DIGITAL CLASSICS ONLINE

Band 10,1 (2024)

ISSN: 2364-7957

DIGITAL CLASSICS ONLINE

Bd. 10,1 (2024)

Aus dem Inhalt

Charlotte Schubert:
Editorial: Die digitale Revolution: Wirklich nichts Neues?……..I–IV

Andreas P. Antonopoulos et al.:
A Digital Timeline of Greek Theatre Events……..……1–15

Sarah Middle:
Linked Ancient World Data: Implementation, Advantages, and Barriers
…….……16–49

Marco Cornaglia:
GIS analysis and the abandoned settlements in Pausanias' Periegesis: a methodological proposal……..50–75

Franziska Schropp et al.:
Manuell-hermeneutische Forschung als Goldstandard: Zur Kalibrierung digitaler Analysetools in den Digital Humanities……...76–98

DIGITAL CLASSICS ONLINE

Editorial

Charlotte Schubert

Die digitale Revolution: Wirklich nichts Neues?

Image of a futuristic city, including a woman **(sic!)** ***at a futuristic desk*****, erstellt mit** ***Dall-E*** **in** ***ChatGPT 4.0*** **(20.02.2024).**[1]

1 Für den Prompt war eine explizite Aufforderung, eine weibliche Person abzubilden, notwendig, da ansonsten entsprechende Visualisierungen immer mit männlichen Personen präsentiert wurden.

2022 hat Franco Moretti (fast jeder in unserem Bereich kennt sein Buch *Graphs, Maps, Trees* mit der Wortprägung ‚*Distant Reading*' als Emblem für die Digital Humanities), ein schmales Bändchen mit dem Titel *Falso Movimento* publiziert, ebenfalls 2022 auf Deutsch unter dem Titel *Falsche Bewegung*. Er macht darin eine Schere auf zwischen der quantitativen Literaturwissenschaft und den digitalen Humanwissenschaften. Für Letztere wird ein höherer Grad an Präzision in Anspruch genommen, für Erstere eine Verschlechterung, fast schon ein Rückschritt in der Forschung. Das Beispiel, an dem er das festmacht, ist unter anderem das *Topic Modeling*: indem Texte als Wortsäcke betrachtet (*Bag of Words*, ein ziemlich alter Hut, der schon vor den Digital Humanities getragen wurde) und damit formale Konzepte (als wesentlich und typisch für die Literaturwissenschaft postuliert) somit verworfen würden. Form als Begriff sei die ästhetische Sphäre der literaturwissenschaftlichen Arbeit und diese wiederum sei durch das *Topic Modeling*, das *Text Mining* und andere Formen der digitalen Analyse verloren gegangen. Moretti spitzt dies sogar noch zu mit dem Ausdruck: „Austreibung der Form" (S. 9, Anm. 3 der deutschen Ausgabe). Er reflektiert in dieser Aufsatzsammlung eine Art des persönlichen Scheiterns: Wie er selbst konstatiert, so sei es ihm nicht gelungen, Morphologie und Geschichte, Quantifizierung und Hermeneutik, Normen und Anomalie in der datengetriebenen Forschung zu einer Synthese zu bringen. Mit einem persönlichen Rückblick beschließt er die Aufsatzsammlung: Am Beispiel der datengetriebenen Untersuchung, der Entwicklung von Segmentierungssystemen (damit ist die grafische Unterscheidung von direkter und indirekter Rede, Navigationssysteme wie Kapiteleinteilungen, etc. gemeint) vs. der Bedeutung des hermeneutischen Zirkels für die Untersuchung und Analyse von Texten kommt er zu dem Ergebnis, dass die Arbeit mit Datenbanken und Algorithmen eine Forschung hervorgebracht habe, die sich selbst genüge, und diese in keiner Verbindung mehr zu der langen und gut entwickelten hermeneutischen Praxis stehe (S. 164 der deutschen Ausgabe). Das gipfelt in folgendem Satz: „Jedes Mal, wenn man ein Korpus durchrechnen ließ, fand man irgendetwas. Es war wie ein Wunder". Statt dieses Phänomen der Zufälligkeit auf die Frage ihrer Scheinbarkeit bzw. kontextualisierenden Plausibilisierung systematisch zu untersuchen, gibt er sich mit Resignation zufrieden.

Ist das Ergebnis dieser Entwicklung also ein Ende der Theorie? Er sieht die Mathematisierung der geisteswissenschaftlichen Texte und Quellen in der Forschung als Verlust von „wissenschaftlichen Ambitionen", als „eine Lawine, kleinere Studien, ohne jede geistige Synthese" (S. 165).

Es ist traurig und enttäuschend, wenn einer der Protagonisten der Forschungsrichtung, die die klassischen Geisteswissenschaften seit mehr als 20 Jahren verändert, bereichert und weiterentwickelt hat, jetzt anscheinend die Segel streicht. Ausgerechnet jetzt, da durch *ChatGPT* und die *LLMs* (*Large Language Models*) weitere, neue Türen aufgestoßen worden sind, die nicht von der Forschung an naturwissenschaftlichen Phänomenen, nicht von der Forschung von sozialen und wirtschaftlichen Phänomenen ausgehen, sondern auf der Grundlage von Texten entwickelt wurden.

Es scheint möglicherweise billig, darauf hinzuweisen, dass Theorien, wie gerade die Hermeneutik und ihre diversen Verzweigungen (der hermeneutische Zirkel ist nicht erst von Dilthey und Schleiermacher entdeckt worden!), sich nicht in 20 Jahren entwickelt haben, sondern seit der Antike bis heute eine mehrtausend-jährige Geschichte haben. Methoden sind kurzlebig, Theorien sind langlebig – Theorien benötigen Zeit und Muße, Austausch und Diskurs. Nicht umsonst ist eines der grundlegenden Prinzipien des Sokratischen Dialogs bei Platon das λόγον τε δοῦναι καὶ δέξασθαι (Rede und Antwort stehen)! Die Forderung, dass durch den Einsatz neuer Methoden auch zeitgleich neue Theorien hervorgebracht werden sollen, erscheint mir als zu weitgehend und die Ansicht, dass die Statistik soweit überhandgenommen habe, dass sie theoretisches Wissen verdrängen würde, erscheint mir als unzulässige Verkürzung. Natürlich ist es für klassisch ausgebildete Geisteswissenschaftlerinnen und Geisteswissenschaftler nicht einfach, sich in die Tiefen von statistischen Modellen und in die Komplexität von Algorithmen einzuarbeiten. Aber dass dies notwendig ist und damit einher auch die theoretische Reflexion der Vorbedingungen und Kontextualisierung gehen muss, ist doch heute offensichtlich. Die Digital Humanities sind – für sich betrachtet – nach wie vor ein sehr kleiner Bereich. Sieht man jedoch die instituti-

onelle Ausweitung der Geisteswissenschaften im Hinblick auf die universitären Institutionen, die großen Forschungsrichtungen, das Vordringen der digitalen Technologien in fast alle Bereiche des Alltags, dann erkennt man doch sofort, womit wir es zu tun haben: eine Veränderung unserer Lebenswelt, die die industrielle Revolution bei weitem in den Schatten stellen wird. Vergleichbar ist meines Erachtens eher der Übergang vom Nomadentum zur Sesshaftigkeit oder von der Mündlichkeit zur Schriftlichkeit. Und wann hat sich eine Theorie entwickelt, die dies reflektiert hat? Die ersten theoretischen Überlegungen, die diesen Übergang (heute als neolithische Revolution bezeichnet) im Rahmen von Evolutionstheorien reflektiert haben, sind die Evolutionstheorien der griechischen Antike (die aber auch schon die solchen Prozessen inhärenten Risiken kannten: Sophokles, *Antigone* 332ff.). Der Begriff ‚neolithische Revolution' ist erst im 20. Jahrhundert geprägt worden, analog zur industriellen Revolution. Die digitale Revolution hat gerade erst begonnen, Begriffe und Reflexion sind noch rudimentär.

Eines jedoch lässt sich festhalten: Drastische Überschätzung vs. Unterschätzung ist typisch für die Einschätzung von Umbruchsituationen, insbesondere bei technologischen Entwicklungen, die oft mehrere Schübe benötigen, bis sie ihre volle Wirkung entfalten (Edisons Glühbirne ist vielleicht eine der bekanntesten). An den heute bei uns verwendeten Begriffen ist deutlich zu erkennen, wie unsicher und z.T. irreführend die Diskussion abläuft: So werden am laufenden Band falsche Bilder wie ‚künstliche Intelligenz' oder ‚neuronale Netze' für etwas verwendet, was schlicht und einfach mit dem korrekten Terminus als ‚*Machine Learning*' zu bezeichnen ist. Da sich bis heute keine allgemein akzeptierte Definition für Intelligenz durchgesetzt hat, sollte man auf diesen umstrittenen Begriff einfach verzichten. *Machine Learning* bzw. maschinelles Lernen, legt den Akzent richtig auf den maschinellen und skalierbaren Prozess, um den es geht.

Der Umbruch, der heute schon sichtbar ist, wird dazu führen, dass einfache, zu automatisierende Tätigkeiten, also geistige Routineaufgaben in den Bereichen der Bürokratie, der Verwaltung, der Behörden und vieler ähnlicher Institutionen alles, was auf der Grundlage von Mustern maschinell operationalisiert werden kann, was als Auswertung und Recherche noch heute unendlich viel menschliche Aufmerksamkeit, Kraft und Fähigkeit bindet, von *Bots*, also Maschinen, übernommen wird. Auch scheinbar schwierigere Aufgabenbereiche dürften betroffen sein: Daniel Katz von der Bucerius Law School Hamburg und Michael Bommarito (Chicago Kent College of Law) haben in einem Test *ChatGPT* das amerikanische Staatsexamen schreiben lassen.[2] *ChatGPT* sollte den Multiple-Choice-Teil dieses Examens absolvieren und in den Kategorien Delikts-, Vertrags-, Verfassungs- und Strafrecht schnitt *ChatGPT* immerhin mit 70% richtigen Antworten ab und im Beweis- und Sachenrecht mit noch 65%. Zum Vergleich: Nach siebenjähriger Ausbildung liegt das Ergebnis amerikanischer Studenten bei durchschnittlich 68%. Die Ergebnisse von *ChatGPT* betreffen nur den Multiple-Choice-Teil und sind auch nicht gerade überdurchschnittlich, allerdings dürften die nächsten *GPT*-Modelle deutlich leistungsfähiger sein. Immerhin ist für das erfolgreiche Ablegen eines juristischen Examens das Beherrschen von komplizierten Argumentationsstrukturen notwendig, ebenso wie für das Erstellen eines Programmcodes, für das *ChatGPT* gleichfalls heute schon häufig eingesetzt wird.

An diesem Beispiel wird sichtbar, welche erheblichen Auswirkungen diese digitale Revolution haben wird: In allen Belangen, in denen die Effizienz, die Schnelligkeit und Skalierbarkeit der *Bots* für eine Operationalisierung angewendet werden kann, werden sie auch eingesetzt werden! Dies ergibt sich schlicht aus der Logik aller Geschäftsmodelle, die auf diesen Prinzipien aufgebaut sind. Es wird ganz sicher so sein, dass alles, was eine Maschine besser kann als ein Mensch, auch dorthin verlagert werden wird – wie es die industrielle Revolution bereits eindrücklich gezeigt hatte.

Entgegen der Skepsis von Franco Moretti ist hier doch deutlich zu erkennen, wie das Zukunftsversprechen der Digitalität nun in den Alltag unserer Gesellschaft Einzug hält, allerdings auch verbunden mit großen Risiken. Beides hängt mit der bisher fehlenden gesellschaftlichen Einbindung der digitalen Revolution zusammen. Einerseits fehlt die von früh auf einzuübende Kompetenzbildung, deren Fehlen zu

2 https://www.legalvisio.de/chatgpt-schreibt-juristische-staatsexamen/ (abgerufen am 20.2.2024).

gefährlichen Auswüchsen insb. im Hinblick auf die Gefährdung von Kindern und Jugendlichen führt. Aber andererseits ist endlich eine Abwendung von den medialen Ruinen von gestern und vorgestern notwendig – denn wer kennt nicht (nur als Beispiel!) die zahllosen ‚Datenbanken' etc., die weder gepflegt, aktualisiert oder überhaupt noch erreichbar sind? Der Blick muss sich auf die Möglichkeiten der Entlastung von Routineaufgaben wenden, aus heutiger Perspektive sicher nicht für alle wünschenswert, aber dass dadurch neue Freiräume entstehen, die Bildung, Wirtschaft, Freizeit etc. komplett verändern werden, ist die große Chance am Horizont. Das – aus meiner Sicht – interessanteste an der Entwicklung ist, dass die Grundlage der Entwicklung nicht die Entdeckung eines Naturgesetzes oder einer mechanischen Entwicklung ist, sondern dass es Texte und Methoden der Textverarbeitung sind. Im Kern dieser allerneuesten Umbrüche stehen Sprachmodelle, die Texte verarbeiten. Mit Sprachmodellen werden nicht nur Computer programmiert, sondern Unterhaltungen sind möglich, medizinische Diagnosen können erstellt, Akten, Anträge, Berichte aller Art analysiert und ausgewertet werden und soviel anderes mehr. Dass Textwissenschaftlerinnen und Textwissenschaftler die darin liegenden Möglichkeiten (noch?) nicht erkannt haben, ist ausgesprochen spannend.

Autoreninformation[3]

Prof. em. Dr. Charlotte Schubert
Universität Leipzig
Historisches Seminar
Lehrstuhl für Alte Geschichte
E-Mail: schubert@uni-leipzig.de

3 Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei der Autorin. Alle Inhalte dieses Beitrages unterstehen, soweit nicht anders gekennzeichnet, der Lizenz CC BY 4.0.

DIGITAL CLASSICS ONLINE

A Digital Timeline of Greek Theatre Events

Andreas P. Antonopoulos, Stylianos Chronopoulos, Nikolaos Ntaliakouras, Panagiota Taktikou, Anastasia Psomiadou, Aikaterini Giakimtsouki-Magaraki, Iraklis Markelis

Abstract: The present paper describes the dataset, the visualisation, and the storage details of the *Timeline* tool developed in the framework of the project *Greek Fragmentary Tragedians Online*. The tool is based on a structured collection of *events* related to the fragmentary tragedians of the 6th and 5th centuries BC and visualises them in a vertical timeline graph.

A tool for the study of Greek Drama: the motivation

The project *Greek Fragmentary Tragedians Online* (*FragTrag.1*) is developing an open-access database[1] containing all available information about 45 Greek tragedians of the 6th and the 5th centuries BC, whose works have come down to us in a fragmentary state.[2] In keeping with the tradition of print editions of fragmentary texts, *FragTrag.1* collects, edits, and translates *testimonia* pertaining to the lives, careers and works of the 'fragmentary' tragic poets.[3]

These *testimonia* are extracted from five types of ancient and medieval sources: (1) Inscriptions related to the history of dramatic contests (*didascaliae*, victor lists, et al.),[4] (2) hypotheseis of plays preserved on ancient papyri and medieval manuscripts, (3) chronographical works (such as the *Marmor Parium* and Eusebius' *Chronicle*), (4) lexicographical, encyclopedic, and biographical works (such as the *Suda*), (5) other literary works.

1 https://fragtrag1.upatras.gr (last access 15.01.2024).

2 For more information on the nature and the specifics of the project, see Antonopoulos et al. (2023). *FragTrag.1* is hosted by the Department of Philology, University of Patras – Greece. Its research and other activities are also supported by the dynamic Institute of Digital Innovation of the University Research Centre of Ioannina. The project has been made possible through a funding grant from the Hellenic Foundation for Research and Innovation (HFRI), in the framework of the *2nd Call for Research Projects to Support Post-Doctoral Researchers* (Project Number: 85). The authors of this paper wish to thank the HFRI for its generosity, as well as the members of the project's Advisory Board, Profs. James Diggle (Cambridge), Eric Csapo (Warwick), George W.M. Harrison (Carleton), Antonis Petrides (Open University of Cyprus), Angeliki Syrkou and Ioannis Chatzi-lygeroudis (Patras) for their support in the implementation of the database. With regard to the present paper, in particular, we are grateful to James Diggle and Brady Kiesling (*ToposText*-Project) for their edits and remarks on our earlier version of it.

3 For the term, see Antonopoulos et al. (2023), 20.

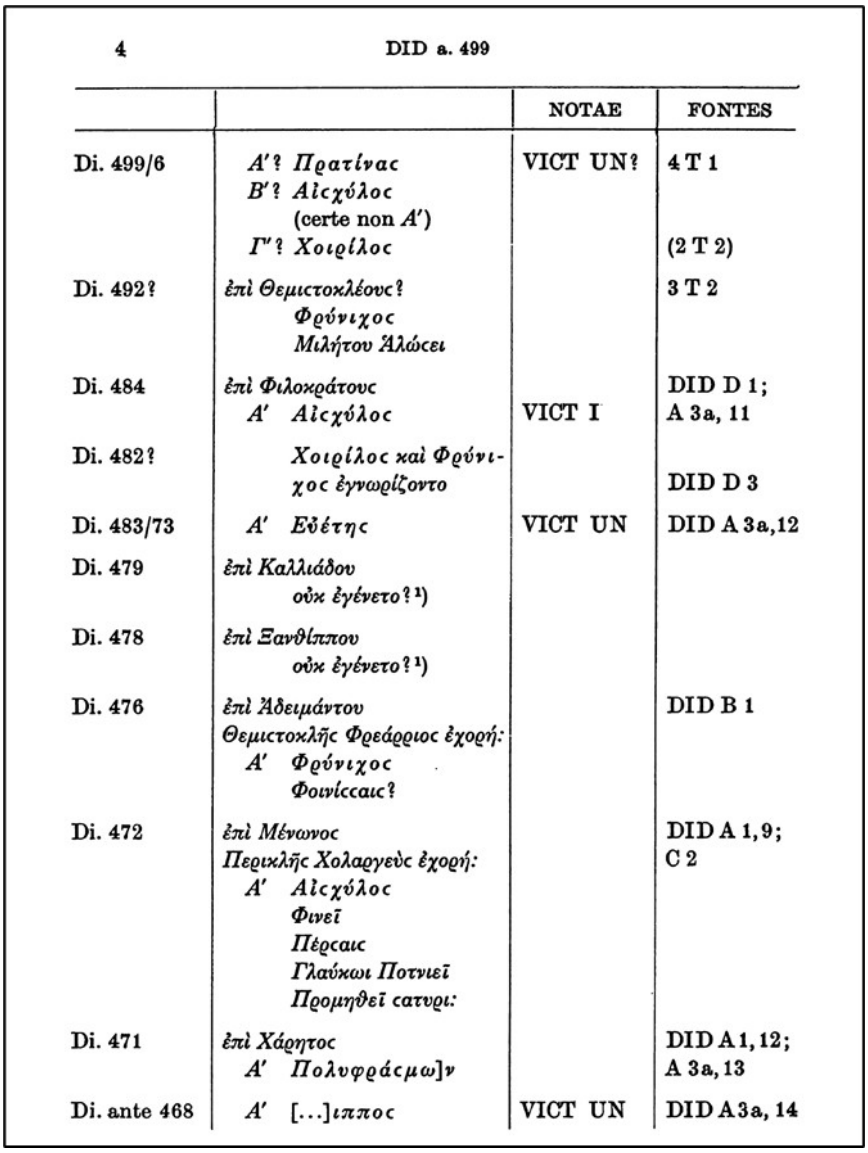
4 For these, see esp. Millis / Olson (2012).

From the early stages of the project, we realised that several of these *testimonia* constitute documentations of events[5] and could, thus, be treated as data with a time dimension.[6] When assembled, these data can contribute significantly to reconstructing the history of ancient Greek theatre. Thus, we decided to create a structured and searchable collection of these data, along with a timeline visualisation and access to the original texts of the testimonies, as a helpful tool for the study of these less-known poets and of the history of Greek drama in general. In this framework, while the known facts about each of these poets are presented with full documentation, they are also correlated with relevant facts of other poets to provide a context for the evolution of Greek drama during the 6th and the 5th centuries BC. In practice, such a collection allows the user to search for and retrieve both individual events related to one or several poets (e.g. participation of one or more poets in a specific tragic contest), as well as groupings of similar events (e.g. first victories at the City Dionysia).

