to the best recommendations of the Linked Open Data (LOD) principles.⁵ Moreover, current new possibilities offered by Large Language Models (LLM) and Artificial Intelligence (AI) request scholars to increase the number of data and control their quality in order to feed and train machines, so that in the future it will be possible to parse more texts and get better results.⁶

The workshop was structured in four sessions (see fig. 1) to introduce the concepts of Named Entities (NEs) and Linked Open Data (LOD) (1), and to deal specifically with Onomastics and Prosopography (2), Geography (3), and Authors and Works (4). International specialists currently working on these topics were invited to Leipzig to discuss various aspects of these themes.⁷

UNIVERSITÄT LEIPZIG
Historisches Seminar / Professor für Alte Geschichte

Gefördert durch **DFG** Deutsche Forschungsgemeinschaft
Projektnummer 434173983

International Workshop

Nomina Omina. Detecting and Preserving Ancient Greek and Latin Proper Names in the Age of Artificial Intelligence

Leipzig University, June 27-29, 2024

Thursday, June 27 - Seminargebäude - S420

Opening

14:30 **Monica Berti** (Universität Leipzig)
15:00 Welcome and Introduction

Session 1a - Named Entities and Linked Open Data for the Ancient World

15:00 **Adam Gtner** (Bayerische Akademie der Wissenschaften), *Digitalizing the Thesaurus Linguae Latinae Index librorum*
15:45
16:15 Coffee Break

16:15 **Andrea Beyer** (Humboldt-Universität zu Berlin), *Daidalos: NER for Literary Studies on Latin and Greek Texts*
17:00

17:00 **Giuseppe G.A. Celano** (Universität Leipzig), *Opera Graeca Adnotata: A Multilayer Corpus*
17:45

Friday, June 28 - Seminargebäude - S420

Session 1b - Named Entities and Linked Open Data for the Ancient World

09:30 **Carina Geldhauser** (Munich Centre for Machine Learning), *Automatic Annotation of Nomina Sacra*
10:15

10:15 **Irine Darchia** (Ivane Javakishvili Tbilisi State University), *Greek and Latin Proper Names in Georgian Epigraphic, Lexicographic and Encyclopaedic Material: Questions and Plans for Digitization*
11:00

11:00 Coffee Break
11:30

Session 2 - Onomastics and Prosopography

11:30 **Matilde Garré** (Université de Paris 1 Panthéon-Sorbonne, UMR 8210 ANHIMA), *LGPN-Ling for the Preservation of Greek Personal Names in a Digital Environment*
12:15

12:15 **Yanne Broux** (KU Leuven), *Of Gods and Men: Theophoric Names in the Intersection of TM People and TM Gods*
13:00

13:00 Lunch (Mensa am Park)
15:00

15:00 **Sylvain Lebreton** (Université Toulouse - Jean Jaurès), *Digital Divine Onomastics? About the Mapping Ancient Polytheisms Database*
15:45

Session 3 - Geography

15:45 **Monica Berti** (Universität Leipzig) and **Pietro Zaccaria** (KU Leuven), *Digital Paradoxography: Toponyms in Paradoxographical Texts*
16:30

16:30 Coffee Break
17:00

17:00 **Chiara Palladino** (Furman University), *So Much More than Names: Modeling Geographical Entities in Ancient Texts*
17:45

19:30 Dinner (Auerbachs Keller - Mädler-Passage)
22:00

Saturday, June 29 - Hörsaalgebäude - HS 8

Session 4 - Authors and Works

09:00 **Margherita Fantoli** (KU Leuven), *The NIKAW Project: Finding and Disambiguating References to People*
09:45

09:45 **Annette von Stockhausen** (Berlin-Brandenburgische Akademie der Wissenschaften), *Named Entities in the Patristic Text Archive (PTA)*
10:30

10:30 Coffee Break
11:00

11:00 **Andrea Balbo** and **Elisa Della Calce** (Università di Torino), *Eastern and Western Names in Latin: the SERICA Corpus*
11:45

11:45 **Ivan Matijašić** (Università Ca' Foscari Venezia), *Detecting Proper Names in Greek Fragmentary Historiography: Between Digital Philology and Prosopography*
12:30

12:30 Conclusions
13:30

KONTAKT
PD Dr. Monica Berti
Universität Leipzig
Historisches Seminar
Alte Geschichte
monica.berti@uni-leipzig.de

LAGL.org
Linked Ancient Greek and Latin

Fig. 1: *Nomina Omina* Workshop – the program.