This collection bears similarities to the chronologically ordered list of reconstructed *didascaliae* printed at the beginning of Snell / Kannicht (1986) *TrGF* vol. I (*didascaliae e fontibus restitutae*, pp. 3–21, see Fig. 1). The difference at the material level is that our collection also includes events other than tragic performances, namely important moments in the life and career of these poets extracted from their respective *testimonia*. Regarding utility, unlike the *TrGF* printed edition, where the reader needs to page back and forth through the entire book to consult the sources of the respective didascalic entries, here all data are easily accessible to the user in parallel windows with the opening of the relevant hyperlinks. Furthermore, presenting our collection in a digital medium opens possibilities for alternative orderings and visualisations of the material, as well as for interconnections with other sources available online.

5 Most of these are unique events in terms of their spatiotemporal constituents; nevertheless, a few of the recorded events, such as ‘dramatist X was active (= presenting his plays) in place P at time T’ (see e.g. the case of Euetes Test. 1 [= *Suda* E. 2766] reporting that Epicharmus, Euetes, Euxenides, and Myllus were active six years before the Persian Wars) do not have uniqueness concerning their temporal constituents. For the purposes of this collection of data, our understanding of event is rather loose and includes occurrences or incidents (see Oxford English Dictionary, s.v. ‘event (n.)’, sense II.3.a, https://doi.org/10.1093/OED/1106581244 [last access 15.01.2024]), which are related to fragmentary tragedians and to which the reporting sources attribute directly, or by implication, a more or less precise date. For a concise introduction to the philosophical discussion about events, see Schneider (2005).

6 The term ‘data’ is used here in a very broad meaning; we are actually dealing with *capta*, as Drucker (2011) defines them, that “are not ‘given’ pieces of information able to be recorded and observed”, but pieces of information produced through an interpretative process.



4 DID a. 499

		NOTAE	FONTES
Di. 499/6	*A'*? *Πρατίνας* *B'*? *Αἰσχύλος* (certe non *A'*) *Γ'*? *Χοιρίλος*	VICT UN?	4 T 1 (2 T 2)
Di. 492?	*ἐπὶ Θεμιστοκλέους*? *Φρύνιχος* *Μιλήτου Ἁλώσει*		3 T 2
Di. 484	*ἐπὶ Φιλοκράτους* *A'* *Αἰσχύλος*	VICT I	DID D 1; A 3a, 11
Di. 482?	*Χοιρίλος καὶ Φρύνιχος ἐγνωρίζοντο*		DID D 3
Di. 483/73	*A'* *Εὐέτης*	VICT UN	DID A 3a,12
Di. 479	*ἐπὶ Καλλιάδου* *οὐκ ἐγένετο*?[1])		
Di. 478	*ἐπὶ Ξανθίππου* *οὐκ ἐγένετο*?[1])		
Di. 476	*ἐπὶ Ἀδειμάντου* *Θεμιστοκλῆς Φρεάρριος ἐχορή:* *A'* *Φρύνιχος* *Φοινίσσαις*?		DID B 1
Di. 472	*ἐπὶ Μένωνος* *Περικλῆς Χολαργεὺς ἐχορή:* *A'* *Αἰσχύλος* *Φινεῖ* *Πέρσαις* *Γλαύκωι Ποτνιεῖ* *Προμηθεῖ σατυρι:*		DID A 1, 9; C 2
Di. 471	*ἐπὶ Χάρητος* *A'* *Πολυφράσμω]ν*		DID A 1, 12; A 3a, 13
Di. ante 468	*A'* [...]*ιππος*	VICT UN	DID A 3a, 14

Fig. 1: Extract from the list of the *didascaliae* printed in *TGrF*.

So far, the most characteristic example of the possibilities offered by a properly annotated collection of events in digital format combined with a timeline is *The Chronology of the Marmor Parium* (*CMP*),[7] which is part of the project *Digital Marmor Parium* directed by M. Berti (2014–). This tool encodes carefully the chronological expressions of the *Marmor Parium* related to the recorded events and visualises them in a timeline through *TimelineJS*[8] (see Fig. 2). It encodes events coming from only one source which, being a chronicle, has a structure very similar to that of a timeline and offers concrete dates for the recorded events. The *CMP* does not offer a specific search interface, but only the possibility to download a fully searchable spreadsheet underlying the visualisation. Even so, the structured data generated by the project make it possible to quickly retrieve any information related to a specific subject (e.g. using the search term *poet*). Moreover, the timeline visualisation offers a clear overview of the *Marmor Parium*'s uneven distribution of events in time, which makes it obvious, for example, which time periods have a high or low concentration of events.

7 https://www.digitalmarmorparium.org/chronology.html (last access 15.01.2024). See also Berti (2016).

8 https://timeline.knightlab.com (last access 15.01.2024).

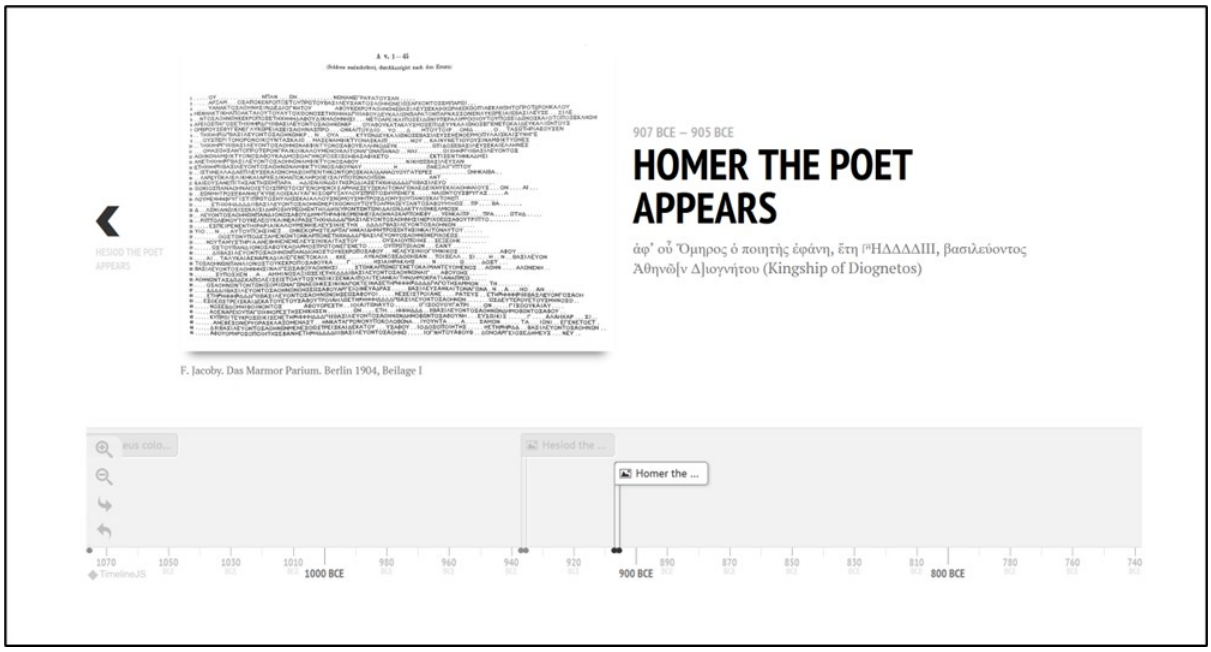


Fig. 2: Extract from the *Digital Marmor Parium*'s Timeline.

In addition, several online sources (mostly educational) offer digital timelines documenting events related to the ancient world. Some of them refer to Greek theatre, for example the *Greek Theatre Timeline*[9] by *The World History Encyclopedia* (which is in turn retrieved from the broader *Ancient History Timeline* of the project). Nevertheless, these timelines focus only on major events, concerning mostly the 'Great Three' tragedians, and do not document the events they are presenting with references to primary sources.

The *FragTrag* project proposes a collection of data and its visualisation as a timeline, with the following targets:

- to provide an easily navigable and complete overview of events related to the fragmentary tragedians of Greek theatre, combining data from various sources;
- to offer an environment that allows for different visualisations based on specific searches;
- to give easy access to the source(s) of each event in the original language and in English translation.

In this article, we are presenting the data-structure, the visualisation, and the storage details of the relevant files. At the moment the *Timeline of Greek Theatre Events* includes data related to the following tragedians: Thespis, Choerilus, Phrynichus, Pratinas, Euetes, Polyphrasmon, Aristias, Mesatus, Aristarchus Tegeates, Euphorion, Ion, and Achaeus. We expect to cover all remaining poets from the 5th century BC by the end of the project (summer 2024).

9 https://www.worldhistory.org/timeline/Greek_Theatre (last access 15.01.2024).

Data-structure

The basic unit of our data structure is the *event*. This might be related to one or more tragedians, a specific dramatic contest, or a particular dramatic performance. These events are or could be, dated, although, obviously, we are not always able to retrieve their dating.

The definition of *events* for our purposes presents two methodological challenges. These arise from the fundamental fact that we are not directly registering events based on primary empirical material, but, rather, we are recording registrations of events as they are represented in our ancient and medieval sources:

1. Given that a significant portion of theatrical life in classical Athens revolved around dramatic contests during major religious festivals, it seems natural to consider the specific contests as a primary organisational unit in a list of theatrical events (e.g. the City Dionysia tragic contest of 483 BC). However, the available material is influenced by structural decisions made by the sources transmitting it, and these decisions do not always use the dramatic contest as the basic unit. For example, the so-called *Victor Lists* (*IG* II.2 2325) documenting the total number of victories of tragic poets use the poet as the basic unit. They record the names of poets in the order of their first victory (without giving a date for that victory), followed by the total number of their victories.[10] Consequently (and often after correlation with other sources) 'the first victory of poet P in the tragic contest of a festival F' is reconstructed as an event involving a single individual. The highly fragmentary nature of our evidence often precludes precisely dating these victories[11] or associating them with specific contests where (potentially) all participating poets are known (as e.g. documented by the so-called *Fasti*).[12] In our *Timeline of Greek Theatre Events*, we adhere to the ordering logic of the sources and register 'first victories' as distinct event units. However, in cases where we can combine information about a poet's first victory with other details related to the same tragic contest, we compile all related pieces of information into the same entry (interconnecting, if necessary, with other *distinct* entries).

2. Given the complexity of theatre contests, which involve the participation and competition of various individuals in different roles (such as poets, *choregoi*, and actors), we often find information about several participants in a specific contest. For instance, at the tragic contest of the City Dionysia in 476 BC, the winner was the tragic poet Phrynichus, alongside his *choregos*, Themistocles.[13] In such cases, there are two possible ways to structure the data: one approach is to break down the complex event into its simplest components, namely the victory of the tragic poet Phrynichus and the victory of the *choregos* Themistocles, recording each as a separate entry while indicating in a machine-readable format that these entries are interconnected. The other way is to record the multifaceted event as a single entry, with

10 See e.g. the portion of the inscription mentioning Euetes, Polyphrasmon and Nothippus: https://fragtrag1.upatras.gr/exist/apps/fragtrag/polyphrasmon/testimonia/Polyphrasmon_TESTIMONIA.xml?action=search&view=page&odd=fragtragodd.odd#3.4.2.2.8.7.4.8.4.2.3 (last access 15.01.2024). For further details on these inscriptional records, see Millis / Olson (2012), 133–149 and 204–207.

11 For instance, the best we can have for the first victory of Polyphrasmon is between 482 and 471 BC (see Polyphrasmon Test. 1: https://fragtrag1.upatras.gr/exist/apps/fragtrag/polyphrasmon/testimonia/Polyphrasmon_TESTIMONIA.xml#Polyphrasmon-test-1 [last access 15.01.2024]).

12 I.e. *IG* II.2 2318, offering a detailed programme for each given contest (at least in the preserved portion of the inscription). For more information, see Millis / Olson (2012), 5–58.

13 See Phrynichus Test. 4 (https://fragtrag1.upatras.gr/exist/apps/fragtrag/phrynichus/testimonia/Phrynichus_TESTIMONIA.xml#Phrynichus-test-4 [last access 15.01.2024]).

its complexity documented in a machine-readable format. In our *Timeline of Greek Theatre Events*, we have chosen the latter approach.

The dating of events is frequently uncertain and a subject of debate. In such instances, we have chosen to provide a range of datings that cover all plausibly proposed dates. While we do not yet explicitly encode this uncertainty into a machine-readable format, we do document it in our human-readable descriptions of each event.

Consistent with the approach adopted for the *Greek Fragmentary Tragedians Online* project, the data collection *Timeline of Greek Theatre Events* structures its data using a TEI/XML-compatible scheme.

All event units are contained within a single <listEvent> element. Each event unit is represented by an individual <event> element, which is supplied with an @xml:id attribute to make it uniquely identifiable.

- Importantly, each <event> is necessarily equipped with an @type, which can hold one or more values identifying the type of event. These are: “tragic-performance”, “tragic-victory”, “choregic-tragic-victory”, “birth”, “death”, or “incident”. We define these values following the categories implied or explicitly used by our ancient sources. By the end of the project we expect to have added some more values and we plan, in a next step, to use the ontological framework “E5 Event” of CIDOC-CRM to represent formally the events we have registered.[14]

- The dating of the event is specified using an @when for exact dates, @notBefore for *post quem* dates (earliest possible dates), and @notAfter for *ante quem* dates (latest possible dates), or even a combination of @notBefore and @notAfter. The values of these attributes follow the format of *W3C XML Schema Part 2: Datatypes Second Edition, 3.2.7*.[15] To aid in the chronological sorting of the entries, we also add an @datingPoint, which takes the value of @when, @notBefore, @notAfter, or, in cases where both @notBefore and @notAfter are present, the value of @notBefore.

- Each <event> element is also necessarily equipped with an @source containing references to the *testimonia* documenting the specific event. Each <event> is optionally equipped with an @corresp with a value pointing to another <event>, in cases where two different events may be considered as one and the same.[16]

- Each <event> element includes a further <desc> element (description), a concise text that presents the event in all its aspects and documents any uncertainties in a human-readable format.[17] Personal names, play titles, and names of festivals mentioned in the <desc> are annotated using the <persName>, <title>, and <name @type=“feast”> elements, respectively. Each <persName> element is equipped with an @role, which accepts values indicating the person’s role in the event, such as “poet” or “choregos”. It also includes an @sameAs, holding a pointer to the xml:id of the person in the *FragTrag*-database, and an @ref with a link to

14 https://www.cidoc-crm.org/entity/e5-event/version-7.1.1 (last access 31.03.2023). For a comprehensive review of existing event ontologies, see Piryani et. al. (2023).

15 https://www.w3.org/TR/xmlschema-2/#dateTime (last access 15.01.2024).

16 For instance, Choerilus first participation in the City Dionysia tragic contest was recorded by Test. 1 (https://fragtrag1.upatras.gr/exist/apps/fragtrag/choerilus/testimonia/Choerilus_TESTIMONIA.xml#Choerilus-test-1 [last access 15.01.2024]) may coincide with his (first?) victory recorder by Test. 3 (https://fragtrag1.upatras.gr/exist/apps/fragtrag/choerilus/testimonia/Choerilus_TESTIMONIA.xml#Choerilus-test-3 [last access 15.01.2024]).

17 We plan to implement a fully machine-readable format in future.

the relevant entry in *Wikidata*.[18] Each <title> element comes with an @type, allowing for three possible values: "tragedy", "satyr-drama", or "comedy". Additionally, all <name> elements, besides the @type="feast", include an @subtype with "city-dionysia" or "lenaea" as potential values (as well as an @ref with a link to the relevant entry in *Wikidata*).

To illustrate how we encode these events, we give here as an example the encoding for the City Dionysia tragic contest of 467 BC (Fig. 3).

```
<event xml:id="ev00013" type="tragic-performance;tragic-victory" when="-0467"
   datingPoint="-0467"
   source="#Pratinas-test-2 #Pratinas-test-3 #Aristias-test-3 #Aristias-test-4 #Polyphrasmon-test-3 #Polyphrasmon-test-4">
   <desc>
      <persName role="tragedian" sameAs="#Aeschylus" ref="https://www.wikidata.org/wiki/Q40939">Aeschylus</persName>
      with the plays
         <title type="tragedy" rend="italic">Laius</title>,
         <title type="tragedy" rend="italic">Oedipus</title>,
         <title type="tragedy" rend="italic">Seven against Thebes</title>
         and <title type="satyr-drama" rend="italic">Sphinx</title>,
      <persName role="tragedian" sameAs="#Aristias" ref="https://www.wikidata.org/wiki/Q3622659">Aristias</persName>
      with the plays
         <title type="tragedy" rend="italic">Perseus</title>,
         <title type="tragedy" rend="italic">Tantalus</title>,
         an unknown tragedy,
         and the satyr drama <title type="satyr-drama" rend="italic">Wrestlers</title>,
      and <persName role="tragedian" sameAs="#Polyphrasmon" ref="https://www.wikidata.org/wiki/Q3907308">Polyphrasmon</persName>
      with a <title type="tragic-tetralogy" rend="italic">tetralogy on Lycurgus</title>,
      competed at the tragic contest of the <name type="feast"
         subtype="city-dionysia" ref="https://www.wikidata.org/wiki/Q21707866">City
         Dionysia</name>.
      Aeschylus won the first and Aristias the second prize. Polyphrasmon occupied the third place. </desc>
</event>
```

Fig. 3: Extract from our XML data collection of *events*.

In the above example, all information about the specific tragic contest is included within a single element <event> (which in our system is in turn included within a single <listEvent> element), which is given a unique identifier, an @xml:id ("ev00013"). The event is further defined by an @type as a "tragic-performance" and "tragic-victory" (the latter referring to the documented victory of Aeschylus), dated with the @when to "-0467" (equivalent of 467 BC) and chronologically arranged in relation to the other events of the *Timeline of Greek Theatre Events* with the aid of an @datingPoint (in this case it gets the same value as the @when, namely "-0467"). In addition, using an @source the event is interconnected with its sources, namely Pratinas Test. 2 and 3, Aristias Test. 3 and 4 and Polyphrasmon Test. 3.

Further details of the event are provided through an additional element <desc> (included within the <event>). The three playwrights participating in this event, Aeschylus, Aristias and Polyphrasmon are encoded with a <persName> element each, which is further defined with an @role (taking the value "tragedian") and an @sameAs pointing to each of these authors' xml:id in our project (respectively "#Aeschylus" "#Aristias" and "#Polyphrasmon"). A further @ref links to each poet's entry in *Wikidata*. The plays with which each tragedian participated in the contest are encoded with the element <title>, which is further defined by the @type (in the case of Polyphrasmon's *Tantalus*, for instance, the attribute takes the value "tragedy", whereas for his *Wrestlers* the value "satyr-drama"). Finally, we specify the festival in whose framework the event took place using the element <name>, which is defined with @type as a "feast" and with the @subtype as the "city-dionysia".

18 https://www.wikidata.org (last access 15.01.2024).

The above encoding results in the following visualisation (Fig. 4):

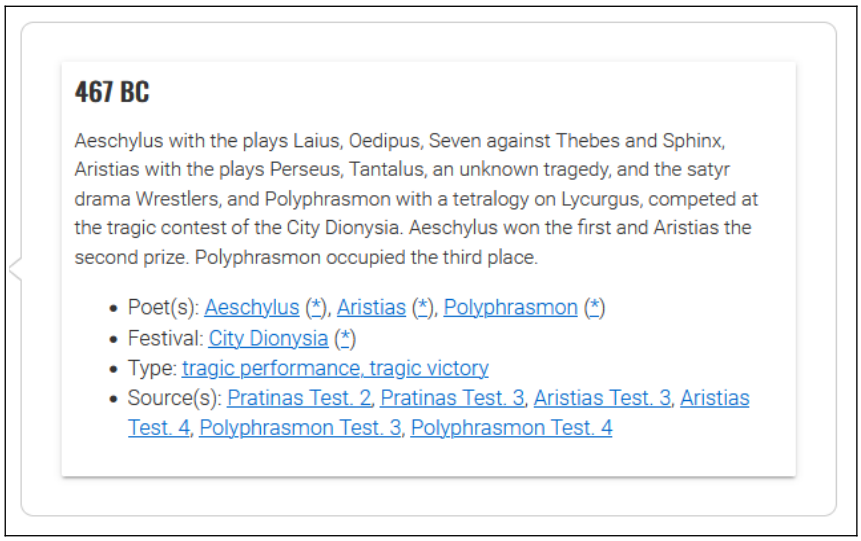


Fig. 4: Extract from the *Timeline of Greek Theatre Events* presenting the City Dionysia tragic contest of 467 BC.