5 On LOD for the Ancient World, see Cayless (2019) and Middle (2024).

6 Berti (2025c) and Berti (2026a).

7 Links to online resources, projects, and corpora cited in the following paragraphs can be found at the end of the paper in the section listing online resources.

After the introduction to the workshop, the first session (Named Entities and Linked Open Data for the Ancient World) was devoted to the current state of research for Named Entity Recognition (NER) of ancient Greek and Latin, on linguistic annotation as a foundational method to generate data for extracting named entities, and on digital projects born with the digitization of print publications and the annotation of proper names. Andrea Beyer (Humboldt-Universität zu Berlin) described methods developed at the DFG project *Daidalos* for Named Entity Recognition (NER) in ancient Greek and Latin, while Carina Geldhauser (Munich Center for Machine Learning – MCML) showed examples for *nomina sacra*; Giuseppe G. A. Celano (Universität Leipzig) presented the current state of linguistic data for ancient Greek in the DFG project *Opera Graeca Adnotata (OGA)*; finally, Adam Gitner (Bayerische Akademie der Wissenschaften) introduced the work done at the *Thesaurus Linguae Latinae (TLL)* project to extract and annotate proper names and bibliographic citations from the *Index Librorum*, and Irine Darchia (Ivane Javakhishvili Tbilisi State University) described Georgian projects to extract data from epigraphic, lexicographic, and encyclopedic material.

The second session (Onomastics and Prosopography) included three projects on different kinds of personal names. Matilde Garré (Université de Paris 1 Panthéon–Sorbonne) presented the work done for linguistic annotations of personal names in the project *LGPN-Ling*, which is a semantical complement to the *Lexicon of Greek Personal Names (LGPN)* Database, while Yanne Broux (KU Leuven) and Sylvain Lebreton (Université Toulouse – Jean Jaurès) introduced projects working on divine onomastics, such as *Trismegistos People*, *Trismegistos God*, and the *Mapping Ancient Polytheisms (MAP)* database.

The third session (Geography) was devoted to projects and methods for extracting toponyms from ancient sources. Pietro Zaccaria (KU Leuven) and Monica Berti (Universität Leipzig) presented the *TM Paradoxography* project of *Trismegistos* to extract place names from paradoxographical texts, and Chiara Palladino (Furman University) discussed geographical entities in ancient texts.

The fourth and last session (Authors and Works) had a focus on the individuation and extraction of proper names related to names of authors and forms referring to their works in ancient sources, which can be defined as bibliographic references in past texts. Margherita Fantoli (KU Leuven) and Annette von Stockhausen (Berlin-Brandenburgische Akademie der Wissenschaften) discussed this kind of entities in the two projects *NIKAW (Networks of Ideas and Knowledge in the Ancient World)* and *Patristic Text Archive (PTA)*; Andrea Balbo and Elisa Della Calce (Università degli Studi di Torino) presented the research currently done to extract proper names from Latin sources on China dated between Hellenistic times and the 19th century in the project *SERICA*, with a focus on the *Confucius Sinarum Philosophus*; finally, Ivan Matijašić (Università Ca' Foscari Venezia) discussed proper names in Greek fragmentary historiography with examples from Attidography.

The *Linked Ancient Greek and Latin (LAGL)* Project

The selection of the aforementioned topics and sections could have been expanded to include other aspects of Named Entity Recognition and the analysis of proper names. However, their choice was driven by two main reasons: 1) current needs and discussions in the community of Digital Classicists to improve NER and increase data for ancient Greek and Latin; 2) the project I am currently leading, which is called *Linked Ancient Greek and Latin (LAGL)* and which is extracting named entities from ancient Greek and Latin sources, and selecting those related to bibliographic references.⁸

Considering the still limited availability of digital data for ancient Greek personal names, the *LAGL* project was established to develop a workflow for extracting these forms from ancient sources, lemmatizing them, and labelling them according to the NE categories employed in computational linguist-

8 <https://www.lagl.org> (last access 23.09.2025).

ics, such as PER and PERderiv for personal names and derivatives (e.g., Θεμιστοκλῆς and Πυθαγορικός), LOC and LOCderiv for place names and derivatives (e.g., Ἀλικαρνασός and Μυτιληναῖος), ORG and ORGderiv for organizations and derivatives like festivals and schools (e.g., Παναθήναια and Στωϊκός), and OTH for other entities such as currencies, months, work titles, etc. (e.g., Ἱστορία and Ὀκτώβριος).⁹ These broad categories allow us to generate training data that can be used to train computational models for NER and extract semi-automatically proper names from ancient sources. For example, the form Λαοδίκειαν can be extracted, lemmatized as Λαοδίκεια, and labeled as a location (LOC). Through the lemma, it is possible to individuate other inflected forms, like for example Λαοδικείας, that can be further disambiguated in context as Laodicea in Syria or Laodicea on the Lycus in Asia Minor.¹⁰

A second part of the project is the *LAGL Catalog*, which collects linguistic forms about author names and their works cited in ancient sources.¹¹ This analysis is related to citation detection and fragmentary literature, even if the focus is not on the content of citations (i.e., the so called fragments), but on those elements that signal the presence of quotations, like names of authors, titles and descriptions of their works, and indirect references to them, such as unnamed authors and works.¹² For this kind of analysis, NER is particularly helpful, because named entities function in the text as anchors to individuate and disambiguate further elements of bibliographic references. This type of research is rooted in my extensive experience with fragmentary texts and citation analysis, which reveal the numerous linguistic forms concealed behind modern scholarly labels like *fragmenta*, *testimonia*, *reliquiae*, etc.¹³

These linguistic forms are important to explore the ancient language of proper names and bibliographic references, advance our knowledge of citation practices in the ancient world, and enrich digital data in the original language. An example among the many others that I could mention is Duris (Δοῦρις: PER), the ancient Greek historian and tyrant of the island of Samos, who is connected by ancient sources to a great variety of named entities. Duris, in fact, is not only labelled as Samius (Σάμιος: LOCderiv) and tyrant of Samos (Σάμου: LOC), but also as a pupil of Theophrastus of Eresus (Θεοφράστου Ἐρεσίου: PER|LOCderiv), brother of the author Lynceus of Samos (Λυγκεὺς Σάμιος: PER|LOCderiv), descendant of Alcibiades (Ἀλκιβιάδου: PER), and probably relative of the Kaios (Καῖος: PER) victorious in the boy's boxing contest and commemorated with a statue in Olympia (Ὀλυμπία: LOC) made by a certain Hippias (Ἱππίου: PER).¹⁴ The name of Duris is cited in lists of ancient Greek and Latin sources with authors like Phylarchos (Φύλαρχον: PER), Polybios (Πολύβιον: PER), Psaon (Ψάωνα: PER), Demetrios of Kallatis (Καλλατιανὸν Δημήτριον: LOCderiv|PER), Hieronymos (Ἱερώνυμόν: PER), Antigonos (Ἀντίγονον: PER), Herakleides (Ἡρακλείδην: PER), Hegesianax (Ἡγησιάνακτα: PER), Hellenicus (Ἑλλάνικοί: PER), Onesicritus (Onesicrito: PER), Clitarchus (Κλιτάρχο: PER), Ctesia (Ctesia: PER), Philistus (Philisto: PER), and many others.¹⁵ Moreover, references to Duris' works show his heterogeneous interests that ranged from Homer (Προβλημάτων Ὀμηρικῶν: OTH|PERderiv), to tragedians like Euripides (Εὐριπίδου: PER) and Sophocles

9 Berti (2019b), Berti (2023b), Berti (2025b).

10 On the two ancient cities, see the gazetteer entries in Pleiades at <https://pleiades.stoa.org/places/668290> (last access 23.09.2025) and <https://pleiades.stoa.org/places/638955> (last access 23.09.2025). For examples of sources mentioning the two places, see *Suda*, alpha 3398 (s.v. Ἀπολιναῖριος) and gamma 450 (s.v. Γρηγόριος).

11 <https://catalog.lagl.org> (last access 23.09.2025).