The visualisation

The visualisation of our data collection takes the form of a simple, vertically oriented timeline (*Timeline of Greek Theatre Events*, Fig. 5), on which by default all event entries appear in chronological order on the basis of their @datingPoint attribute. If the @datingPoint value of two entries coincides the entries are sorted according to the order in which they are placed in the XML-document.

We have chosen vertical scrolling for presenting the events as it is a widely adopted and expected behaviour across various platforms and devices. Considering general web viewing patterns, researchers have observed that nowadays most users do not utilize fully horizontal scrolling, as most of the viewing content is on the left half of the screen.[19] Moreover, users tend to dislike horizontal scrolling, whereas they are accustomed to vertical scrolling on websites, mobile apps, and other digital interfaces, creating a consistent user experience.[20]

19 See Fessenden (2017).

20 See Nielsen (2018).

Fig. 5: Extract from the *Timeline of Greek Theatre Events*.

In the visualisation of each entry, the attributed dating appears as the heading. This may be a single date, with or without attributes (such as “after” or “before”), or a date-span indicating an approximation. The main body of the visualisation contains the description of the event.

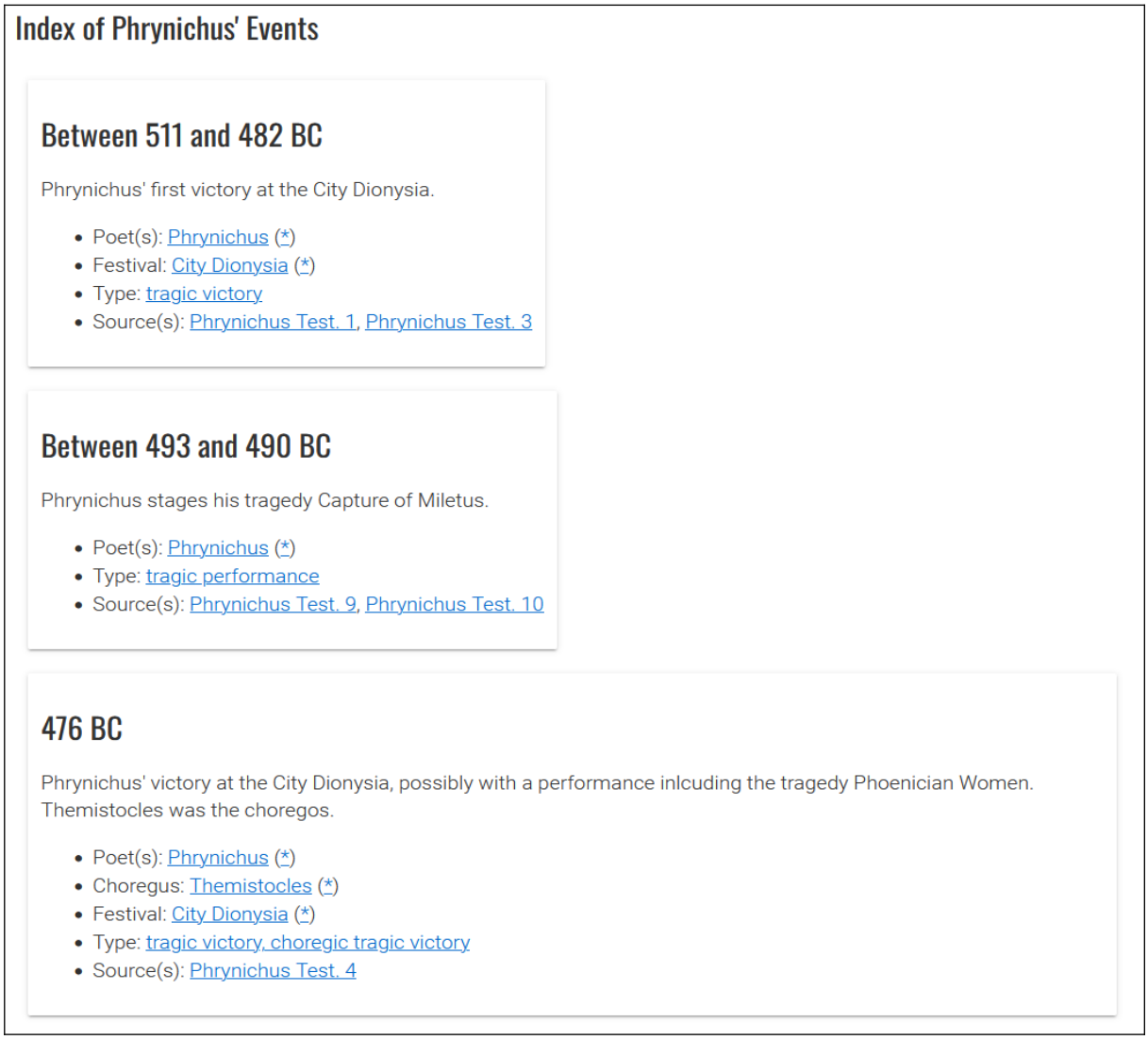


Fig. 6: Visualised aggregation of events related to Phrynichus.

The description is followed by a list of entries. Among them, "Poet(s)", "Choregus" and "Festival" (where applicable) are extracted from the <desc> element and correspond to the values assigned to, respectively, the <persName type="tragedian">, <persName type="choregus"> and <name type="feast"> elements. Personal names and names of feasts appear as interactive links. Upon clicking, these links lead to a new visualisation aggregating events related to the specific person or feast (see e.g. Fig. 6). Additionally, these elements are interconnected with corresponding entries in *Wikidata*. Users are directed to these *Wikidata* entries by clicking on the asterisk symbol (*) next to the annotated term. Additional entries, "Type", "Source" and "Corresponding event(s)" are extracted directly from the <event> element and correspond to the values assigned to the @type, @source, and @corresp attributes. The visualisation of these values includes interactive links with the following behaviours upon clicking:

- each "type" value links to a new visualisation aggregating events categorised under the same type.
- each "source" value links to the referred *testimonium* in our edition (Fig. 7).
- each "corresp" value links to the corresponding event.

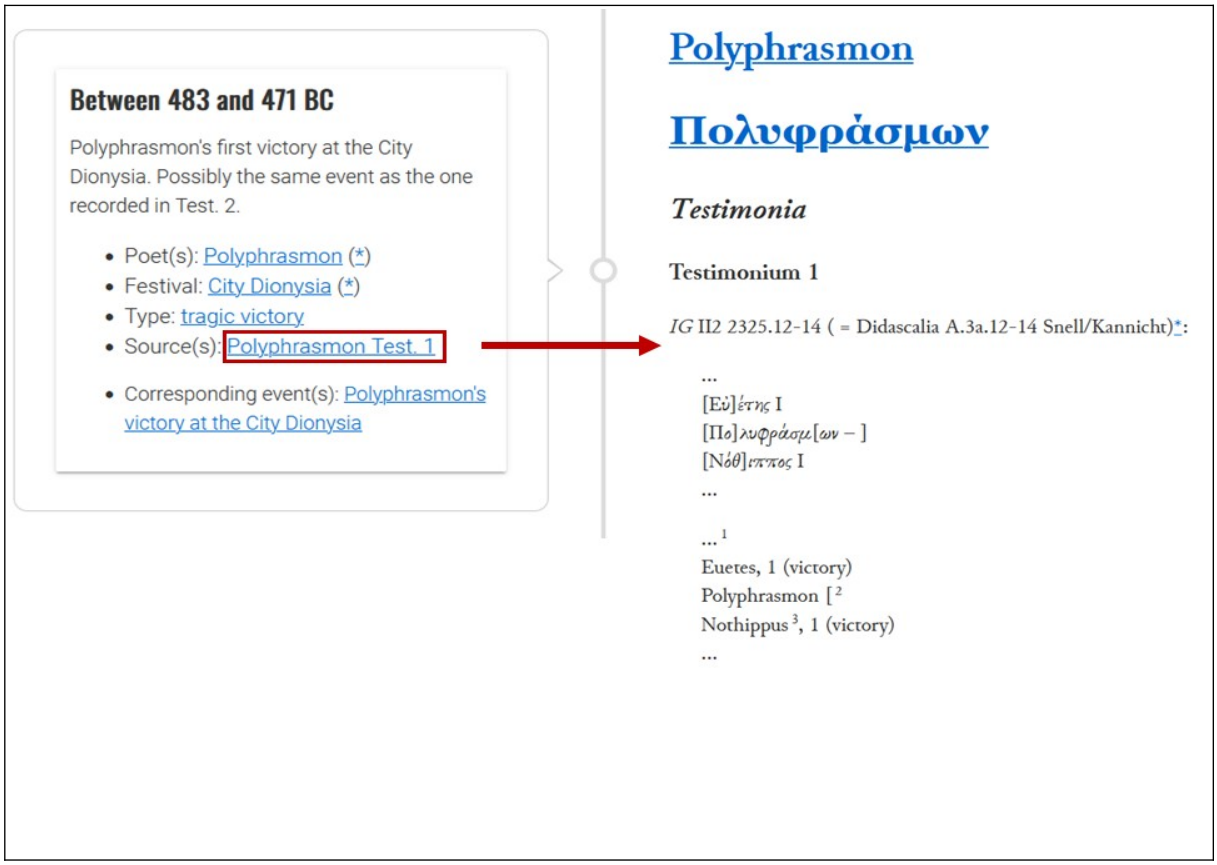


Fig. 7: From the *Timeline of Greek Theatre Events* to the original source documenting Polyphrasmon's first victory at the City Dionysia.

The visualisation is generated through a simple transformation script combining XQuery (app.xql file), HTML (events.html, indexe.html, indexesameas.html and indexesubtype.html files), and JavaScript (timeline.min.js file).

With regard to the four HTML files used to visualise the data:

1. *events.html* implements the *Timeline of Greek Theatre Events of Events* via the call to the *app:fragtragevents* function (see in the app.xql file) and the use of timeline.min.js, a JavaScript to produce the format of the Timeline.

2. *indexe.html* gathers the different types of events via the call to the *app:searchevents* function (from the app.xql file) and the use of JavaScript to produce the title of the document.

3. *indexesameas.html* filters the different personal names, so as to visualise the relevant events for each person via the call to the *app:searcheventsp* function (from the app.xql file) and the use of JavaScript to produce the title of the document.

4. *indexesubtype.html* filters the different types of festivals, so as to visualise the relevant events for each festival via the call to the *app:searcheventsf* function (from the app.xql file) and the use of JavaScript to produce the title of the document.

The app.xql file, includes the following functions that process the source XML-file of the events and, according to the XML-tags, these functions extract all the appropriate information to be visualised for the user:

1. *app:fragtragevents* selects all events from the database and sorts them according to each event's @datingPoint. In this function, there is a call to another function, the *app:show-events* function (see below).

2. *app:searchevents* selects the events belonging to each event's type and sorts them according to the @datingPoint attribute. In this function, there is also a call to the *app:show-events* function (see below).

3. *app:searcheventsp* selects the events related to each person (tragedians, choregoi and actors) and sorts them according to the @datingPoint. In this function, there is also a call to the *app:show-events* function (see below).

4. *app:searcheventsf* selects each festival's events (City Dionysia or Lenaea) and sorts them according to the @datingPoint. In this function, there is also a call to the *app:show-events* function (see below).

5. *app:show-events* organises the selected data, formats them in the appropriate way to be used in the HTML page, and calls the *app:get-event* function (see below).

6. *app:get-event* is the most important function as it gives the data its final format, creates the hot-words with links to the indexes (the various event aggregations) and the *FragTrag*'s basic texts (*fragmenta*, *testimonia*, biographies), and checks if there are any restrictions or conditions to take into account while building each event's card, as shown to the user.

Data storage and accessibility

The XML file which contains the data, the XQL file which extracts the relevant information, along with the HTML and the JavaScript used to generate the visualisation, are stored in a location under our database's "Documentation" menu and are openly accessible. The direct link is https://fragtrag1.upatras.gr/exist/apps/fragtrag/resources/fragtrag_events.zip (last access 15.01.2024; all files are concentrated in a compressed file).

Finally, the *Timeline of Greek Theatre Events of Events* tool can be accessed at https://fragtrag1.upatras.gr/exist/apps/fragtrag/events.html (last access 15.01.2024). Hosting these resources (as well as our entire database) on the servers of the University of Patras, a public institution, ensures their long-term maintenance and open-access availability.

Potential expansion and reusability of data

Our data collection currently focuses on material related to fragmentary tragedians, which derives from the *Greek Fragmentary Tragedians Online* edition of their *testimonia*. It has the potential to be expanded to well-known playwrights such as Aeschylus, Sophocles, Euripides. In fact, we have already included some material from them in cases of events where their own paths meet with those of the fragmentary tragedians. For instance, we know from our sources that in a tragic contest during the

70th Olympiad (i.e. between 500 and 497 BC) Aeschylus competed against Pratinas and Choerilus,[21] and this information has been added to our Timeline of Greek Theatre Events. Our plan is to expand the material pertaining to the 'Great Three' so as to provide an all-inclusive chronological tool for the events of Greek tragic theatre as a whole. But that will have to wait until we have completed our work with the fragmentary tragedians.

Furthermore, given the specialised nature of our data and its format in TEI/XML – a format prevalent in Digital Humanities and easily convertible to formats like JSON – combined with its open accessibility, we believe that databases of cultural or historical events in the ancient Greek world could effectively reuse and repurpose our data in their projects.

21 See Choerilus Test. 2 (https://fragtrag1.upatras.gr/exist/apps/fragtrag/choerilus/testimonia/Choerilus_TESTIMONIA.xml#Choerilus-test-2 (last access 15.01.2024).

Sources

Online-Sources

https://www.cidoc-crm.org/entity/e5-event/version-7.1.1 (last access 31.03.2024).

https://www.digitalmarmorparium.org/chronology.html (last access 15.01.2024).

https://fragtrag1.upatras.gr/ (last access 15.01.2024).

https://www.oed.com/dictionary/ (last access 15.01.2024).

https://timeline.knightlab.com (last access 15.01.2024).

https://www.worldhistory.org/timeline/Greek_Theatre (last access 15.01.2024).

https://www.wikidata.org (last access 15.01.2024).

Text editions

Snell / Kannicht (1986): B. Snell (ed.), Tragicorum Graecorum Fragmenta. Vol. 1. Didascaliae tragicae. Catalogi tragicorum et tragoediarum. Testimonia et fragmenta tragicorum minorum (revised by R. Kannicht), Göttingen 1986.

References

Antonopoulos et al. (2023): A. P. Antonopoulos / S. Chronopoulos / N. Ntaliakouras / P. Taktikou / A. Psomiadou / I. Markelis, Developing a Database for the Greek Fragmentary Tragedians, Digital Classics Online 9 (2023), 15–29, https://doi.org/10.11588/dco.2023.9.95214 (last access 15.01.2024).

Berti (2014–): M. Berti, The Chronology of the Marmor Parium, in: Digital Marmor Parium, hosted by the University of Leipzig, 2014–, https://www.digitalmarmorparium.org/chronology.html (last access 15.01.2024).

Berti (2016): M. Berti, The Digital Marmor Parium, in: M. Berti (ed.), Epigraphy Edit-a-thon. Editing chronological and geographic data in ancient inscriptions, April 20–22, 2016, Leipzig https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa-221511 (last access 31.03.2024).

Drucker (2011): J. Drucker, Humanities Approaches to Graphical Display, Digital Humanities Quarterly 5 (2011), https://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html (last access 15.01.2024).

Fessenden (2017): T. Fessenden, Horizontal Attention Leans Left, in: Nielsen Norman Group, 2017, https://www.nngroup.com/articles/horizontal-attention-leans-left/ (last access 15.01.2024).

Millis / Olson (2012): B. Millis / S. D. Olson, Inscriptional Records for the Dramatic Festivals in Athens. IG II2 2318-2325 and Related Texts, Leiden 2012.

Nielsen (2018): J. Nielsen, Scrolling and Scrollbars, in: Nielsen Norman Group, 2018, https://www.nngroup.com/articles/scrolling-and-scrollbars/ (last access 15.01.2024).

Piryani et. al. (2023): R. Piryani1 / N. Aussenac-Gilles / N. Hernandez, Comprehensive Survey on Ontologies about Event, in: SEMMES'23: Semantic Methods for Events and Stories co-located with ESWC, May 29 2023, https://ceur-ws.org/Vol-3443/ESWC_2023_SEMMES_XPEvent.pdf (last access 31.03.2024).

Schneider (2005): S. Schneider, Events, in: Internet Encyclopedia of Philosophy, https://iep.utm.edu/events (last access 15.01.2024).

Figure References

Fig. 1: Extract from the list of the *Didascaliae* printed in *TGrF.*

Fig. 2: Extract from the *Digital Marmor Parium*'s Timeline, https://www.digitalmarmorparium.org/chronology.html (last access 15.01.2024).

Fig. 3: Extract from our XML data collection of *events*.

Fig. 4: Extract from the timeline presenting the City Dionysia tragic contest of 467 BC, https://fragtrag1.upatras.gr/exist/apps/fragtrag/events.html (last access 15.01.2024).

Fig. 5: Extract from the *Timeline of Greek Theatre Events*, https://fragtrag1.upatras.gr/exist/apps/fragtrag/events.html (last access 15.01.2024).

Fig. 6: Visualised aggregation of events related to Phrynichus, https://fragtrag1.upatras.gr/exist/apps/fragtrag/indexesameas.html?p=Phrynichus (last access 15.01.2024).

Fig. 7: From the *Timeline of Greek Theatre Events of Events* to the original source documenting Polyphrasmon's first victory at the City Dionysia.

Author contact information[22]

Andreas P. Antonopoulos
Assist. Professor of Ancient Greek Literature
Department of Philology
University of Ioannina
Ioannina 45100, Greece
E-Mail: andreas.antonopoulos@uoi.gr

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

DIGITAL CLASSICS ONLINE

Linked Ancient World Data: Implementation, Advantages, and Barriers

Sarah Middle

Abstract: Linked Data technologies can be hugely beneficial for modelling and representing the complexity and nuances of Humanities data, while also facilitating discoverability and reuse. In the Ancient World domain, their implementation appears to be relatively mature, with many tools and resources available that demonstrate the benefits of this approach in different ways; however, barriers to Linked Data production and consumption still remain. In this article, I provide some brief background information on the principles behind a Linked Data approach, before surveying the literature on Linked Ancient World Data. Using examples provided by existing initiatives, I explore how these tools and resources work together to widen access to digital materials, reveal connections between them, and facilitate collaboration, while also recognising obstacles faced by both users and producers. Issues discussed include identification and disambiguation, modelling complexity, communicating uncertainty, reasoning, making connections, discoverability and usability, awareness and training, openness and collaboration, and sustainability. My conclusions point towards the key role of social and institutional factors in barriers to Linked Data uptake, instead of a series of solely technological problems requiring technological solutions.

1. Introduction

Of the various approaches available for modelling and publishing Humanities data, Linked Data is perhaps the most effective for representing its complexity and nuance, while also facilitating discoverability and reuse. The term 'Linked Data' refers to a set of technologies that can be used to describe entities, such as places, people, or objects, and connect them based on features they have in common. Its rich semantic descriptions, disambiguation capabilities, and interoperability can unlock opportunities to address new research questions and reveal previously undiscovered relationships between entities. However, uptake of Linked Data among Humanities researchers has been relatively low, compared with approaches such as relational databases or text encoding. In the Ancient World domain, however, Linked Data implementation is relatively mature. Additionally, as study of the Ancient World encompasses multiple disciplines, including Archaeology, Art, History, Literature and Philosophy, it might be considered a microcosm of the Humanities, rendering the conclusions from this article more broadly applicable.