12 Berti (2013), Berti (2018), and Berti (2025a).

13 Berti (2021), Berti (2023a), and Berti (2024b) with examples, bibliography, and projects related to this topic.

14 For the ancient sources on these assertions and commentaries on them, see FGrHist (= BNJ) 76 and BNP s.v. Duris (3).

15 See FGrHist (= BNJ) 76.

(Σοφοκλέους: PER), and to historical works on Samos (Σαμίων Ἱστοροί: LOCderiv|OTH), Agathocles (Ἀγαθοκλέα: PER), and Macedonia (Μακεδονικά: LOCderiv).¹⁶

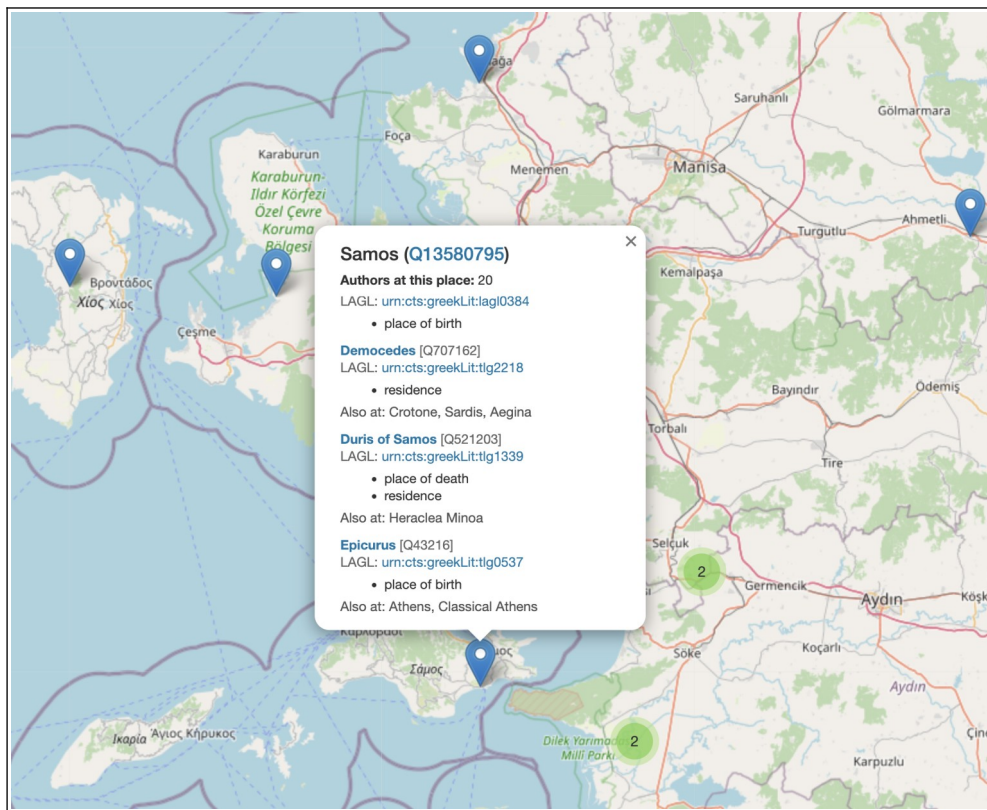


Fig. 2: LAGL Catalog – experimental map with Wikidata properties.

The LAGL project is collecting linguistic annotations like those mentioned in the previous paragraph to create a dataset of named entities in ancient Greek and Latin, and further disambiguate author names and work titles. To accomplish these goals, the project started with the semi-automatic annotation of the texts of Athenaeus of Naucratis (*Deipnosophists*), Valerius Harpocration (*Lexicon of the Ten Orators*), and the *Suda*, because they are very rich in named entities for ancient Greek, which is a language where data is still very scarce.¹⁷ In order to get an estimate, the project currently contains the following occurrences of inflected proper names in ancient Greek: 23,583 personal names (PER) and 837 derivatives (PERderiv), 3,916 place names (LOC) and 6,649 derivatives (LOCderiv), 299 organization names (ORG) and 59 derivatives (ORGderiv), and 2,721 other entities not classifiable in the previous categories (OTH).¹⁸ As far as authors and works are concerned, the current coverage of the LAGL Catalog is of 1,242 authors and 3,228 works cited in the above mentioned sources.¹⁹ The project will expand to other sources in order to increase the number of named entities in ancient Greek and add those in Latin.²⁰

16 For a complete collection of sources citing Duris' works and commentaries on them, see FGrHist (= BNJ) 76. The Named Entity labels listed in this paragraph are examples of inflected Greek and Latin forms as they occur in ancient sources that are extracted and disambiguated in the context.

17 On the reasons for starting with these sources, see Berti (2026a).

18 These numbers refer to occurrences of inflected forms not lemmatized (e.g., every occurrence of Κτησιφῶντος). The annotation of the *Suda* is in progress, and data is now available for the entries of the letters alpha, beta, and gamma: see <https://www.lagl.org/tools/suda/> (last access 23.09.2025). For the meaning of these NE labels, see above in the paper.

19 These numbers refer to the total number of authors and works with a unique identifier in the LAGL Catalog under which linguistic annotations referring to them are collected. These authors and their annotations can be found at <https://catalog.lagl.org/> (last access 23.09.2025).

20 Berti (2024a) and Berti (2026a).

LAGL annotations of authors and works are enriched with CTS URNs and *Wikidata* IDs that function as unique identifiers to connect the annotated forms with metadata and external resources.²¹ For example, annotations referred to Duris of Samos are collected under [urn:cts:greekLit:tlg1339](https://catalog.perseus.org/catalog/urn:cts:greekLit:tlg1339), which is the unique identifier of this author in the *Perseus Catalog* that preserves the four-digit number [tlg1339](https://catalog.perseus.org/catalog/urn:cts:greekLit:tlg1339) of the *Thesaurus Linguae Graecae (TLG)*.²² The *Wikidata* ID [Q521203](https://www.wikidata.org/wiki/Q521203) accompanies the CTS URN [urn:cts:greekLit:tlg1339](https://catalog.perseus.org/catalog/urn:cts:greekLit:tlg1339) in the *LAGL Catalog* for Duris of Samos. The corresponding *Wikidata* item also includes a *LAGL* author ID property (P12869) that collects *LAGL* CTS URNs, linking the *LAGL Catalog* to metadata of the *Wikidata* knowledge base, and vice versa, linking *Wikidata* to the linguistic annotations of the *LAGL* project.²³

Metadata deriving from *Wikidata* can be extracted and used for many possible views and analyses. For example, locations currently connected with *LAGL* authors in *Wikidata* permit the creation of maps. Fig. 2 shows an experimental map with Duris as one of the twenty authors currently collected in the *LAGL Catalog* and listed under ancient Samos ([Q13580795](https://www.wikidata.org/wiki/Q13580795) with *Pleiades* ID [599925](https://pleiades.stoa.org/places/599925)), because ancient sources connect this place to these authors in assertions referring to their place of birth or intellectual activity.²⁴

Nomina Omina – The Proceedings

These proceedings follow the structure and order of the workshop, excluding four papers that were not submitted to the editor, and adding three new ones by speakers who attended the workshop in person at Leipzig University and have an interest in its topics.

The paper by Anna Clara Maniero Azzolini (University of London) presents prosopographical data extracted from inscriptions of the Roman period from the ancient city of Altinum in Veneto, Italy, and its insertion into *Wikidata*. Camillo Carlo Pellizzari di San Girolamo (Scuola Normale Superiore Pisa) discusses named entities in the Wikibase *Hypotheseis*, which is a relational database of Greek rhetorical exercises written in Greek from the Hellenistic to the Byzantine age. Finally, Farnoosh Shamsian (Universität Leipzig) and Monica Berti (Universität Leipzig) present the translation alignment, annotation, and analysis of named entities in the trilingual inscription (Greek, Middle Persian, and Parthian) of Shapur I at Ka'ba-ye Zartošt (ŠKZ) located in Naqsh-e Rostam, Fars province, Iran.

These three new articles represent a significant addition to the topics covered in the workshop, as they provide further data from epigraphic sources and ancient languages other than Greek and Latin, demonstrating the potential and importance of Digital Classics methods in addressing the new challenges that await us in the face of the fascinating and important technological revolution of Artificial Intelligence.

21 On *Wikidata* and its use in the humanities, see Vrandečić et al. (2023) and Zhao (2023). On the interactions between the *LAGL* Project and *Wikidata*, see Berti (2025c).

22 See <https://catalog.perseus.org/catalog/urn:cts:greekLit:tlg1339> (last access 23.09.2025). On the *Perseus Catalog* and the use of CTS URNs, see Babeu (2019).