In this article, I present a survey of Linked Ancient World Data (LAWD) implementation, illustrating the advantages of and barriers to this approach. I start by providing some background to the areas of Linked Data and knowledge organisation, before discussing in detail how these technologies have been applied to tools and resources that facilitate study of the Ancient World. To conclude, I will high-

light the interplay between technological and social/cultural factors in relation to the advantages of and barriers to Linked Data implementation for Ancient World research.[1]

2. Background: Linked Data and Knowledge Organisation

2.1 Linked Data Principles

Linked Data technologies are used to describe digital resources, facilitating machine-readable connections between datasets, with the potential to transform the way they are consumed – leading to new insights that would not have been possible previously. Applying a Linked Data approach facilitates the implementation of the Semantic Web: a 'Web of Data', where online resources are semantically linked in a machine-readable way, based on the information about them.[2] For a dataset to be accurately described as Linked Data, it must comply with Berners-Lee's Linked Data principles:

> "1. Use URIs as names for things.
>
> 2. Use HTTP URIs so that people can look up those names.
>
> 3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL).
>
> 4. Include links to other URIs, so that they can discover more things."[3]

Firstly, for a resource to be described using Linked Data, that resource must have a unique and persistent identifier, known as a Uniform Resource Identifier (URI) (point 1). A URI can refer to any concept or entity in the physical or digital world (i.e., it does not necessarily denote a digital resource), but in a Linked Data representation, a URI must start with 'http://' (point 2). 'HTTP' refers to 'Hypertext Transfer Protocol', which is the standard mechanism by which digital resources are accessed via the World Wide Web.[4]

Each piece of information about the resource (e.g., its title, creator, or location) is expressed using three components: a subject (e.g., the resource, expressed using its URI), an object (e.g., the resource's creator, expressed using their URI), and a predicate (the relationship between the subject and object, e.g., 'has creator', expressed using the URI that describes this property). If the subject or object cannot be identified using a URI (e.g. the title of a book, or an object's latitude/longitude), a literal value, such as a number or text string, can be used instead.[5] The three components are expressed as a 'triple' of the form <subject> <predicate> <object>, using the Resource Description Framework (RDF), the standard format for expressing Linked Data (point 3). RDF data can be queried using the SPARQL Protocol and RDF Query Language (SPARQL) (point 3) via a SPARQL endpoint. Both RDF and

1 This work was supported by the Arts and Humanities Research Council (AHRC) through the Consortium for Humanities and Arts South-east England (CHASE) Doctoral Training Partnership (ref. AH/L503861/1). Thanks are also owed to my PhD supervisors Elton Barker, Phil Perkins, Mathieu D'Aquin and Alessandro Adamou, for their advice and support in carrying out the research that led to this article.

2 Berners-Lee (1998).

3 Berners-Lee (2010).

4 W3C Network Working Group (2004).

5 van Hooland / Verborgh (2014), 48.

SPARQL are open standards, "reusable agreements that make it easier for people and organisations to publish, access, share and use better quality data".[6]

As more resources are described using RDF, connections build up between triples originating from multiple sources, which share the same URI as either the subject or object (point 4).[7] Such links are facilitated through provision of authority files, usually managed by well-known, trusted institutions, such as national libraries, which serve as central and authoritative sources of information about particular topics.

It is only possible to take full advantage of this interlinking when the datasets concerned are openly available via the World Wide Web. In response, Berners-Lee developed a five-star model for Linked *Open* Data (LOD):

> "★ Available on the web (whatever format) but with an open licence, to be Open Data
>
> ★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table)
>
> ★★★ as (2) plus non-proprietary format (e.g. CSV instead of excel)
>
> ★★★★ All the above plus, Use [sic] open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
>
> ★★★★★ All the above, plus: Link your data to other people's data to provide context"[8]

The five-star model prioritises openness, with full compliance requiring that the dataset meets Berners-Lee's original Linked Data principles. An extension to improve data reusability has since been proposed by Hyvönen et al., which adds sixth and seventh stars where a dataset has been adequately described:

> "– The 6th star is given if the schemas (vocabularies) used in the dataset are explicitly described and published alongside the dataset, unless the schemas are already available somewhere on the Web.
>
> – For the 7th star, the quality of the dataset against the schemas used in it must be explicated, so that the user can evaluate whether the data quality matches her needs."[9]

In a similar vein, the more broadly applicable *FAIR* data management principles state that research data should be Findable, Accessible, Interoperable and Reusable.[10] Although these principles were originally developed with the Sciences in mind, their application to Ancient World research (and Humanities disciplines more broadly) are likely to improve usability of the resulting data, as well as that of the tools and resources through which it is accessed. Findability and Accessibility require persistent

6 Open Data Institute (2018).

7 van Hooland / Verborgh (2014), 49.

8 Berners-Lee (2010).

9 Hyvönen et al. (2014), 227.

10 Wilkinson et al. (2016).

identifiers, such as URIs, while Interoperability and Reusability require open, non-proprietary standards. As such, Linked Open Data implemented using the above seven-star model should also comply with the *FAIR* principles.

More recently, work led by Robert Sanderson and championed by the *Linked Art*[11] initiative has sought to implement and advocate for 'Linked Open *Usable* Data' (LOUD). The five LOUD 'Design Principles' are aimed at data scientists and ontologists, and focus on usability by developers, to achieve a compromise between rich semantic description and comprehensibility by others, ultimately facilitating data reuse and interoperability. These principles recommend providing data that is "The right Abstraction for the audience", with "Few Barriers to entry", that is "Comprehensible by introspection", that includes "Documentation with working examples", and that has "Few Exceptions, instead many consistent patterns".[12]

2.2 Authorities and Ontologies

In addition to each individual entity in a Linked Data resource being associated with a URI, it is often helpful to define them in relation to specific classification systems. For example, Euripides might be defined as a person, author, or playwright, while Classical Athens might be defined as a place, settlement, or *polis* (city-state). Such classifications are provided by Knowledge Organisation Systems (KOS), such as thesauri, which express a series of conceptual terms and the relationships between them, often organised using some form of hierarchy.[13]

Ontologies are like thesauri, but with a greater level of expressivity. In this context, the word 'ontology' refers to a type of controlled vocabulary whose structure facilitates complex relationships between terms, which cannot be represented solely via a hierarchical format.[14] Ontologies contain classes, used to classify the subjects or objects of triples, as in the Euripides and Athens examples above, as well as properties, used as the predicates that define the relationships between them. Using terms from a well-defined ontology enhances the machine-readability of resources and enables tools and systems to search more intelligently based on meanings and relationships rather than arbitrary keywords. Several broader, more extensive, resources provide both authority files and ontologies that describe them, including *Dbpedia*[15] and *Wikidata*[16]. *Wikidata* in particular has been used widely in Digital Humanities projects, largely due to its broad coverage and freely available reconciliation tools, although it is usually combined with more subject-specific data sources.[17] Some projects have additionally used *Wikidata* to store connections between entities such as objects, places, people and chronological periods, creating new concepts and properties as required. Among the advantages of using such a platform is its provision of links between representations of the same concept in different ontologies, thereby facilitating interoperability.[18]

The benefits of a Linked Data approach therefore include richer description of digital objects, leading to more effective integration of multiple collections and datasets. Together, these information sources form an infrastructure on which to build and link other digital tools and resources, both within and outside the organisation that produced them. This process facilitates the integration of previously separate

11 https://linked.art/ (last access 20.12.2023).

12 Delmas-Glass / Sanderson (2020), 22; Linked Art Contributors (n.d.-b).

13 International Organization for Standardization (2011), 2.62; Mayr et al. (2016).

14 Hughes et al. (2016), 163.

15 https://www.dbpedia.org/ last access 20.12.2023).

16 https://www.wikidata.org/ (last access 20.12.2023).

17 Zhao (2023).

18 Schmidt et al. (2022), 341.

datasets, with the potential to transform research by providing insights that could not have been discovered by looking at each one in isolation. However, when Linked Data is applied in practice, various barriers emerge, as demonstrated through my discussion on its implementation in Ancient World research in the following sections.

3. Linked Ancient World Data in Practice

There are multiple examples that serve to demonstrate how Linked Data technologies have been applied to Ancient World research, with these examples themselves illustrating many of the advantages of, and barriers to, this approach. Topics covered in this section include technological issues, such as identification and disambiguation, modelling complexity, reasoning over Linked Datasets, and making connections between them. I will then go on to discuss broader social and cultural issues relating to the implementation of Linked Data technologies for Ancient World research. These include discoverability and usability of the resulting resources, researchers' awareness of Linked Data and related training, and the benefits and obstacles Linked Data can bring to the key areas of openness, collaboration, and sustainability. Initiatives included in my examples were selected due to their relatively high uptake by the research community, as evidenced by their scale, as well as the frequency with which they are mentioned in the literature. Some initiatives have broader applicability beyond the Ancient World context, particularly those from the cultural heritage domain.

3.1 Identification and Disambiguation

When Humanities researchers are exploring a topic using a digital tool or resource, their search terms usually revolve around particular concepts, or "contextual entities", categorised by Lee as Object, Agent, Occurrence, Purpose, Time, Place, Form of expression, Concept or Abstraction, and Relationship.[19] Accurate identification and description of these concepts is intrinsic to effective digital representation that facilitates their discovery. Linked Humanities Data tools, resources, and initiatives, including those relating to the Ancient World, often revolve around one of these contextual entities, while incorporating their relationships to others. In the following examples, Places, Agents (people), Time, and Objects are identified using URIs, providing a unique and consistent mechanism for referring to them, as well as disambiguating those with similar names and uniting equivalent terms in multiple languages. Place (and, more recently, time) URIs are often included in gazetteers, with person URIs are often included in prosopographies, and various academic and cultural heritage resources providing URIs for objects. Assigning URIs to digital resources is recommended by the first of Berners-Lee's Linked Data principles (section 2.1, above); even if this is the only Linked Data principle followed by a project, it can nonetheless assist interoperability, as well as the understanding of terminology by both humans and machines.

Place is a key component of research involving historical events or people, or the movement of objects and materials. Real and mythological places appear in art or are mentioned in literature, demonstrating that place additionally permeates many research topics with a less obvious geographic component. *GeoNames*[20] provides URIs for places throughout the world, integrating geographic data from multiple official sources, such as the National Geospatial-Intelligence Agency and Ordnance Survey.[21] Its near-comprehensive coverage and open licence have made it attractive to many Digital Humanities projects that incorporate a spatial element; however, it was designed to represent the world in its current state. Although alternative names for a place are permitted (and can include historic names), it is not pos-

19 Lee (2011), 106.

20 https://www.GeoNames.org/ (last access 20.12.2023).

21 Geonames (n.d.).

sible to associate them with dates or boundary changes. Berman et al. and Simon et al. suggest that one way to address this issue is to produce suitable historical gazetteers that link places to their modern equivalents in *GeoNames*.[22] This recommendation emphasises the importance of specialist resources that are semantically rich enough for Humanities research, while recognising the value of connection to major information sources.

One such specialist resource is *Pleiades*[23], a gazetteer providing persistent URIs for ancient places, which stemmed originally from the digitisation of the *Barrington Atlas of the Greek and Roman World*,[24] but has evolved into a community-driven resource with an increasingly broad geographical scope. *Pleiades* URIs are additionally linked to *GeoNames* URIs, where an appropriate place record is available;[25] however, *Pleiades* is more fluid in its definition of 'place'. Its creators were influenced by Tuan's conception of 'place' as "a center of meaning constructed by experience", with most places lying on a spectrum between "points in a spatial system" and "strong visceral feelings".[26] Unlike *GeoNames*, a *Pleiades* place need not always correspond to a physical location, which provides the flexibility to include mythological places, or those whose location is unknown.[27] *Pleiades* is therefore better suited to represent the multiplicity of places encountered in Humanities research. Furthermore, to provide historical nuance and disambiguation, *Pleiades* permits multiple names to be associated with each place, with each name having a start and end date.[28] Researchers are therefore able to identify historical places with greater precision, rather than approximating to their counterparts in the contemporary world. As such, *Pleiades* URIs have been used in multiple digital tools and resources, most notably *Pelagios* (discussed in section 3.5, below).

While initiatives to identify places using URIs are particularly mature, significant progress has also been made to identify people in a similar way. People appear in a wide variety of contexts in ancient sources. These include creators of objects and authors of texts, as well as those depicted or mentioned within them. Some might have existed in real life, while others are either fictional, or their existence cannot be confirmed with certainty. There are inherent complexities in the digital representation of person information, for which Linked Data can be a particularly effective solution. For example, multiple people often have the same name, an issue that can be addressed by assigning a distinct URI to each individual,[29] which can be aligned if further evidence is discovered.[30] Similarly, an individual can have many names, or epithets, all of which can be linked to the same URI, potentially with further information about the context in which each name is used. Person URIs can additionally be linked to relevant places, chronological periods, and events, as well as other people. Two key sources of linked person data, *VIAF* and *SNAP:DRGN*, are discussed below (section 3.5) in the context of connecting resources.

Yet more complexity is inherent in the modelling of our next contextual entity, time. While many disciplines incorporate the modelling of time as an absolute concept, where exact days, hours, minutes, seconds and beyond can be identified with absolute precision, this is not the case when representing information about the premodern world. Dates must often be reconciled to different calendrical sys-

22 Berman et al. (2016), 124; Simon et al. (2016b), 107.

23 https://pleiades.stoa.org/ (last access 20.12.2023).

24 Elliott / Gillies (2009), para. 41.

25 Simon et al. (2016b), 102.

26 Tuan (1975), 152.

27 Gillies (2015).

28 Schneider et al. (2018), 15.

29 Bodard et al. (2017), 29.

30 Varga et al. (2018), 39.

tems (if they are provided at all). More often, dates are unclear or unavailable, in which case a relative chronology must be applied, based on aspects such as an object's style, typology or context. Such definitions are frequently contested and are subject to change if new evidence comes to light. As these characteristics vary by geographical region, each chronological period defined in this way is bound up with the concept of place; for example, the Greek Bronze Age spanned approximately 3200 to 1050 BCE, while in Britain the Bronze Age is dated to between 2500 and 700 BCE. Initiatives and resources that are primarily focused on describing place therefore often incorporate a time component.

While several initiatives seek to represent time using a Linked Data approach, *Periods*, *Organized* (*PeriodO*)[31], a gazetteer of chronological periods, appears to be the most advanced in terms of development, and appears most frequently in the literature. Rather than seek to provide an authoritative identifier for each chronological period, *PeriodO* provides a URI for each *assertion* of a chronological period.[32] Each assertion includes the period name, date range and the geographical area to which it applies, linked to a URI from a spatial gazetteer such as *Pleiades*.[33] In this way, *PeriodO* "attempt[s] to mirror scholarly practice" by representing the 'fuzziness' and disagreements in this area of scholarship.[34] *PeriodO* has therefore been developed with existing research processes and the research community in mind: acknowledging that much Ancient World data cannot be considered authoritative, and that representing it in a way that implies otherwise would deter potential users.[35] Assertions are linked based on relationships between them, e.g. whether a particular term provides a broader or narrower definition than another for a similar chronological period.[36] Users can therefore maintain consistency by linking to periods that fall within the same sequence, as well as comparing similar definitions. *Pleiades* incorporates *PeriodO* URIs to allow datasets to define their temporal scope.[37]

Humanities research in general, and Ancient World research in particular, often involves the study of objects, including texts, artworks, and material culture, with many occupying more than one of these categories. As an example, *Trismegistos*[38] provides access to metadata about documentary texts from the Ancient World, while linking externally to the texts themselves. These texts were predominantly written on papyri, but also include ostraca (pottery sherds), wooden tablets, and inscriptions; their information and contents were previously contained in other databases and print publications.[39] Its remit was originally restricted to Egypt between 800 BCE and 800 CE but has since expanded geographically and is beginning to expand chronologically.[40] In addition to descriptive metadata about the texts, entities mentioned within them, such as places, people and chronological periods have been extracted using a combination of named entity recognition (NER) and manual editing.[41] Although *Trismegistos* data is modelled using relational databases,[42] it provides URIs in the form of TM-numbers, which are used in Linked Data resources such as *EAGLE* and *Peripleo* (both discussed below).[43]

31 https://perio.do/ (last access 20.12.2023).

32 Buchanan et al. (2016), 3; Rabinowitz (2021).

33 Shaw et al. (2018).

34 Rabinowitz (2014).

35 Rabinowitz et al. (2016), 51.

36 Rabinowitz et al. (2018), 207.

37 Rabinowitz (2021).

38 https://www.trismegistos.org/ (last access 20.12.2023).

39 Depauw (2018), 193–195; Depauw / Gheldof (2014).

40 Broux (2017b).

41 Broux (2017a), 13; Broux / Depauw (2015); Depauw (2018), 196–197.

42 Trismegistos (n.d.); Verreth (2017), 202.

43 Depauw (2018), 199; Simon et al. (2016a).

Where appropriate vocabularies and authority files to represent contextual entities already exist, their URIs can be used by other initiatives, to ensure clarity and consistency of meaning and to connect their resources to machine-readable information from external sources. For example, rather than creating a new resource-specific list of places, place information in an Ancient World resource can be represented using the relevant *Pleiades* URI(s).[44] As well as providing efficiency by avoiding duplication of effort and potential redundancy,[45] linking to established URIs (where possible) enables integration. Integrating related resources in this way leads to richer descriptions of concepts and entities throughout interlinked datasets, enhancing discoverability. Vitale et al. describe how this can facilitate the interoperability of future developments, particularly if an agreed set of standards can be used by a research community.[46] Additionally, Prag and Chartrand attest that the approach of linking to established URIs can bring much-needed consistency to resources that rely heavily on project-specific implementations of standards that are more open to interpretation, such as EpiDoc[47].[48]

However, in many cases established URIs do not sufficiently represent what needs to be described. For example, when focusing on place, *GeoNames* is often inadequate for describing historical places, while locations in *Pleiades* do not have sufficient precision to allow detailed spatial analysis.[49] There is also limited capacity in existing vocabularies for representing nuanced levels of difference and similarity between entities, i.e. values that fall between 'same as' or 'different from'.[50] As a result, data producers are often left with the choice between linking to established URIs that may not accurately represent the entity or concept they wish to describe or creating new URIs that link to their more established (but not exactly equivalent) counterparts. The latter would usually be more desirable in a scholarly context; however, it requires considerably more time than using existing resources and is often not a realistic prospect for fixed-term projects, due to the implications for long-term sustainability.

Additionally, people differ in their interpretation of the meaning of terms used to represent entities and concepts, resulting in inconsistent application, particularly in cases where there is a distinction between a URI used to represent an entity itself and a URI used to represent information about that entity.[51] Even where clear definitions are available, the suitability of a particular term in a particular context remains subject to the implementer's interpretation of its meaning, which may differ from that of the user. This discrepancy could result in misunderstandings, although these can be mitigated with clear documentation to describe the producer's understanding of their terminology.

Similarly, there are often subtle disciplinary or theoretical differences between definitions of terms that might initially appear to be equivalent, which can cause difficulties in aligning ontologies, affecting subsequent information retrieval.[52] Such differences potentially lead to incompatibility with some researchers' views, or require them to adjust existing ways of working.[53] To compound this issue, Limp states that in the archaeology domain, individual ontology development is seen as an integral part of

44 Horne (2020a).

45 Buzi et al. (2018), 40.

46 Vitale et al. (2021), 9.

47 EpiDoc (https://sourceforge.net/p/epidoc/wiki/Home/ [last access 20.12.2023]) is an extension of standards produced by the Text Encoding Initiative (TEI), initially developed for encoding epigraphic documents, but which has since been applied to papyri and manuscripts.

48 Prag / Chartrand (2018), 248.

49 Horne (2020b), 218; Middle et al. (2022).

50 Brown (2022), 4.

51 Brown (2022), 6–7.

52 Gerth (2016), 31–32.

53 Geser (2016), 12, 73; Meroño-Peñuela et al. (2014), 20.

scholarship, even when suitable terms and vocabularies already exist.[54] Without some level of cooperation between data producers, this situation potentially results in project-specific data silos with models that are more difficult to align.[55] Projects such as *Federated Archaeological Information Management Systems* have sought to mitigate issues of conflicting terminology by mapping local terms for particular concepts and entities to a core ontology;[56] however, if there is not exact alignment between term definitions, mapping is either not possible or requires extensive documentation to advise users of potential inconsistencies.

3.2 Modelling Complexity

While any form of data modelling about real-life concepts and entities is an approximation, Linked Data can provide a more accurate means of representing these complexities, facilitating the process of finding commonalities and eliciting meaning. Modelling complexity is particularly important where the information about an entity comprises a collection of relationships that do not fall into a neat, hierarchical structure, as in the examples discussed below. In many cases, these entities can often be most accurately represented in relation to events and interactions, rather than a table of characteristics or an increasingly complicated set of relational databases.