23 See <https://www.wikidata.org/wiki/Q521203> (last access 23.09.2025).

24 See <https://www.wikidata.org/wiki/Q13580795> (last access 23.09.2025) and <https://pleiades.stoa.org/places/599925> (last access 23.09.2025). The experimental map – available at <https://catalog.lagl.org/> (last access 23.09.2025) – shows all *Wikidata* items with a *LAGL* author ID (P12869) and their associated *Wikidata* location properties: place of birth (P19), place of death (P20), country of citizenship (P27), ancestral home (P66), place of burial (P119), residence (P551), location of formation (P740), work location (P937). This map is updated regularly to correct possible mistakes and inconsistencies, and keep track of changes in *Wikidata*. For this map and other *Wikidata*-related resources in the *LAGL* Project, see Berti (2025c).

In conclusion to this introduction, I express my sincere thanks to the contributors, who accepted the invitation to come to Leipzig for the workshop and submitted their articles, providing stimulating ideas for the future of data on the ancient world and for our disciplines. I am also very grateful to the German Research Foundation (Deutsche Forschungsgemeinschaft) for supporting this publication, to Leipzig University for facilitating the organization of the workshop, to Alexander Plate for working on the layout of this publication, and to the Editorial Team of Digital Classics Online for hosting these proceedings as a special issue of the journal.

List of Abbreviations

- BNJ I. Worthington (ed.), Brill's New Jacoby, Leiden 2006–.
- BNP H. Schneider / M. Landfester / H. Cancik (eds.), Brill's New Pauly, Leiden 1996–.
- FGrHist F. Jacoby (ed.), Die Fragmente der griechischen Historiker, I–III, Leiden 1923–1958.

Sources

Online sources

- Altinum: a Wikidata Project for Digital Epigraphy: https://www.wikidata.org/wiki/User:Anna_Clara_Maniero_Azzolini/Altinum (last access 23.09.2025).
- Daidalos-Projekt: NLP in der Klassischen Philologie: <https://daidalos-projekt.de> (last access 23.09.2025).
- Hypotheseis: <https://hypotheseis.wikibase.cloud/> (last access 23.09.2025).
- LGPN – The Lexicon of Greek Personal Names: <https://www.lgpn.ox.ac.uk/> (last access 23.09.2025).
- LGPN-Ling. Etymology and semantic of ancient Greek Personal Names: <https://lgpn-ling.humanum.fr/> (last access 23.09.2025).
- Linked Ancient Greek and Latin (LAGL) Project: <https://www.lagl.org/> (last access 23.09.2025).
- Linked Ancient Greek and Latin (LAGL) Catalog: <https://catalog.lagl.org/> (last access 23.09.2025).
- Mapping Ancient Polytheisms (MAP): <https://base-map-polytheisms.humanum.fr> (last access 23.09.2025).
- NIKAW (Networks of Ideas and Knowledge in the Ancient World): <https://research.kuleuven.be/portal/en/project/3H220323> (last access 23.09.2025).
- Opera Graeca Adnotata (OGA): <https://varro.informatik.uni-leipzig.de/oga/en/> (last access 23.09.2025).
- Patristic Text Archive (PTA) – An Open Access Archive of Ancient Christian Texts: <https://pta.bbaw.de/> (last access 23.09.2025).
- Perseus Catalog: <https://catalog.perseus.org> (last access 23.09.2025).
- Pleiades Gazetteer: <https://pleiades.stoa.org/> (last access 23.09.2025).
- SERICA Project: <https://serica.unipi.it> (last access 23.09.2025).
- Thesaurus Linguae Graecae (TLG): <https://stephanus.tlg.uci.edu/> (last access 23.09.2025).
- Thesaurus Linguae Latinae (TLL): <https://thesaurus.badw.de/> (last access 23.09.2025).
- Trismegistos God: <https://www.trismegistos.org/god/> (last access 23.09.2025).
- Trismegistos Paradoxography: <https://www.trismegistos.org/paradoxography/> (last access 23.09.2025).
- Trismegistos People: <https://www.trismegistos.org/ref/> (last access 23.09.2025).
- Word-level Alignment and Named Entities in the Trilingual Inscription at Ka'ba-ye Zartošt (ŠKZ): <https://zenodo.org/records/15050878> (last access 23.09.2025).

References

- Babeu (2019): A. Babeu, The Perseus Catalog: of FRBR, Finding Aids, Linked Data, and Open Greek and Latin, in: M. Berti (ed.), *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*, Berlin 2019, 53–72, <https://doi.org/10.1515/9783110599572-005> (last access 15.04.2026).
- Beersmans et al. (2023): M. Beersmans / E. de Graaf / T. Van de Cruys / M. Fantoli, Training and Evaluation of Named Entity Recognition Models for Classical Latin, in: *Proceedings of the Ancient Language Processing Workshop, Varna (Bulgaria) 2023*, 1–12, <https://aclanthology.org/2023.alp-1.1/> (last access 23.09.2025).
- Beersmans et al. (2024): M. Beersmans / A. Keersmaekers / E. de Graaf / T. Van de Cruys / M. Depauw / M. Fantoli, “Gotta catch ‘em all!”: Retrieving people in Ancient Greek texts combining transformer models and domain knowledge, in: *Proceedings of the 1st Workshop on Machine Learning for Ancient Languages (ML4AL 2024)*, 152–164, <https://aclanthology.org/2024.ml4al-1.16/> (last access 23.09.2025).
- Berti (2013): M. Berti, Collecting Quotations by Topic: Degrees of Preservation and Transtextual Relations among Genres, *Ancient Society* 43 (2013), 269–288.
- Berti (2018): M. Berti, Annotating Text Reuse within the Context: the Leipzig Open Fragmentary Texts Series (LOFTS), in: U. Tischer / A. Forst / U. Gärtner (eds.), *Text, Kontext, Kontextualisierung. Moderne Kontextkonzepte und antike Literatur*, Hildesheim 2018, 223–234.
- Berti (2019a): M. Berti (ed.), *Digital Classical Philology. Ancient Greek and Latin in the Digital Revolution*, Berlin 2019, <https://doi.org/10.1515/9783110599572> (last access 15.04.2026).
- Berti (2019b): M. Berti: Named Entity Annotation for Ancient Greek with INCEpTION, in: K. Simov / M. Eskevich (eds.), *Proceedings of CLARIN Annual Conference 2019*, Leipzig 2019, 1–4.
- Berti (2021): M. Berti, *Digital Editions of Historical Fragmentary Texts*, Heidelberg 2021, <https://doi.org/10.11588/propylaeum.898> (last access 15.04.2026).
- Berti (2023a): M. Berti, Ancient Greek Historians in the Digital Age, in: M. Althage / M. Dröge / T. Hiltmann / C. Prinz (eds.), *Digitale Methoden in der geschichtswissenschaftlichen Praxis: Fachliche Transformationen und ihre epistemologischen Konsequenzen: Konferenzbeiträge der Digital History 2023*, Berlin 2023, <https://doi.org/10.5281/zenodo.8322062> (last access 15.04.2026).
- Berti (2023b): M. Berti, Named Entity Recognition for a Text-Based Catalog of Ancient Greek Authors and Works, in: A. Baillot / W. Scholger / T. Tasovac / G. Vogeler (eds.), *Digital Humanities 2023: Book of Abstracts*, University of Graz: Austrian Centre for Digital Humanities 2023, 557–558, <https://doi.org/10.5281/zenodo.8108058> (last access 15.04.2026).
- Berti (2024a): M. Berti, Digital Canons and Catalogs of Fragmentary Literature, in: F. Neuerburg / T. Tsiampokalos / P. Wozniczka (eds.), *Fragmente einer fragmentierten Welt. Zur Problematik des Umgangs mit Fragmenten in der gegenwärtigen klassisch-philologischen Forschung*, Berlin 2024, 217–236, <https://doi.org/10.1515/9783111508788-009> (last access 23.09.2025).
- Berti (2024b): M. Berti, Digital Practice for Studying the Indirect Transmission of Classical Authors and Works, in: V. Mastellari / F. Favi (eds.), *Treasuries of Literature: Anthologies, Lexica, Scholia, and the Indirect Tradition of Classical Texts in the Greek World*, Berlin 2024, 189–205, <https://doi.org/10.1515/9783111386010-010> (last access 15.04.2026).