Data about objects can be particularly complex: each object is a product of a series of interventions (usually by people) over time, often in multiple places; objects can be classified into different types, with increasing levels of granularity, and their contents are often rich with textual and/or visual information. Linked Data technologies are ideally poised to capture the layered and multi-faceted nature of such objects in a way that would be difficult (if not impossible) to achieve accurately using tabular formats. In turn, object-based data is particularly amenable to a Linked Data approach, largely because it is often already available as a structured format, in collections or archaeological databases.

Many LAWD tools and resources, particularly those in the cultural heritage domain, use the *CIDOC Conceptual Reference Model* (*CIDOC CRM*)[57], an ontology that semantically describes entities, concepts and relationships relating to objects, facilitating interoperability between collections. It is the only cultural heritage ontology to be recognised as an ISO standard[58],[59] indicating its maturity and positive reception. In addition to its extensive core ontology, members of its user community have developed extensions to describe domains such as archaeological excavations[60] and buildings[61], ancient texts[62], spacetime[63], and the provenance of digital objects[64]. Rather than placing the individual object at the centre of the data model, *CIDOC CRM* models a series of events in which it was involved (e.g., production, acquisition and transformation), thereby providing an informative way of connecting the object to other entities.[65] Its relevance is not limited to a cultural heritage context; both Grossner and

54 Limp (2011), 278.

55 Geser (2016), 17.

56 Ross et al. (2015), 126–127.

57 http://www.cidoc-crm.org/ (last access 20.12.2023).

58 https://www.iso.org/obp/ui/#iso:std:iso:21127:ed-2:v1:en (last access 20.12.2023).

59 Bruseker et al. (2017), 108.

60 CRMarchaeo: http://www.cidoc-crm.org/crmarchaeo/ (last access 20.12.2023).

61 CRMba: http://www.cidoc-crm.org/crmba (last access 20.12.2023).

62 CRMtex: http://www.cidoc-crm.org/crmtex (last access 20.12.2023).

63 CRMgeo: http://www.cidoc-crm.org/crmgeo/ (last access 20.12.2023).

64 CRMdig: http://www.cidoc-crm.org/crmdig (last access 20.12.2023).

65 Bruseker et al. (2017), 113.

Hill, and Meroño-Peñuela et al. advocate the use of event-based ontologies for representing historical information in an academic research context.[66]

Elsewhere, there are relatively mature Linked Data solutions for the representation and discovery of specific objects such as coins (*Nomisma*[67]), inscriptions (*EAGLE*[68]), and papyri (*Trismegistos* and *Papyri.info*[69]). These objects form particularly interesting use cases due to their combination of textual, visual, and material culture characteristics. A complex aspect of many papyri and some inscriptions is their fragmentary nature. Parts of the same text can appear on multiple physical objects, many are incomplete, and scholars must often use their best judgement regarding missing words. Integration of multiple collections is therefore essential for a more comprehensive understanding of their contents and meaning.[70] Contrastingly, coins are mass-produced objects with a specific set of attributes that applies to all instances, irrespective of chronological period or geographical area, which can often be classified using established typologies.[71] For example, every coin has obverse (heads) and reverse (tails) sides, and every coin was minted in a particular place. This predictable structure, incorporating discrete entities and concepts that can be identified with persistent URIs, as well as the numerous potential connections to external resources, makes Linked Data particularly appropriate for modelling numismatic data.[72]

Different approaches to modelling Linked Humanities Data therefore have varying levels of complexity, with accompanying advantages and disadvantages. Simple data models might provide efficiency and a low barrier to implementation, but often do not capture the degree of nuance required for research purposes.[73] Conversely, complex models like *CIDOC CRM* provide a rich level of detail; however, as the degree of nuance increases, so do the barriers to its effective implementation. For example, incorporating multiple terms with similar definitions can cause ambiguity and inconsistency in their application and complicate analysis of the resulting data.[74] Once such a complex ontology has been applied, it becomes more difficult to make inferences from the data (see below, section 3.4) because the number of conditions required for an inference to be confirmed increases with the number of terms used.[75] Additionally, several projects have chosen not to use *CIDOC CRM* (in particular) due to the considerable time and resources required to model data in this way.[76] The latter is a particular issue for academic projects, due to the short-term nature of research grants. Producing and consuming data involving complex ontologies like *CIDOC CRM* might therefore deter researchers from working with Linked Data at all in future.

A compromise could involve implementing a simpler approach at the outset, while ensuring there is scope for adding further complexity.[77] Such an approach might be achieved by encouraging producers

66 Grossner / Hill (2017), 9; Meroño-Peñuela et al. (2014), 13.

67 http://nomisma.org/ (last access 20.12.2023).

68 https://www.eagle-network.eu/ (last access 20.12.2023).

69 https://papyri.info/ (last access 20.12.2023).

70 Celano (2018), 139.

71 Heath (2018), 36.

72 Gruber / Meadows (2021).

73 Liu et al. (2017), 350.

74 Cayless (2019), 46; Gerth (2016), 21; Liu et al. (2017), 349.

75 Isaksen (2011), 155.

76 Bodard et al. (2017), 35; Kansa et al. (2018), 501.

77 Vitale et al. (2021), 9–10.

to map some or all of their data to a set template,[78] or by applying a simplified version of the ontology. An example of the latter approach can be found in *Linked Art*[79], a community-led data model for describing art objects. Acknowledging its complexity issues, *Linked Art* uses a simplified version of *CIDOC CRM*, to minimise confusion and ambiguity.[80] Such an approach might therefore be a potential solution to *CIDOC CRM* implementation in a Humanities research context.

3.3 Communicating Uncertainty

An additional dimension to consider when modelling in LAWD is that of uncertainty. Much of the data that supports Ancient World research is based on interpretation rather than fact. Some data will never reach the point where it can be considered 'factual'; for example, representations of mythological entities about whom there is conflicting information from different sources.[81] To take advantage of the benefits Linked Data provides, producers must consider how uncertainty can best be modelled within the data structure. For cases where several alternative values are possible, Thaller advocates a mathematical approach, where a probability is assigned to each value;[82] Niccolucci and Hermon demonstrate how a similar approach might be modelled using *CIDOC CRM*.[83] However, in many cases it is not feasible to perform a reliable calculation of probability (e.g., if the relative frequencies of different possible values cannot readily be estimated), and assigning probability based on the data producer's level of certainty would be extremely subjective.

Another example is that of the Linked Places format[84], where period definitions can start and end with a date range, rather than a single date, and geographic locations can be assigned a value of "'certain', 'less certain' or 'uncertain'",[85] although again these values rely heavily on the judgement of the person producing the data. As an alternative, *PeriodO* and *Pleiades* mirror more traditional scholarly practice by modelling information as cited assertions that link to the original source, which Golden and Shaw refer to as "nanopublications".[86] Despite these efforts, there is still no universally agreed method for expressing and communicating uncertainty in Linked Humanities Data and this is likely to remain the case. Humanities researchers should therefore always treat both digital and physical sources with an element of caution; their critical evaluation can be assisted by appropriate information about these sources' limitations.

3.4 Reasoning

Once a dataset is modelled using established ontologies and linked to appropriate identifiers, a reasoning (or inferencing) system can be used to automatically produce new knowledge, inferred from the relationships between entities, which would be difficult or impossible for a single user to discover manually. For example, in the *Sharing Ancient Wisdoms* project, automated reasoning was used to detect texts that were linked to the same translation but were not linked to each other.[87] Once the reasoner has elicited this information, it is the researcher's task to verify its likelihood based on any known gaps or

78 Binding et al. (2019), 371.

79 https://linked.art/ (last access 20.12.2023).

80 Linked Art Contributors (n.d.-a).

81 Bodard et al. (2017), 36.

82 Thaller (2020).

83 Niccolucci / Hermon (2017).

84 https://github.com/LinkedPasts/linked-places-format (last access 20.12.2023).

85 Grossner / Mostern (2021), 43.

86 Golden / Shaw (2016).

87 Tupman / Jordanous (2014).

issues inherent in the data.[88] Datasets aligned using Linked Data technologies therefore have the potential to become more than the sum of their parts. However, few publications relating to LAWD refer to the application of these capabilities, perhaps because thus far, emphasis has been placed on building tools, resources and infrastructures; substantial work on inferencing might then form the next stage of the process.

3.5 Making Connections

Humanities research topics often involve the analysis and interpretation of evidence from multiple sources of different types; using a Linked Data approach brings these research objects together, allowing them to be explored in the same virtual space. Once datasets are connected using Linked Data techniques, multiple collections and repositories that previously existed in separate silos can be searched via a single federated query, with the potential to visualise their combined data.[89] The ability to conduct one search instead of many reduces the time required for the discovery/data collection phase of a project, allowing more time to be spent on analysis and interpretation.[90] Linked Data tools and resources therefore have the potential to facilitate investigation into new research questions that would either be difficult or impossible to address using other technologies.

Pelagios[91] is one of the best known LAWD initiatives, although (like *Pleiades*, 3.1, above), it has subsequently increased its temporal and geographical scope.[92] In *Pelagios*, place names in a digitised text or image are identified and annotated with relevant gazetteer URIs, using the World Wide Web Consortium (W3C) Web Annotation Data Model[93]. Rather than expecting contributing organisations to adjust their data structures, *Pelagios* provides interconnectivity by hosting only these annotations as "stand-off markup", linking to the record for each object in its original dataset. In this way, the gazetteers provide a "central backbone" to connect multiple datasets, based on their relationships to place.[94] To ensure interoperability between different gazetteers, *Pelagios* developed a Gazetteer Interconnection Format[95], which has since been superseded by the Linked Places format.[96] Its implementation enables gazetteers with very different approaches to defining and representing the concept of 'place' to be reconciled and used alongside each other,[97] and it is now considered a "de facto standard" for "exchange of historical place name data".[98] Overall, *Pelagios* provides an example of how Linked Data from gazetteers can be brought together into an infrastructure and made available to the research community via user-friendly tools and resources. Its relative success in this regard could be due to several factors, such as openness, low barrier to entry, and an active user community.

Elsewhere, the *Virtual International Authority File* (*VIAF*)[99] aggregates information about people from authority files produced by libraries and cultural heritage institutions throughout the world; some or-

88 Meroño-Peñuela et al. (2014), 20–21.

89 Bagnall / Heath (2018), 184; Geser (2016), 26; Meroño-Peñuela et al. (2014), 19.

90 Gruber (2018), 59.

91 https://pelagios.org/ (last access 20.12.2023).

92 Simon et al. (2014), 105.

93 https://www.w3.org/TR/annotation-model/ (last access 20.12.2023).

94 Simon et al. (2017), 114.

95 https://github.com/pelagios/pelagios-cookbook/wiki/Pelagios-Gazetteer-Interconnection-Format (last access 20.12.2023).

96 Grossner / Mostern (2021), 41.

97 Simon et al. (2016b), 106.

98 Berman et al. (2016), 123–124.

99 http://viaf.org/ (last access 20.12.2023).

ganisations, events and places are also included, albeit to a lesser extent.[100] In recent years there has been increasing engagement with the research community, with a view to enriching the information available. In particular, an initiative called *Scholars' Contributions to VIAF* resulted in the incorporation of Ancient Greek data from the *Perseus Digital Library*[101] and Syriac data from *Syriaca.org: the Syriac Reference Portal*[102].[103] As well as broadening *VIAF*'s linguistic diversity, these additions have increased its relevance to Ancient World researchers. However, like *GeoNames*, *VIAF* does not provide the degree of nuance required for many Ancient World projects; for example, it contains relatively little information about mythological entities.[104] Again, more domain-specific resources are needed, ideally linked to *VIAF* to ensure connection to the wider web of data.

Perhaps the most extensive and well-known such resource for ancient people is *Standards for Networking Ancient Prosopographies: Data and Relations in Greco-Roman Names* (*SNAP:DRGN*)[105]. Like *VIAF*, *SNAP:DRGN* provides authority files for people, but with a focus on the Ancient World. Existing catalogues of personal names (such as that provided by *Trismegistos*, section 3.1, above) are aligned and combined into one searchable dataset, and connected based on their relationships with each other, using the SNAP data model.[106] Similarly to *Pelagios*, *SNAP:DRGN* has opted for relatively lightweight methods for data linking, to best reflect the ambiguities and incompleteness inherent in Ancient World person data, particularly where mythological entities are concerned.[107]

There are many cases where data connection has been achieved through the implementation of *CIDOC CRM*, with notable examples in the area of archaeological infrastructure. In *Arachne*[108], the object database of the German Archaeological Institute, *CIDOC CRM* has greatly facilitated its interoperability and potential for information exchange.[109] Indeed, the *CIDOC CRM* framework was used to link *Arachne* data with that of *Perseus*' art and archaeology collection as part of the *Hellespont Project*.[110] Additionally, the *Advanced Research Infrastructure for Archaeological Dataset Networking in Europe* (*ARIADNE*)[111] project integrated datasets from European national repositories by mapping the entities within them to *CIDOC CRM*. Users can access this data via the *ARIADNE Portal*, which includes visualisation and query functionality,[112] demonstrating the potential for ontologies such as *CIDOC CRM* to form the basis of LAWD access via a usable interface.

Other examples of initiatives that connect object data include those focused on specific types such as inscriptions, coins and papyri. The *Europeana Network for Ancient Greek and Latin Epigraphy* (*EAGLE*)[113] is an initiative by *Europeana*[114], a digital library based on a Linked Data model that ag-

100 Angjeli et al. (2014), 2–3; Hickey / Toves (2014).

101 http://www.perseus.tufts.edu/hopper/ (last access 20.12.2023).

102 http://syriaca.org/ (last access 20.12.2023).

103 OCLC (2019); Smith-Yoshimura (2013); Smith-Yoshimura (2014).

104 Gerth et al. (2016), 15.

105 http://snapdrgn.net (last access 20.12.2023).

106 Bodard et al. (2017), 31; Lawrence / Bodard (2015).

107 Bodard et al. (2017), 31–36.

108 https://arachne.dainst.org/ (last access 20.12.2023).

109 Scriba / Stockinger (2016).

110 Crane (2014).

111 https://ariadne-infrastructure.eu/ (last access 20.12.2023).

112 Aloia et al. (2017); Meghini et al. (2017).

113 https://www.eagle-network.eu/ (last access 20.12.2023).

114 http://www.europeana.eu/portal/en (last access 20.12.2023).

gregates multiple European digital collections. *EAGLE* provides access to images and metadata for ancient inscriptions by bringing together existing databases. Inscription texts are structured using EpiDoc, with artefacts modelled using the *EAGLE Common Metadata Model and Schema*, predominantly comprising selected terms from *CIDOC CRM*. These approaches are united using *EAGLE*'s *Epigraphy Aggregation Conceptual Model* (*EACM*),[115] with duplicate entries disambiguated using *Trismegistos* identifiers. *Nomisma*[116] is a Linked Data ontology and resource provided by the American Numismatic Society (ANS)[117] to facilitate integration and discovery of data about coins, predominantly from the Ancient World. *Nomisma* links coin hoards, types and specific instances,[118] using terms from established ontologies and authority files, such as *CIDOC CRM* and *Pleiades*, to facilitate integration with external datasets.[119] *Papyri.info* is an aggregated collection of digitised documentary papyri that integrates several online resources[120], with disambiguation provided by *Trismegistos* URIs.[121]

Caution must be taken, however, in aligning datasets modelled using different approaches. For example, papyri often exist in the form of multiple fragments that originally belonged to the same text; simultaneously, the same writing surface may have been reused for two or more texts. This situation has resulted in different approaches to identifying and numbering texts, fragments and writing surfaces.[122] In the case of *Papyri.info*, some source collections choose to assign each text by a different scribe to a different URI (even if they appear on the same writing surface), while others choose to identify all texts written on the same surface with the same URI.[123] Therefore, using Linked Data technologies to aggregate multiple collections does not always provide a straightforward means of aligning data models, potentially resulting in inconsistency and confusion for the user.

3.6 Discoverability and Usability

Linked Data tools and resources that bring together data from multiple sources have great potential for serendipity, exposing collections and datasets previously unknown to the researcher,[124] such as those from relevant, but distantly related, subject domains,[125] or those in other languages.[126] As these datasets are often produced by a variety of organisations, for a wide range of purposes, they complement each other's information and provide the user with broader context for their object of study, potentially resulting in more accurate interpretations.[127] Searching multiple datasets simultaneously can additionally reveal relationships between (for example) objects held in separate collections,[128] as well as facilitating virtual reunification of objects from the same place.

115 Mannocci et al. (2014), 291.

116 http://nomisma.org/ (last access 20.12.2023).

117 http://numismatics.org/ (last access 20.12.2023).

118 Heath (2018), 41.

119 Gruber (2016), 100; Gruber / Meadows (2021).

120 The *Duke Databank of Documentary Papyri*, the *Advanced Papyrological Information System*, and the *Heidelberger Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens.*

121 Reggiani (2017), 227.

122 Ast / Essler (2018), 69; Polis / Razanajao (2016), 25.

123 Baumann (2013), 97; Cayless (2011), 32; Reggiani (2017), 74–75.

124 Gruber (2018), 62.

125 Geser (2016), 50; Pagé-Perron (2017), 11.

126 Koch / Koch (2017), 171.

127 Kansa et al. (2018), 503.

128 Geser (2016), 12; Rabinowitz et al. (2018), 212–213.

Use of Linked Data resources for discovery and analysis need not be restricted to researchers with an understanding of SPARQL. Queries can be facilitated via detailed filters and faceting options,[129] with data made meaningful to a variety of audiences by building narratives around curated subsets of digital objects.[130] Connections can be visualised as maps or networks and incorporated into discovery tools, to enable powerful cross-collection searching and identifying key entities and relationships.[131] New tools are being developed to facilitate the process of building user interfaces to explore and consume Linked Data.[132] These include the JavaScript-based *Sampo-UI*[133] framework, which provides a customisable filtering and faceting interface to query SPARQL endpoints,[134] as well as the collaborative knowledge base software *Wikibase*[135], which Koho et al. describe as "an easy, high-return, out-of-the-box solution".[136] Such tools and resources have the potential to break down the barriers between datasets, improve the efficiency of the research process, and provide a more holistic view of a subject domain, while reducing the need for specialist training to take advantage of the benefits that LAWD can provide.

Pelagios, *EAGLE* and *Nomisma* provide some examples of how LAWD resources can be usable even for those with limited technical expertise or subject knowledge. For example, *Peripleo*[137] is a visualisation and discovery tool, developed as part of *Pelagios*, which displays data via a map interface, with searching, filtering and faceting functionality.[138] As such, it provides an intuitive, visual means of exploring and querying multiple datasets.[139] While *Peripleo* was initially restricted to data associated with *Pelagios*, the new *Peripleo Lite*[140] tool allows the user to select the datasets they require for geographical exploration and visualisation.[141]

Similarly, *EAGLE* facilitates discovery across its component databases via a single search interface that incorporates filtering, faceting, and image recognition.[142] Alongside this functionality, *EAGLE* also incorporates mobile and storytelling applications, to fulfil their aim of broadening access to inscriptions to wider audiences beyond academia. The former provides visual search capabilities to enable in situ identification of inscriptions, while the latter enables users to view and create narratives using *EAGLE* resources, with the option to incorporate maps from *Pelagios*.[143] In a review, Hedrick praised the storytelling app for its "elegant way" of integrating *EAGLE* materials, but found that *EAGLE*'s main search interface can be difficult to use effectively for users unfamiliar with its constituent data-

129 Simou et al. (2017), 221.

130 Liuzzo et al. (2017), 521.

131 Isaksen et al. (2018).

132 For a recent survey, see Bernasconi et al. (2023).

133 https://seco.cs.aalto.fi/tools/sampo-ui/ (last access 20.12.2023).

134 Rantala et al. (2023).

135 https://wikiba.se/ (last access 20.12.2023).

136 Koho et al. (2023), 56:5.

137 http://peripleo.pelagios.org (last access 20.12.2023).

138 Simon et al. (2016a).

139 Vitale et al. (2021).

140 http://pelagios.org/peripleo-lite/ (last access 20.12.2023).

141 Barker (2021).

142 Prandoni et al. (2017).

143 Liuzzo et al. (2017); Orlandi et al. (2014).

bases.[144] These comments suggest that *EAGLE*'s researcher-focused search interface might benefit from some of the usability measures incorporated into its more creative, public-facing resources.