- Berti (2025a): M. Berti, Canoni e cataloghi collaborativi per una filologia sostenibile in ambiente digitale, in: S. Cannavale / V. Casapulla / M. Verstraete / M. Vitali-Rosati (eds.), *Orizzonti della filologia digitale. L'Antologia Greca per ripensare formati, paradigmi e collaborazione*, Napoli 2025, 131–152.
- Berti (2025b): M. Berti, Digital Catalogs of Ancient Greek Authors and Works through Papyrological Data, in: N. Reggiani (ed.), *Digital Papyrology III. The Digital Critical Edition of Greek Papyri: Issues, Projects, and Perspectives*, Berlin 2025, 89–106, <https://doi.org/10.1515/9783111070162> (last access 15.04.2026).
- Berti (2025c): M. Berti, Linked Ancient Greek and Latin (LAGL) and Wikidata: Structuring and Re-using Data of Classical Literature (discussion paper): in *Journal of Open Humanities Data* 11/72 (2025), 1–12, <https://doi.org/10.5334/johd.423> (last access 15.04.2026).
- Berti (2026a): M. Berti, Traces of Lost Libraries in the Works of Alexandrian and Byzantine Scholars: Bibliographic open data from antiquity to Wikimedia, in: G. Bodard / V. Vitale (eds.), *Open Data in Ancient and Byzantine Studies*, London 2026 (forthcoming).
- Berti (2026b): M. Berti, Annotating the Ancient World. Critical Annotations and Digital Editions, in: P. d'Hoine / D. Kohler / W. Decock (eds.), *Charting the Future of Historical Humanities*, Turnhout 2026, 21–46.
- Berti (2026c): M. Berti, Digitale Philologie: in V. Schulz / A. Schwab (eds.), *Moderne Zugänge zu antiken Texten. Theorien, Konzepte, Methoden und ihre Anwendung*, Stuttgart 2026 (forthcoming).
- Cayless (2019): H. A. Cayless, Sustaining Linked Ancient World Data, in: M. Berti (ed.), *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*, Berlin 2019, 35–50, <https://doi.org/10.1515/9783110599572-004> (last access 15.04.2026).
- Middle (2024): S. Middle, Linked Ancient World Data: Implementation, Advantages, and Barriers, *Digital Classics Online* 10/1 (2024), 16–49, <https://doi.org/10.11588/dco.2024.10.104105> (last access 15.04.2026).
- Palladino / Yousef (2024): C. Palladino / T. Yousef, Development of Robust NER Models and Named Entity Tagsets for Ancient Greek, in: *Proceedings of the Third Workshop on Language Technologies for Historical and Ancient Languages (LT4HALA) @ LREC-COLING-2024*, Torino 2024, 89–97, <https://aclanthology.org/2024.lt4hala-1.11/> (last access 23.09.2025).
- Ripoll Alberola et al. (2025): L. Ripoll Alberola / L. D'Addario / M. Burghardt / M. Berti / M. Depauw, Tracing Antiquity: References to Greco-Roman Authors in Modern Academic Discourse, in: *Digital Humanities 2025* (accepted long paper forthcoming).
- Vrandečić et al. (2023): D. Vrandečić / L. Pintscher / M. Krötzsch, Wikidata: The Making Of, in: *Companion Proceedings of the ACM Web Conference 2023 (WWW '23 Companion)*, New York (NY) 2023, 615–624, <https://doi.org/10.1145/3543873.3585579> (last access 15.04.2026).
- Zhao (2023): F. Zhao, A Systematic Review of Wikidata in Digital Humanities Projects, in: *Digital Scholarship in the Humanities* 38/2 (2023), 852–874, <https://doi.org/10.1093/lc/fqac083> (last access 15.04.2026).

Figure References

Fig. 1: *Nomina Omina* Workshop – The Program

Fig. 2: *LAGL Catalog* – Experimental Map with *Wikidata* Properties

Author Contact Information²⁵

PD Dr. Monica Berti
Universität Leipzig
Lehrstuhl für Alte Geschichte
Beethovenstraße 15
04107 Leipzig
E-mail: monica.berti@uni-leipzig.de

²⁵ The rights pertaining to content, text, graphics, and images, unless otherwise noted, are reserved by the author. This contribution is licensed under CC BY-SA 4.0.