Nomisma also benefits from other discovery mechanisms, rather than relying solely on a SPARQL endpoint. The ANS also provides searchable resources such as *Online Coins of the Roman Empire* (*OCRE*)[145], *Coin Hoards of the Roman Republic* (*CHRR*)[146], *Coinage of the Roman Republic Online* (*CRRO*)[147] and *PELLA: Coinage of the Kings of Macedonia*[148], as well as its overarching catalogue, *MANTIS*[149]. Providing multiple methods for querying *Nomisma* data ensures its consumption is neither restricted to more technically advanced users, nor limited by the confines of a visual interface.

A more recent initiative, *Modelling Ancient Narratives, Territories, Objects* (*MANTO*)[150], took a different approach. Their dataset of Greek mythological people (and person-like entities) is managed using the *Nodegoat*[151] platform, which provides flexible data modelling and visualisation options for Humanities projects. As such, existing data in *MANTO* can already be displayed as maps and networks, without requiring specific technical skills. The project team additionally consulted their potential user community during development, by asking them to test a prototype version of the tool and complete an online questionnaire.[152]

However, due to its relatively complex structure (and often, producers' lack of familiarity thereof), building a Linked Data driven resource and ensuring its usability by non-technical researchers can be more difficult and time-consuming than a similar resource based on tabular data. If insufficient development time is available within the funded period of a project to produce a usable interface, its absence significantly limits the number of people who could potentially engage with and benefit from the dataset. As a result, some resources, such as *SNAP:DRGN,* exist in the form of rich but largely inaccessible datasets that must be queried via a SPARQL endpoint. While the data has the potential to be extremely useful, the proportion of interested parties who might be able to access it is relatively small, limiting the scope for its consumption. This situation additionally demonstrates the importance of funding projects to enhance and consume existing Linked Data, rather than prioritising the production of new datasets.

Even users who are familiar with SPARQL might not always be able to make effective use of a SPARQL endpoint, however, as they first need to be familiar with the way in which new and existing ontologies and data models have been implemented in this particular context.[153] The W3C recommends that providers describe linked datasets using the Vocabulary of Interlinked Datasets (VoID), which includes specifying ontologies used.[154] However, such a description is not always included; out of 427 SPARQL endpoints, Buil-Aranda et al. found that approximately two thirds were not accompanied by a VoID description.[155] Additionally, a single query has the potential to generate huge

144 Hedrick (2018).

145 http://numismatics.org/ocre/ (last access 20.12.2023).

146 http://numismatics.org/chrr/ (last access 20.12.2023).

147 http://numismatics.org/crro/ (last access 20.12.2023).

148 http://numismatics.org/pella/ (last access 20.12.2023).

149 http://numismatics.org/search/ (last access 20.12.2023).

150 https://www.manto-myth.org/manto (last access 20.12.2023).

151 https://nodegoat.net/ (last access 20.12.2023).

152 Hawes (2020).

153 Calvanese et al. (2016), 214.

154 Alexander et al. (2011).

155 Buil-Aranda et al. (2013), 280.

amounts of data, which is both unhelpful to the consumer and computationally intensive for the host. Therefore, even for technically experienced users, direct access to a SPARQL endpoint is not always advisable; providing access via an interface or API is usually preferable.[156]

Another barrier to querying data from multiple external sources via a SPARQL endpoint is the lack of control over its availability.[157] For example, Gerth et al. had originally intended to incorporate data from the British Museum in their experiments with integrating sculpture datasets, but were prevented from doing so due to unavailability of the SPARQL endpoint.[158] Additionally, in their 27-month study of 427 SPARQL endpoints encompassing different subject areas, Buil-Aranda et al. found that only 32.2% were available 99–100% of the time, while 29.3% were available less than 5% of the time. They note that many endpoints in the latter category were produced through experimentation with the technology, rather than long-term resource provision, and are now permanently unavailable.[159] As well as causing functional issues, resource unavailability can affect citations, a particular concern when referring to online sources in scholarly publications,[160] and may perpetuate the idea that online material is not stable enough to be cited. There are, therefore, considerable sustainability issues with Linked Data tools and resources; while the same could be said of any Digital Humanities project, the impact of Linked Data unavailability has the potential for more far-reaching repercussions, as it directly affects the availability of the resources linked to it. However, Schmidt et al. suggest that hosting this data via an established, openly-available data hub such as *Wikidata*, could mitigate these issues.[161]

Similarly, it is often unclear whether data held in external resources is accurate and up to date;[162] therefore, any inaccuracies could potentially be reproduced across multiple resources.[163] Such reliability issues might explain why many 'linked datasets' contain only internal links, rather than connecting to external resources;[164] however, this phenomenon may also be due to an actual or perceived lack of relevant external datasets and ontologies, or not being aware of the benefits of linking to more general resources. Quality issues can be mitigated with appropriate documentation, including provenance information, to assure potential users that the data is trustworthy.[165]

As with any type of technology or data structure, users with the requisite skills and experience are likely to derive the most benefit from Linked Data technologies; however, as the above examples have demonstrated, if resources are carefully designed with potential audiences in mind, all interested researchers should be able to engage with them.

3.7 Awareness and Training

Many Humanities researchers are not aware of Linked Data technologies, the potential benefits of this approach in the context of their research topics, or how it might be implemented. This situation is likely due to Linked Data rarely being covered in standard institutional training offered to staff and students, which usually focuses on spreadsheets and relational databases. As a result, researchers often tend to think in a tabular format, without considering that their data might be better represented by a

156 Gunter et al. (2019), 4–5; Schweizer / Geer (2019), 1.

157 Geser (2016), 15.

158 Gerth et al. (2016), 17.

159 Buil-Aranda et al. (2013), 289–290.

160 Hannemann / Kett (2010), 2.

161 Schmidt et al. (2022), 341.

162 Calvanese et al. (2016), 213.

163 Angjeli et al. (2014), 2.

164 Isaksen (2011), 64.

165 Geser (2016), 69–71.

networked, graph structure.[166] Even for those researchers who are aware of the potential benefits of Linked Data technologies, substantial training is often required for their effective implementation, which may not seem a realistic approach to take within the tight time constraints of funded projects.[167] For example, Granados-García found that using a Linked Open Data approach required far more time and training than initially anticipated and recommended that future data producers ensure they have sufficient technical knowledge and support (or the time to acquire them) before embarking on similar projects.[168] Similarly, Polczynski and Polczynski caution future gazetteer builders against a Linked Data approach, as their experience indicated that any benefits were outweighed by the amount of time and training required for its effective implementation.[169] In fact, Smith-Yoshimura found when surveying Linked Data producers from the cultural heritage domain that the main barrier to producing Linked Data was a "steep learning curve for staff", indicating the importance of addressing this fundamental obstacle.[170]

While I have noted implications for project timescales throughout the above discussion, time constraints are a particular concern for researchers who are unfamiliar with Linked Data and its production. This situation is compounded by difficulty in accessing, or unawareness of, adequate training, and often results in the prioritisation of immediate research objectives rather than longer-term usage of the dataset, tool, or resource beyond the lifetime of the project. To avoid these issues, and ensure that high quality, reusable data is produced, Gerth recommends that Linked Data should be produced via a collaborative process involving both technology experts and domain specialists.[171] However, Geser acknowledges that, while ideal, expert support is not always available, and that the development of usable tools to support Linked Data production could be a more effective solution to minimise the need for training.[172]

Such user-friendly tools that allow non-technical researchers to produce or enhance Linked Humanities Data are few in number.[173] Additionally, as these tools are often produced with the goal of academic experimentation rather than long-term usability, they have rarely been tested sufficiently to ensure consistency and reliability, and are often not adequately maintained.[174] Until the learning curve for developing Linked Data resources is reduced, the temptation will likely be to work with familiar data structures; a potentially 'safer' option that ensures development of a usable resource within the funded period, while maximising the amount of time available for research.

Some notable exceptions are *Pelagios*' annotation platform *Recogito*[175], and the Semantic Technologies Enhancing Links and Linked data for Archaeological Resources (STELLAR) applications[176].[177] *Recogito* provides a visual interface that allows users to annotate places, people and events in texts or images, either with tags or links to gazetteers. NER functionality is incorporated, while also allowing

166 Barbera (2013), 96; Ross et al. (2015), 118.

167 Isaksen (2011), 153–154; van Hooland / Verborgh (2014), 51.

168 Granados-García (2020), 261–264.

169 Polczynski / Polczynski (2022), 330.

170 Smith-Yoshimura (2018).

171 Gerth (2016), 14.

172 Geser (2016), 12, 56.

173 Barbera (2013), 98; Thiery / Engel (2016), 259.

174 Geser (2016), 55–56.

175 http://recogito.pelagios.org/ (last access 20.12.2023).

176 https://hypermedia.research.southwales.ac.uk/kos/stellar/stellar-applications/ (last access 20.12.2023).

177 Geser (2016), 12, 58.

manual annotation and confirmation.[178] *Recogito*'s development was substantially informed by user consultation,[179] resulting in an intuitive and widely-adopted research tool that renders Linked Data production (at least on a small scale) possible for researchers without specific training.[180] The STELLAR applications consist of tools and templates that allow conversion of archaeological data in CSV or SQL format to RDF, using the CRM-EH extension of CIDOC CRM, developed by English Heritage[181].[182] Templates were developed following a series of workshops with archaeologists to establish the most helpful and broadly applicable use cases and workflows within which they could be applied. Usability by data producers was therefore at the heart of both initiatives, although only *Recogito* remains in a current and readily available state.[183]

In cases where a usable interface can be produced in the time available, producers should acknowledge that methods for presenting and visualising Linked Data may be unfamiliar to Humanities researchers. For example, many potential users may not have prior experience or training in network analysis, which could result in an inability to use the resource to its full potential, or to inaccurate interpretations of their query results.[184] Accurate interpretation by the user is particularly important for resources with inferencing capabilities; as Hickey and Toves warn, results will reflect any ambiguities or inconsistencies in the original data, and as such should always be checked by domain experts.[185]

3.8 Openness and Collaboration

Even if Linked Data producers are not able to build a resource themselves, the fact that their data uses open standards and vocabularies means it can more easily be accessed and used by others. Similarly, providing an open dataset using Linked Data formats can encourage new collaboration opportunities that might result in the funding and expertise required to build a user-friendly interface.[186] Therefore, as well as connecting datasets, Linked Data has the potential to bring researchers together and facilitate the process of building on previous work in overlapping subject areas. As such, the majority of tools, resources and initiatives discussed in this article are a result of collaboration, usually involving some combination of Ancient World researchers, computer scientists, and information specialists. For example, *Nomisma* was developed as a result of international collaboration between experts in the field, ensuring accurate representation of information and acceptance by the research community.[187] *EAGLE*'s development took a similarly collaborative approach, with feedback on models and processes actively sought from contributors to optimise the end result.[188] On a related note, the previous section included examples where collaboration with potential users was integral to the development process.

178 Simon et al. (2017); (2019).

179 Simon et al. (2015).

180 Vitale et al. (2021), 19.

181 https://hypermedia.research.southwales.ac.uk/kos/star/cidoc-crm-implementation-and-crm-eh/ (last access 20.12.2023).

182 May et al. (2012).

183 The source code for the STELLAR applications continues to be available via GitHub at https://github.com/cbinding/stellar (last access 20.12.2023) but, at the time of writing, has not been maintained for over four years.

184 Barker (2020).

185 Hickey / Toves (2014).

186 Buzi et al. (2018), 50.

187 Gruber (2018), 55.

188 Rocco (2017), 128; Santucci et al. (2016), 53.

This emphasis on collaboration within the Linked Ancient World Data community means that users are often active contributors to these resources, rather than being solely passive consumers of their data. Several initiatives have developed methods to encourage and facilitate these contributions, while maintaining scholarly quality and credibility. For example, to include their data in *Nomisma*, contributors need not structure it as Linked Data themselves. Instead, they provide a Google spreadsheet, with headings that can be directly mapped to *Nomisma* terms, which is then validated and converted to RDF.[189] Similarly, as well as ensuring discoverability through the *Papyrological Navigator*, *Papyri.info* additionally provides the *Papyrological Editor*, a collaborative editing tool. Texts are encoded using EpiDoc, and connected using RDF.[190] The resulting resource not only comprises a searchable collection of texts, it also provides open, dynamic and collaborative digital critical editions, to which any user can contribute.[191] Bambaci et al. have highlighted the relative ease with which users can encode text in the *Papyrological Editor*, using familiar papyrological conventions without compromising data quality and richness.[192]

Creating an open, usable resource for exploring Linked Data can, however, have some negative implications; making this data more discoverable can amplify information now known to be incorrect, terminology now considered to be offensive,[193] or potentially distressing images.[194] This issue is common to all data types, but its effects can be more pronounced where discoverability is aided via semantic enrichment. Avoiding or mitigating this issue often requires extensive work to update the dataset before linking can be considered, which poses a particular barrier for smaller institutions with limited funding. Otherwise, publishing the original dataset as Linked Data could have harmful consequences for users, as well as potentially perpetuating outdated language and interpretations in subsequent academic research.

Similarly, the exact locations of heritage sites are often considered sensitive information due to the potential for looting or vandalism.[195] To mitigate this, while allowing some discoverability, Kansa et al. reduced the level of precision to which locations are described in the *Digital Index of North American Archaeology*.[196] As well as altering newly published datasets in this way, similar actions may need to be taken for data that already exists online, but where integration with other datasets or richer description may significantly increase its visibility.

3.9 Sustainability

In some regards, the openness of Linked Data can have positive implications for sustainability, both in terms of immediate reuse potential and the relative ease of preserving data that uses open standards. However, the sustainability of any tool or resource through which that data can be accessed without requiring specific technical skills or knowledge is often a concern. The initiatives discussed above have taken differing approaches to ensuring that such tools and resources can continue to be usable over time.

189 Gruber (2016), 104–105.

190 Papyri.info (n.d.).

191 Reggiani (2017), 241, 269.

192 Bambaci et al. (2019), 28.

193 Gibson / Kahn (2016); Modest / Lelijveld (2018).

194 Holterhoff (2017).

195 Tolle / Wigg-Wolf (2016), 277.

196 Kansa et al. (2018), 494.

The *Trismegistos* team has used their data to perform statistical analyses, uncovering trends over time,[197] and to produce visualisations, particularly in the area of social network analysis.[198] However, while the data itself remains openly available, since January 2020 this additional functionality, as well as performing more complex queries via the user interface, requires users (or their institutions) to pay a subscription fee.[199] Although this move was deemed necessary to ensure long-term sustainability of the resource, it is unfortunate that features with the potential to open up *Trismegistos* to a wider audience are now only available to a restricted subset of users.

It is clear from a recent fundraising campaign to ensure *Papyri.info*'s sustainability, led by the Association Internationale de Papyrologues and the American Society of Papyrologists, that *Papyri.info* now occupies a position of fundamental importance to the wider papyrological community, rather than appealing only to those who consider themselves digital humanists.[200] These scholarly organisations have recognised that, despite considerable contributions by volunteers, a dedicated role is required for effectively managing the resource, which should not be funded by introducing subscription fees, ensuring that *Papyri.info* will remain openly available.

After initially being supported through a series of grant-funded projects, *Pelagios* now operates as a community-driven network of contributors, who drive improvements to its tools and share best practice.[201] Individual membership of the *Pelagios* community and joining the network as a partner project are both free of charge, with the expectation that partners will actively contribute to community activities. Hosting and maintenance for *Pelagios* tools and resources continue to be provided through institutional support, which is mitigated by the fact that all source code is made openly available, and that *Pelagios* does not host any of the data that can be accessed or annotated through their tools.[202]

A concern throughout this article has been the impact of short-term funding models on sustainability and usability. In the longer term, as no specific software is required to engage with Linked Data formats, the data itself should continue to be (re)usable for a variety of purposes, even if the interface for its consumption reaches the end of its life.[203] Therefore, although issues remain with the sustainability of tools and interfaces for consuming Linked Data, choosing a Linked Data format is a positive step towards ensuring longevity and reuse of the underlying data.

4. Conclusions

In this article, I have explored the current situation with regard to LAWD, including examples of relevant tools, resources and data structures; the advantages of this approach, and potential barriers to its implementation. Example initiatives focused on representation and discovery of different entity types, such as place, time, people, and objects, and applied different data modelling approaches with varying levels of complexity and granularity. Exploring these examples demonstrated the potential for using Linked Data as the basis for creative and engaging outputs for a wide variety of users, while remaining academically rigorous. Most initiatives took a collaborative approach to development, with the aim of ensuring usability and usefulness of the resulting tools and resources. Although they primarily focus on separate entity types, their producers have recognised the impact of these entities on each other; as

197 Depauw / Stolk (2015).

198 Broux (2017b); (2017c); Depauw (2018), 197–198.

199 Trismegistos (2019).

200 Torallas Tovar / Schubert (2019).

201 Barker et al. (2021), 2.

202 Kahn et al. (2021).

203 Biston-Moulin / Thiers (2018), 160; Jordanous et al. (2012).

such, there are many connections between them. Some resources are broad and multidisciplinary, and therefore ideal for making connections to the wider Linked Data ecosystem, with domain-specific initiatives often required to provide more nuanced representations.

Advantages that Linked Data can bring to Humanities research, as demonstrated through these examples, include consistent and unambiguous identification, alignment of different data models and vocabularies, the production of new knowledge via reasoning, and the effective modelling of complex concepts and entities. Once the data has been produced, there are lightweight mechanisms to connect multiple datasets, on which usable resources can be built. Incorporation of open standards facilitates collaboration and ensures sustainability of the data (if not the tools or interfaces). All the above advantages result in Linked Data facilitating research processes such as discovery, annotation, and visualisation.

However, there has been less uptake of LAWD than one might expect after exploring these advantages. This discrepancy is due to challenges and obstacles to its implementation, which often have implications for its subsequent use. Many such barriers are technical in nature and predominantly apply to tool and resource producers who are already familiar with Linked Data. These include the accurate representation of concepts and entities, the difficulty in aligning data from sources structured in different ways, the disadvantages of selecting either a simple or complex data model, the particular importance in Humanities research of effectively communicating uncertainty, and the reliability issues inherent in full or partial dependence on external sources. Even researchers with significant experience in Linked Data production will need to assess the potential impacts of these barriers on project timescales before opting to implement this approach.

Before these researchers even consider the technical obstacles to Linked Data implementation, however, there are wider issues that must be addressed. The first concerns a general lack of awareness among Humanities researchers of the benefits of Linked Data, with little training available to rectify this, and few use cases demonstrating its effective implementation. The second is the relatively small number of usable tools or resources for producing or consuming Linked Data, potentially due to their being produced for experimentation with the technology, rather than long-term availability. The third concerns instances where enhanced data discoverability is undesirable, e.g., for collections containing sensitive material. These higher-level barriers must be addressed before potential producers can even consider the above technological challenges.

Although the challenges and obstacles to production and consumption of Linked Data are substantial, they are not insurmountable. In the above discussion, I have largely focused on barriers to Linked Data production; however, the way in which each of these barriers is addressed (or not) has significant implications for the usability and uptake of the resulting tools and resources. While expert consultation is usually an ideal solution, many barriers can be mitigated with access to information, training, and the production of more usable tools and resources. In many cases, it can be preferable to start with a simpler, incremental, approach to linking data, with scope to increase the complexity at a later stage. This approach additionally allows more time to develop a usable interface; however, even where such interfaces have been developed and are well-used, it is extremely difficult to guarantee their long-term sustainability while simultaneously ensuring that they remain free to access. Time and money, therefore, continue to be the ultimate obstacles to Linked Ancient World Data implementation.

List of Abbreviations

ANS	American Numismatic Society
API	Application Programming Interface
ARIADNE	Advanced Research Infrastructure for Archaeological Dataset Networking in Europe
CHRR	Coin Hoards of the Roman Republic
CIDOC CRM	CIDOC (International Council of Museums' International Committee for Documentation) Conceptual Reference Model
CRRO	Coinage of the Roman Republic Online
CSV	Comma Separated Values
EAGLE	Europeana Network for Ancient Greek and Latin Epigraphy
EACM	Epigraphy Aggregation Conceptual Model
FAIR	Findable, Accessible, Interoperable, Reusable
HTTP	Hypertext Transfer Protocol
ISO	International Organization for Standardization
KOS	Knowledge Organisation System
LAWD	Linked Ancient World Data
LOD	Linked Open Data
LOUD	Linked Open Usable Data
MANTIS	MANTIS: A Numismatic Technologies Integration Service
MANTO	Modelling Ancient Narratives, Territories, Objects
NER	Named Entity Recognition
OCRE	Online Coins of the Roman Empire
PeriodO	Periods, Organized
RDF	Resource Description Framework
SNAP:DRGN	Standards for Networking Ancient Prosopographies: Data and Relations in Greco-Roman Names
SPARQL	SPARQL Protocol and RDF Query Language
SQL	Structured Query Language
STELLAR	Semantic Technologies Enhancing Links and Linked data for Archaeological Resources
UI	User Interface
URI	Uniform Resource Identifier
VIAF	Virtual International Authority File
VoID	Vocabulary of Interlinked Datasets
W3C	World Wide Web Consortium

References

Alexander et al. (2011): K. Alexander / R. Cyganiak / M. Hausenblas / J. Zhao, Describing Linked Datasets with the VoID Vocabulary, W3C (2011), https://www.w3.org/TR/void/ (last access 19.12.2023).

Aloia et al. (2017): N. Aloia / C. Binding / S. Cuy / M. Doerr / B. Fanini / A. Felicetti / J. Fihn / D. Gavrilis / G. Geser / H. Hollander / C. Meghini / F. Niccolucci / F. Nurra / C. Papatheodorou / J. Richards / P. Ronzino / R. Scopigno / M. Theodoridou / D. Tudhope / H. Wright, Enabling European Archaeological Research: The ARIADNE E-Infrastructure, Internet Archaeology 43 (2017), https://doi.org/10.11141/ia.43.11 (last access 19.12.2023).

Angjeli et al. (2014): A. Angjeli / A. Mac Ewan / V. Boulet, ISNI and VIAF – Transforming ways of trustfully consolidating identities, IFLA WLIC 2014, Lyon, France (2014), http://library.ifla.org/985/ (last access 19.12.2023).

Ast / Essler (2018): R. Ast / H. Essler, Anagnosis, Herculaneum, and the Digital Corpus of Literary Papyri, in: N. Reggiani (ed.), Digital Papyrology II: Case Studies on the Digital Edition of Ancient Greek Papyri, Berlin / Boston 2018, 63–74, https://doi.org/10.1515/9783110547450-003 (last access 19.12.2023).

Bagnall / Heath (2018): R. S. Bagnall / S. Heath, Roman Studies and Digital Resources, The Journal of Roman Studies 108 (2018), 171–189, https://doi.org/10.1017/S0075435818000874 (last access 19.12.2023).

Bambaci et al. (2019): L. Bambaci / F. Boschetti / R. D. Gratta, Qohelet Euporia: A Domain specific Language for the Encoding of the critical Apparatus, International Journal of Information Science and Technology 3/5 (2019), Article 5.

Barbera (2013): M. Barbera, Linked (open) Data at Web Scale: Research, social and engineering Challenges in the digital humanities, JLIS.It 4/1 (2013), 91–101.

Barker (2020): E. Barker, What would Indy do? Resisting post-truth through the practice of annotation, in: C. Roosevelt (ed.), Spatial Webs: Mapping Anatolian Pasts for Research and the Public, Koç University Research Centre for Anatolian Civilizations 2020, 5–31, http://oro.open.ac.uk/69161/ (last access 19.12.2023).

Barker (2021): E. Barker, Early steps in digitally mapping Pausanias's Description of Greece, Pelagios (2021), https://medium.com/pelagios/early-steps-in-digitally-mapping-pausaniass-description-of-greece-548301b2a54d (last access 19.12.2023).

Barker et al. (2021): E. Barker / M. O'Doherty / L. Isaksen, Introduction, International Journal of Humanities and Arts Computing 15/1–2 (2021), 1–4, https://doi.org/10.3366/ijhac.2021.0259 (last access 19.12.2023).

Baumann (2013): R. Baumann, The Son of Suda On-Line, The Digital Classicist (2013), 91–106.

Berman et al. (2016): M. L. Berman / J. Åhlfeldt / M. Wick, Historical Gazetteer System Integration: CHGIS, Regnum francorum Online, and GeoNames, in: M. L. Berman / R. Mostern / H. Southall (eds.), Placing Names: Enriching and Integrating Gazetteers, Bloomington / Indianapolis 2016, 110–125, http://lup.lub.lu.se/record/f279fef4-4d0b-4242-9b23-328a8eb9d816 (last access 19.12.2023).

Bernasconi et al. (2023): E. Bernasconi / M. Ceriani / D. Di Pierro / S. Ferilli / D. Redavid, Linked Data Interfaces: A Survey, Information 14/9 (2023), Article 9, https://doi.org/10.3390/info14090483 (last access 19.12.2023).

Berners-Lee (1998): T. Berners-Lee, Semantic Web Road map, W3C – Design Issues (1998), https://www.w3.org/DesignIssues/Semantic.html (last access 19.12.2023).

Berners-Lee (2010): T. Berners-Lee, Linked Data, WSC – Design Issues, W3C (2010), https://www.w3.org/DesignIssues/LinkedData.html (last access 19.12.2023).

Binding et al. (2019): C. Binding / D. Tudhope / A. Vlachidis, A study of semantic integration across archaeological data and reports in different languages, Journal of Information Science 45/3 (2019), 364–386, https://doi.org/10.1177/0165551518789874 (last access 19.12.2023).

Biston-Moulin / Thiers (2018): S. Biston-Moulin / C. Thiers, The Karnak Project: A Comprehensive Edition of the Largest Ancient Egyptian Temple, in: A. De Santis / I. Rossi (eds.), Crossing Experiences in Digital Epigraphy. From Practice to Discipline, Warsaw / Berlin 2018, 155–164, https://doi.org/10.1515/9783110607208-013 (last access 19.12.2023).

Bodard et al. (2017): G. Bodard / H. Cayless / M. Depauw / L. Isaksen / F. Lawrence / S. Rahtz, Standards for Networking Ancient Person data: Digital approaches to problems in prosopographical space, Digital Classics Online 3/2 (2017), 28–43, https://doi.org/10.11588/dco.2017.0.37975 (last access 19.12.2023).

Broux (2017a): Y. Broux, Ancient Profiles Exploited: First Results of Named Entity Recognition applied to Latin Inscriptions, U Schyłku Starożytności – Studia Źródłoznawcze XVI (2017), 11–33.

Broux (2017b): Y. Broux, Egyptian Names and Networks in Trismegistos (800 BC – AD 800), in: G. Rosati / M. C. Guidotti (eds.), Proceedings of the XI International Congress of Egyptologists: Florence, Italy 23–30 August 2015, Oxford 2017, 64–68.

Broux (2017c): Y. Broux, Towards a Universal Facebook of the Ancient World, in: S. Orlandi / R. Santucci / F. Mambrini / P. M. Liuzzo (eds.), Digital and Traditional Epigraphy in Context: Proceedings of the EAGLE 2016 International Conference, Rome 2017, 343–352, http://www.editrice-sapienza.it/node/7610 (last access 19.12.2023).

Broux / Depauw (2015): Y. Broux / M. Depauw, Developing Onomastic Gazetteers and Prosopographies for the Ancient World Through Named Entity Recognition and Graph Visualization: Some Examples from Trismegistos People, in: L. M. Aiello / D. McFarland (eds.), Social Informatics, Cham 2015, 304–313, https://doi.org/10.1007/978-3-319-15168-7_38 (last access 19.12.2023).

Brown (2022): S. Brown, Same Difference: Identity and Diversity in Linked Open Cultural Data, International Journal of Humanities and Arts Computing 16/1 (2022), 1–16, https://doi.org/10.3366/ijhac.2022.0273 (last access 19.12.2023).

Bruseker et al. (2017): G. Bruseker / N. Carboni / A. Guillem, Cultural Heritage Data Management: The Role of Formal Ontology and CIDOC CRM, in M. L. Vincent / V. M. López-Menchero Bendicho / M. Ioannides / T. E. Levy (eds.), Heritage and Archaeology in the Digital Age: Acquisition, Curation, and Dissemination of Spatial Cultural Heritage Data, Cham 2017, 93–131, https://doi.org/10.1007/978-3-319-65370-9_6 (last access 19.12.2023).

Buchanan et al. (2016): S. Buchanan / P. Golden / E. C. Kansa / A. Rabinowitz / R. Shaw, Periods, Organized (PeriodO): A Gazetteer of Period Assertions for Linking and Visualizing Periodized Data (2016), https://hcommons.org/deposits/item/hc:12295/ (last access 19.12.2023).

Buil-Aranda et al. (2013): C. Buil-Aranda / A. Hogan / J. Umbrich / P.-Y. Vandenbussche, SPARQL Web-Querying Infrastructure: Ready for Action?, in: H. Alani / L. Kagal / A. Fokoue / P. Groth / C. Biemann / J. X. Parreira / L. Aroyo / N. Noy / C. Welty / K. Janowicz (eds.), The Semantic Web – ISWC 2013, Berlin / Heidelberg 2013, 277–293, https://doi.org/10.1007/978-3-642-41338-4_18 (last access 19.12.2023).

Buzi et al. (2018): P. Buzi / J. Bogdani / F. Berno, The 'PAThs' Project: An Effort to Represent the Physical Dimension of Coptic Literary Production (Third–Eleventh centuries), Comparative Oriental Manuscript Studies Bulletin 4/1 (2018), 39–58.

Calvanese et al. (2016): D. Calvanese / P. Liuzzo / A. Mosca / J. Remesal / M. Rezk / G Rull, Ontology-based Data Integration in EPNet: Production and Distribution of Food during the Roman Empire, Engineering Applications of Artificial Intelligence 51 (2016), 212–229, https://doi.org/10.1016/j.engappai.2016.01.005 (last access 20.12.2023).

Cayless (2011): H. A. Cayless, The Papyrological Navigator: Project Integration with RDF, Digital Humanities 2011, Stanford University, Stanford, CA, USA June 19–22, 2011, 31–34, https://dh2011.stanford.edu/wp-content/uploads/2011/05/DH2011_BookOfAbs.pdf (last access 20.12.2023).

Cayless (2019): H.A. Cayless, Sustaining Linked Ancient World Data, in: M. Berti (ed.), Digital Classical Philology, Berlin / Boston 2019, 35–50, https://doi.org/10.1515/9783110599572-004 (last access 20.12.2023).

Celano (2018): G. G. A. Celano, An Automatic Morphological Annotation and Lemmatization for the IDP Papyri, in: N. Reggiani (ed.), Digital Papyrology II: Case Studies on the Digital Edition of Ancient Greek Papyri, Berlin / Boston 2018, 139–148, https://doi.org/10.1515/9783110547450-008 (last access 20.12.2023).

Crane (2014): G. R. Crane, The Hellespont Project: Integrating Arachne and Perseus (100476), Tufts University 2014, http://dx.doi.org/10.17613/M6HH2Q (last access 20.12.2023).

Delmas-Glass / Sanderson (2020): E. Delmas-Glass / R. Sanderson, Fostering a Community of PHAROS Scholars through the Adoption of Open Standards, Art Libraries Journal 45/1 (2020), 19–23, https://doi.org/10.1017/alj.2019.32 (last access 20.12.2023).

Depauw (2018): M. Depauw, Trismegistos: Optimizing Interoperability for Texts from the Ancient World, in: A. De Santis / I. Rossi (eds.), Crossing Experiences in Digital Epigraphy: From Practice to Discipline, Warsaw / Berlin 2018, 193–201, https://doi.org/10.1515/9783110607208-016 (last access 20.12.2023).

Depauw / Gheldof (2014): M. Depauw / T. Gheldof, Trismegistos: An Interdisciplinary Platform for Ancient World Texts and Related Information, in: Ł. Bolikowski / V. Casarosa / P. Goodale / N. Houssos / P. Manghi / J. Schirrwagen (eds.), Theory and Practice of Digital Libraries–TPDL 2013 Selected Workshops (Vol. 416), Cham 2014, 40–52, https://doi.org/10.1007/978-3-319-08425-1_5 (last access 20.12.2023).

Depauw / Stolk (2015): M. Depauw / J. Stolk, Linguistic Variation in Greek Papyri: Towards a New Tool for Quantitative Study, Greek, Roman, and Byzantine Studies 55 (2015), 196–220.

Elliott / Gillies (2009): T. Elliott / S. Gillies, Digital Geography and Classics, Digital Humanities Quarterly 3/1 (2009).

Geonames (n.d.): Geonames, About GeoNames, Geonames, https://www.geonames.org/about.html (last access 20.12.2023).

Gerth (2016): P. Gerth, ARIADNE D14.2: Pilot Deployment Experiments, ARIADNE (2016), http://legacy.ariadne-infrastructure.eu/wp-content/uploads/2019/01/D14.2_Pilot_deployment_experiments_Final.pdf (last access 20.12.2023).

Gerth et al. (2016): P. Gerth / D. M. Beck / W. Schmidle / S. Cuy, Sculptures in the Semantic Web Using Semantic Technologies for the Deep Integration of Research Items in ARIADNE, in: M. Matsumoto / E. Uleberg (eds.), Proceedings of the 44th Annual Conference on Computer Applications and Quantitative Methods in Archaeology CAA 2016 – Oceans of data (2016), 1122.

Geser (2016): G. Geser, ARIADNE WP15 Study: Towards a Web of Archaeological Linked Open Data, ARIADNE (2016), http://legacy.ariadne-infrastructure.eu/wp-content/uploads/2019/01/ARIADNE_archaeological_LOD_study_10-2016-1.pdf (last access 20.12.2023).

Gibson / Kahn (2016): L. Gibson / R. Kahn, Digital Museums in the 21st Century: Global Microphones or Universal Mufflers?, Museological Review 20 (2016), 39–51.

Gillies (2015): S. Gillies, Technical Introduction to Places, Pleiades: A Gazetteer of Past Places (2015), https://pleiades.stoa.org/help/technical-intro-places (last access 20.12.2023).

Golden / Shaw (2016): P. Golden / R. Shaw, Nanopublication beyond the Sciences: The PeriodO period gazetteer, PeerJ Computer Science 2 (2016), https://doi.org/10.7717/peerj-cs.44 (last access 20.12.2023).

Granados-García (2020): P. L. Granados-García, Cultural Contact in Early Roman Spain through Linked Open Data, PhD, The Open University (2020), http://oro.open.ac.uk/73887/ (last access 20.12.2023).

Grossner / Hill (2017): K. Grossner / T. Hill, From Linking Places to a Linked Pasts Network (2017), http://kgeographer.com/pubs/LinkedPastsNetwork_7Dec.pdf (last access 20.12.2023).

Grossner / Mostern (2021): K. Grossner / R. Mostern, Linked Places in World Historical Gazetteer, Proceedings of the 5th ACM SIGSPATIAL International Workshop on Geospatial Humanities (2021), 40–43, https://doi.org/10.1145/3486187.3490203 (last access 20.12.2023).

Gruber (2016): E. Gruber, Using XForms to Create, Publish, and Manage Linked Open Data, XML London 2016 Conference Proceedings (2016), 99–107.

Gruber (2018): E. Gruber, Linked Open Data for Numismatic Library, Archive and Museum Integration, in: M. Matsumoto / E. Uleberg (eds.), Oceans of Data: Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology, Oxford 2018, 55–62, https://www.archaeopress.com/Archaeopress/Products/9781784917302 (last access 20.12.2023).

Gruber / Meadows (2021): E. Gruber / A. Meadows, Numismatics and Linked Open Data, ISAW Papers 20 (2021), http://dlib.nyu.edu/awdl/isaw/isaw-papers/20-6/ (last access 20.12.2023).

Gunter et al. (2019): V. Gunter / S. Matthias / G. Vogeler, Data Exchange in Practice: Towards a Prosopographical API (Preprint), Biographical Data in a Digital World 2019, Varna, Bulgaria, https://hcommons.org/deposits/item/hc:29017/ (last access 20.12.2023).

Hannemann / Kett (2010): J. Hannemann / J. Kett, Linked Data for Libraries, World Library and Information Congress: 76th IFLA General Conference and Assembly, Gothenburg, Sweden (2010), https://www.ifla.org/past-wlic/2010/149-hannemann-en.pdf (last access 20.12.2023).

Hawes (2020): G. Hawes, User testers needed!, MYTHLAB (2020), https://www.manto-myth.org/blog/user-testers-needed (last access 20.12.2023).

Heath (2018): S. Heath, SPARQL as a First Step for Querying and Transforming Numismatic Data: Examples from Nomisma.org, in: A. Meadows / F. Duryat / S. Glenn (eds.), Alexander the Great. A Linked Open World, Bordeaux 2018, 35–52.

Hedrick (2018): C. Hedrick, Review: Searching EAGLE (The Electronic Archive of Greek and Latin Epigraphy), Society for Classical Studies (2018), https://classicalstudies.org/scs-blog/charles-hedrick/review-searching-eagle-electronic-archive-greek-and-latin-epigraphy (last access 20.12.2023).

Hickey / Toves (2014): T. B. Hickey / J. A. Toves, Managing Ambiguity In VIAF, D-Lib Magazine 20/7–8 (2014), https://doi.org/10.1045/july2014-hickey (last access 20.12.2023).

Holterhoff (2017): K. Holterhoff, From Disclaimer to Critique: Race and the Digital Image Archivist, Digital Humanities Quarterly 11/3 (2017).

Horne (2020a): R. Horne, Beyond lists: Digital Gazetteers and Digital History, The Historian 82/1 (2020), 37–50, https://doi.org/10.1080/00182370.2020.1725992 (last access 20.12.2023).

Horne (2020b): R. Horne, Mapping Power: Using HGIS and Linked Open Data to Study Ancient Greek Garrison Communities, in: C. Travis / F. Ludlow / F. Gyuris (eds.), Historical Geography, GIScience and Textual Analysis: Landscapes of Time and Place, Cham 2020, 213–227, https://doi.org/10.1007/978-3-030-37569-0_13 (last access 20.12.2023).

Hughes et al. (2016): L. Hughes / P. Constantopoulos / C. Dallas, Digital Methods in the Humanities: Understanding and Describing their Use across the Disciplines, in: S. Schreibman / R. G. Siemens / J. Unsworth (eds.), A new companion to digital humanities, Chichester / Malden 2016, 150–170.

Hyvönen et al. (2014): E. Hyvönen / J. Tuominen / M. Alonen / E. Mäkelä, Linked Data Finland: A 7-star Model and Platform for Publishing and Re-using Linked Datasets, in: V. Presutti / E. Blomqvist / R. Troncy / H. Sack / I. Papadakis / A. Tordai (eds.), The Semantic Web: ESWC 2014 Satellite Events, Cham 2014, 226–230, https://doi.org/10.1007/978-3-319-11955-7_24 (last access 20.12.2023).

International Organization for Standardization (2011): International Organization for Standardization, ISO 25964-1:2011(en), Information and Documentation – Thesauri and Interoperability with other Vocabularies – Part 1: Thesauri for Information Retrieval, International Organization for Standardization (2011), https://www.iso.org/obp/ui/#iso:std:iso:25964:-1:ed-1:v1:en (last access 20.12.2023).

Isaksen (2011): L. Isaksen, Archaeology and the Semantic Web, PhD, University of Southampton (2011), http://eprints.soton.ac.uk/206421 (last access 20.12.2023).

Isaksen et al. (2018): L. Isaksen / G. del Rio Riande / R. De León / N. Hernández, Getting to Grips with Semantic and Geo-annotation using Recogito 2, Digital Humanities 2018, Mexico City, https://dh2018.adho.org/en/getting-to-grips-with-semantic-and-geo-annotation-using-recogito-2/ (last access 20.12.2023).

Jordanous et al. (2012): A. Jordanous / K. F. Lawrence / M. Hedges / C. Tupman, Exploring Manuscripts: Sharing Ancient Wisdoms across the Semantic Web, Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics (2012), 1–12, https://doi.org/10.1145/2254129.2254184 (last access 20.12.2023).

Kahn et al. (2021): R. Kahn / L. Isaksen / E. Barker / R. Simon / P. de Soto / V. Vitale, Pelagios – Connecting Histories of Place. Part II: From Community to Association, International Journal of Humanities and Arts Computing 15/1–2 (2021), 85–100, https://doi.org/10.3366/ijhac.2021.0263 (last access 20.12.2023).

Kansa et al. (2018): E. C. Kansa / S. W. Kansa / J. J. Wells / S. J. Yerka / K. N. Myers / R. C. DeMuth / T. G. Bissett / D. G. Anderson, The Digital Index of North American Archaeology: Networking Government Data to Navigate an Uncertain Future for the Past, Antiquity 92/362 (2018), 490–506, https://doi.org/10.15184/aqy.2018.32 (last access 20.12.2023).

Koch / Koch (2017): W. Koch / G. Koch, Aggregation and Management of Metadata in the Context of Europeana. Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare 70/2 (2017), Article 2.

Koho et al. (2023): M. Koho / L. P. Coladangelo / L. Ransom / D. Emery, Wikibase Model for Premodern Manuscript Metadata Harmonization, Linked Data Integration, and Discovery, Journal on Computing and Cultural Heritage 16/3 (2023), 56:1–56:25, https://doi.org/10.1145/3594723 (last access 21.12.2023).

Lawrence / Bodard (2015): K. F. Lawrence / G. Bodard, Prosopography is Greek for Facebook: The SNAP:DRGN Project, Proceedings of the ACM Web Science Conference (2015), 44:1–44:2, https://doi.org/10.1145/2786451.2786496 (last access 21.12.2023).

Lee (2011): C. A. Lee, A Framework for Contextual Information in Digital Collections, Journal of Documentation 67/1 (2011), 95–143, https://doi.org/10.1108/00220411111105470 (last access 21.12.2023).

Limp (2011): W. F. Limp, Web 2.0 and Beyond, or On the Web, Nobody Knows You're an Archaeologist, in: E. C. Kansa / S. W. Kansa / E. Watrall (eds.), Archaeology 2.0: New Approaches to Communication and Collaboration, Los Angeles 2011, 265–280, https://escholarship.org/uc/item/1r6137tb (last access 21.12.2023).

Linked Art Contributors (n.d.-a): Linked Art Contributors, CIDOC-CRM Class Analysis, Linked Art, https://linked.art/model/profile/class_analysis.html (last access 21.12.2023).

Linked Art Contributors (n.d.-b): Linked Art Contributors, LOUD: Linked Open Usable Data, Linked Art, https://linked.art/loud/ (last access 21.12.2023).

Liu et al. (2017): D. Liu / A. Bikakis / A. Vlachidis, Evaluation of Semantic Web Ontologies for Modelling Art Collections, in: M. Kirikova / K. Nørvåg / G. A. Papadopoulos / J. Gamper / R. Wrembel / J. Darmont / S. Rizzi (eds.), New Trends in Databases and Information Systems, Cham 2017, 343–352, https://doi.org/10.1007/978-3-319-67162-8_34 (last access 21.12.2023).

Liuzzo et al. (2017): P. Liuzzo / F. Mambrini / P. Franck, Storytelling and Digital Epigraphy-Based Narratives in Linked Open Data, in: M. Ioannides / N. Magnenat-Thalmann / G. Papagiannakis (eds.), Mixed Reality and Gamification for Cultural Heritage, Springer International Publishing 2017, 507–523, https://doi.org/10.1007/978-3-319-49607-8_20 (last access 21.12.2023).

Mannocci et al. (2014): A. Mannocci / V. Casarosa / P. Manghi / F. Zoppi, The Europeana Network of Ancient Greek and Latin Epigraphy Data Infrastructure, in: S. Closs / R. Studer / E. Garoufallou / M.-A. Sicilia (eds.), Metadata and Semantics Research, Springer International Publishing 2014, 286–300, https://doi.org/10.1007/978-3-319-13674-5_27 (last access 21.12.2023).

May et al. (2012): K. May / C. Binding / D. Tudhope / S. Jeffrey, Semantic Technologies Enhancing Links and Linked Data for Archaeological Resources, in: M. Zhou / I. Romanowska / Z. Wu / P. Xu / P. Verhagen (eds.), Revive the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 39th International Conference, Beijing, April 12–16, Amsterdam 2012, 261–272, http://proceedings.caaconference.org/paper/29_may_et_al_caa2011/ (last access 21.12.2023).

Mayr et al. (2016): P. Mayr / D. Tudhope / S. D. Clarke / M. L. Zeng / X. Lin, Recent Applications of Knowledge Organization Systems: Introduction to a Special Issue, International Journal on Digital Libraries 17/1 (2016), 1–4, https://doi.org/10.1007/s00799-015-0167-x (last access 21.12.2023).

Meghini et al. (2017): C. Meghini / R. Scopigno / J. Richards / H. Wright / G. Geser / S. Cuy / J. Fihn / B. Fanini / H. Hollander / F. Niccolucci / A. Felicetti / P. Ronzino / F. Nurra / C. Papatheodorou / D. Gavrilis / M. Theodoridou / M. Doerr / D. Tudhope / C. Binding / A. Vlachidis, ARIADNE: A Research Infrastructure for Archaeology, Journal on Computing and Cultural Heritage 10/3 (2017), 18:1–18:27, https://doi.org/10.1145/3064527 (last access 21.12.2023).

Meroño-Peñuela et al. (2014): A. Meroño-Peñuela / A. Ashkpour / M. van Erp / K. Mandemakers / L. Breure / A. Scharnhorst / S. Schlobach / F. van Harmelen, Semantic Technologies for Historical research: A survey, Semantic Web 6/6 (2014), 539–564, https://doi.org/10.3233/SW-140158 (last access 21.12.2023).

Middle et al. (2022): S. Middle / R. Horne / D. A. McMeekin / C. Zuanni / A. Butterworth, Geographies of Place in Digital Art History, International Journal of Humanities and Arts Computing 16/1 (2022), 94–109, https://doi.org/10.3366/ijhac.2022.0279 (last access 21.12.2023).

Modest / Lelijveld (2018): W. Modest / R. Lelijveld (eds.), Words Matter: An Unfinished Guide to Word Choices in the Cultural Sector, Tropenmuseum, Afrika Museum, Museum Volkenkunde, Wereldmuseum 2018, https://www.materialculture.nl/sites/default/files/2018-08/words_matter-.pdf.pdf (last access 21.12.2023).

Niccolucci / Hermon (2017): F. Niccolucci / S. Hermon, Expressing Teliability with CIDOC CRM, International Journal on Digital Libraries 18/4 (2017), 281–287, https://doi.org/10.1007/s00799-016-0195-1 (last access 21.12.2023).

OCLC (2019): OCLC, Scholars' Contributions to VIAF, OCLC Research (2019), https://www.oclc.org/research/themes/data-science/viaf-scholars.html (last access 21.12.2023).

Open Data Institute (2018): Open Data Institute, Open Standards for Data, Open Standards for Data Guidebook (2018), https://standards.theodi.org/ (last access 21.12.2023).

Orlandi et al. (2014): S. Orlandi / L. M. C. Giberti / R. Santucci, EAGLE: Europeana Network of Ancient Greek and Latin Epigraphy, Making the Ancient Inscriptions Accessible, Lexicon Philosophicum: International Journal for the History of Texts and Ideas 2 (2014), Article 2.

Pagé-Perron (2017): É. Pagé-Perron, Expanding Digital Assyriology With Open Access and Machine Learning: A Cuneiform Digital Library Initiative White Paper for the Global Philology Project, Global Philology Open Conference (2017), https://doi.org/10.17613/M6MP5H (last access 21.12.2023).

Papyri.info (n.d.): Papyri.info, About Papyri.info, Papyri.Info, http://papyri.info/docs/about (last access 21.12.2023).

Polczynski / Polczynski (2022): M. Polczynski / M. Polczynski, Lessons Learned from Using Historical Maps to Create a Digital Gazetteer of Historical Places, International Journal of Cartography 8/3 (2022), 326–342, https://doi.org/10.1080/23729333.2021.2007444 (last access 21.12.2023).

Polis / Razanajao (2016): S. Polis / V. Razanajao, Ancient Egyptian Texts in Context. Towards a Conceptual Data Model (the Thot Data Model – Tdm), Bulletin of the Institute of Classical Studies 59/2 (2016), 24–41, https://doi.org/10.1111/j.2041-5370.2016.12036.x (last access 21.12.2023).

Prag / Chartrand (2018): J. R. W. Prag / J. Chartrand, I. Sicily: Building a Digital Corpus of the Inscriptions of Ancient Sicily, in: A. De Santis / I. Rossi (eds.), Crossing Experiences in Digital Epigraphy, Warsaw / Berlin 2018, 240–252, https://doi.org/10.1515/9783110607208-020 (last access 21.12.2023).

Prandoni et al. (2017): C. Prandoni / A. Fresa / N. Alfarano / G. Amato / F. Zoppi / A. Mannocci / P. Liuzzo / F. Mambrini / S. Orlandi / R. Santucci, Searching Inscriptions through the EAGLE Portal, in: S. Orlandi / R. Santucci / F. Mambrini / P. M. Liuzzo (eds.), Digital and Traditional Epigraphy in Context: Proceedings of the EAGLE 2016 International Conference, Rome 2017, 173–185, http://www.editricesapienza.it/node/7610 (last access 21.12.2023).

Rabinowitz (2014): A. Rabinowitz, It's about Time: Historical Periodization and Linked Ancient World Data, ISAW Papers 7 (2014), http://dlib.nyu.edu/awdl/isaw/isaw-papers/7/rabinowitz/ (last access 21.12.2023).

Rabinowitz (2021): A. Rabinowitz, Time for Linked Open Data, ISAW Papers 20 (2021), http://dlib.nyu.edu/awdl/isaw/isaw-papers/20-3/ (last access 09.02.2024).

Rabinowitz et al. (2016): A. Rabinowitz / R. Shaw / S. Buchanan / P. Golden / E. Kansa, Making Sense of the Ways We Make Sense of the Past: The PeriodO Project, Bulletin of the Institute of Classical Studies 59/2 (2016), 42–55, https://doi.org/10.1111/j.2041-5370.2016.12037.x (last access 09.02.2024).

Rabinowitz et al. (2018): A. Rabinowitz / R. Shaw / P. Golden, Making up for Lost Time: Digital Epigraphy, Chronology, and the PeriodO Project, in: A. De Santis / I. Rossi (eds.), Crossing Experiences in Digital Epigraphy, Warsaw / Berlin 2018, 202–215, https://doi.org/10.1515/9783110607208-017 (last access 09.02.2024).

Rantala et al. (2023): H. Rantala / A. Ahola / E. Ikkala / E. Hyvönen, How to Create Easily a Data Analytic Semantic Portal on Top of a SPARQL Endpoint: Introducing the Configurable Sampo-UI Framework, Proceedings of the 8th International Workshop on the Visualization and Interaction for Ontologies, Linked Data and Knowledge Graphs Co-Located with the 22nd International Semantic Web Conference (ISWC 2023), CEUR Workshop Proceedings 3508 (2023), https://ceur-ws.org/Vol-3508/paper3.pdf (last access 09.02.2024).

Reggiani (2017): N. Reggiani, Digital Papyrology I: Methods, Tools and Trends, Berlin / Boston 2017. https://doi.org/10.1515/9783110547474 (last access 09.02.2024).

Rocco (2017): A. Rocco, EDB 2.0: How Eagle Europeana Project Improved the Epigraphic Database Bari, in: S. Orlandi / R. Santucci / F. Mambrini / P. M. Liuzzo (eds.), Digital and Traditional Epigraphy in Context: Proceedings of the EAGLE 2016 International Conference, Rome 2017, 115–130, http://www.editricesapienza.it/node/7610 (last access 09.02.2024).

Ross et al. (2015): S. Ross / B. Ballsun-Stanton / A. Sobotkova / P. Crook, Building the Bazaar: Enhancing Archaeological Field Recording Through an Open Source Approach, in: A. T. Wilson / B. Edwards (eds.), Open Source Archaeology: Ethics and Practice, Warsaw / Berlin 2015, 111–129, https://doi.org/10.1515/9783110440171-009 (last access 09.02.2024).

Santucci et al. (2016): R. Santucci / S. Orlandi / A. Ferraro / G. Ragone / L. M. C. Giberti / C. Prandoni / A. Fresa, EAGLE Sustainability Plan: De 6.6, Europeana (2016), https://www.eagle-network.eu/wp-content/uploads/2016/05/EAGLE_D6.6_Sustainability_Plan_V1.0.pdf (last access 09.02.2024).

Schmidt et al. (2022): S. C. Schmidt / F. Thiery / M. Trognitz, Practices of Linked Open Data in Archaeology and Their Realisation in Wikidata, Digital 2/3 (2022), Article 3, https://doi.org/10.3390/digital2030019 (last access 12.02.2024).

Schneider et al. (2018): P. Schneider / J. Jones / T. Hiltmann / T. Kauppinen, Challenge-derived Design Practices for a Semantic Gazetteer for Medieval and Early Modern Places, Semantic Web 1 (2018), 1–20, https://doi.org/10.3233/SW-200394 (last access 12.02.2024).

Schweizer / Geer (2019): T. Schweizer / B. Geer, Gravsearch: Transforming SPARQL to Query Humanities Data, Semantic Web Journal, Special Issue Cultural Heritage (2019), http://www.semantic-web-journal.net/system/files/swj2412.pdf (last access 12.02.2024).

Scriba / Stockinger (2016): G. Scriba / U. Stockinger, Timeo interretem et dona ferentem. On ARACHNE and the Potential and Limits of Publishing Archaeological Catalogues Online, in I. Pinto Coelho / J. Bento Torres / L. Serrão Gil / T. Ramos (eds.), Entre ciência e cultura: Da interdisciplinaridade à transversalidade da arqueologia. Actas das VIII Jornadas de Jovens em Investigação Arqueológica, Lisbon (2016), 497–502.

Shaw et al. (2018): R. Shaw / A. Rabinowitz / P. Golden, A Deep Gazetteer of Time Periods, Digital Humanities 2018, Mexico City (2018). https://dh2018.adho.org/en/a-deep-gazetteer-of-time-periods/ (last access 16.02.2024).

Simon et al. (2017): R. Simon / E. Barker / L. Isaksen / P. de Soto Cañamares, Linked Data Annotation Without the Pointy Brackets: Introducing Recogito 2, Journal of Map & Geography Libraries 13/1 (2017), 111–132, https://doi.org/10.1080/15420353.2017.1307303 (last access 16.02.2024).

Simon et al. (2015): R. Simon / E. Barker / L. Isaksen / P. de Soto Cañamares, Linking Early Geospatial Documents, One Place at a Time: Annotation of Geographic Documents with Recogito, E-Perimetron, 10/2 (2015), 49–59.

Simon et al. (2016a): R. Simon / L. Isaksen / E. Barker / P. de Soto Cañamares, Peripleo: A Tool for Exploring Heterogeneous Data through the Dimensions of Space and Time, The Code4Lib Journal 31 (2016), http://journal.code4lib.org/articles/11144 (last access 16.02.2024).

Simon et al. (2016b): R. Simon / L. Isaksen / E. Barker / P. de Soto Cañamares, The Pleiades Gazetteer and the Pelagios Project, in M. L. Berman / R. Mostern / H. Southall (eds.), Placing Names: Enriching and Integrating Gazetteers, Bloomington / Indianapolis (2016), 97–109.

Simon et al. (2014): R. Simon / P. Pilgerstorfer / L. Isaksen / E. Barker, Towards semi-automatic annotation of toponyms on old maps, E-Perimetron 9/3 (2014), 105–112.

Simon et al. (2019): R. Simon / V. Vitale / R. Kahn / E. Barker / L. Isaksen, Revisiting Linking Early Geospatial Documents with Recogito, E-Perimetron 14/3 (2019), Article 3.

Simou et al. (2017): N. Simou / A. Chortaras / G. Stamou / S. Kollias, Enriching and Publishing Cultural Heritage as Linked Open Data, in M. Ioannides / N. Magnenat-Thalmann / G. Papagiannakis (eds.), Mixed Reality and Gamification for Cultural Heritage, Springer International Publishing (2017), 201–223 https://doi.org/10.1007/978-3-319-49607-8_7 (last access 16.02.2024).

Smith-Yoshimura (2018): K. Smith-Yoshimura, Analysis of 2018 International Linked Data Survey for Implementers, The Code4Lib Journal 42 (2018), https://journal.code4lib.org/articles/13867 (last access 16.02.2024).

Smith-Yoshimura (2013): K. Smith-Yoshimura, First Scholars' Contributions to VIAF: Greek!, Hanging Together (2013, November 25), https://hangingtogether.org/?p=3455 (last access 16.02.2024).

Smith-Yoshimura (2014): K. Smith-Yoshimura, New Scholars' Contributions to VIAF: Syriac!, Hanging Together (2014, March 11), https://hangingtogether.org/?p=3634 (last access 16.02.2024).

Thaller (2020): M. Thaller, On Vagueness and Uncertainty in Historical Data, A Digital Ivory Tower (2020, July 8), https://ivorytower.hypotheses.org/88 (last access 16.02.2024).

Thiery / Engel (2016): F. Thiery / T. Engel, The Labelling System: A Bottom-up Approach for Enriched Vocabularies in the Humanities, in S. Campana / R. Scopigno / G. Carpentiero / M. Cirillo (eds.), CAA2015. Keep the Revolution Going: Proceedings of the 43rd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Oxford (2016), 259–268, https://www.archaeopress.com/Archaeopress/Products/9781784913373 (last access 29.02.2024).

Tolle / Wigg-Wolf (2016): K. Tolle / D. Wigg-Wolf, How To Move from Relational to 5 Star Linked Open Data – A Numismatic Example, in: S. Campana / R. Scopigno / G. Carpentiero / M. Cirillo (eds.), CAA2015. Keep the Revolution Going: Proceedings of the 43rd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Oxford (2016), 275–281. https://www.archaeopress.com/Archaeopress/Products/9781784913373 (last access 29.02.2024).

Torallas Tovar / Schubert (2019): S. Torallas Tovar / P. Schubert, The Campaign for Papyri.info, Securing the Future of Papyrology (2019, November 18), https://www.supportpapyri.info/campaign/ (last access 29.02.2024).

Trismegistos (n.d.): Trismegistos, Database Structure, Trismegistos (n.d.), https://www.trismegistos.org/about_databasestructure.php (last access 29.02.2024).

Trismegistos (2019): Trismegistos, What Changes on 1 January 2020?, Trismegistos (2019, July), https://www.trismegistos.org/keeptrismegistosalive_whatchanges.php (last access 29.02.2024).

Tuan (1975): Y.-F. Tuan, Place: An Experiential Perspective, Geographical Review 65/2 (1975), 151–165, https://doi.org/10.2307/213970 (last access 29.02.2024).

Tupman / Jordanous (2014): C. Tupman / A. Jordanous, Sharing Ancient Wisdoms across the Semantic Web Using TEI and Ontologies, in T. L. Andrews / C. Macé (eds.), Analysis of Ancient and Medieval Texts and Manuscripts: Digital Approaches, Vol. 1, Brepols Publishers (2014), 213–228, https://doi.org/10.1484/M.LECTIO-EB.5.102572 (last access 29.02.2024).

van Hooland / Verborgh (2014): S. van Hooland / R. Verborgh, Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish Your Metadata, Facet Publishing (2014).

Varga et al. (2018): R. Varga / A.-I. Pázsint / I. Boda / D. A. Deac, Romans 1 by 1: Overview of a Research Project, Digital Classics Online 4/2 (2018), 37–63, https://doi.org/10.11588/dco.2018.2.45878 (last access 29.02.2024).

Verreth (2017): H. Verreth, Trismegistos Places, a Geographical Index for All Latin Inscriptions, in S. Orlandi / R. Santucci / F. Mambrini / P. M. Liuzzo (eds.), Digital and Traditional Epigraphy in Context: Proceedings of the EAGLE 2016 International Conference, Rome (2017), 201–208, http://www.editricesapienza.it/node/7610 (last access 29.02.2024).

Vitale et al. (2021): V. Vitale / P. de Soto / R. Simon / E. Barker / L. Isaksen / R. Kahn, Pelagios – Connecting Histories of Place. Part I: Methods and Tools, International Journal of Humanities and Arts Computing 15/1–2 (2021), 5–32, https://doi.org/10.3366/ijhac.2021.0260 (last access 29.02.2024).

W3C Network Working Group (2004): W3C Network Working Group, Hypertext Transfer Protocol – HTTP/1.1, W3C – Protocols (2004, September 1), https://www.w3.org/Protocols/rfc2616/rfc2616.html (last access 08.03.2024).

Wilkinson et al. (2016): M. D. Wilkinson / M. Dumontier / Ij. J. Aalbersberg / G. Appleton / M. Axton / A. Baak / N. Blomberg / J.-W. Boiten / L. B. da Silva Santos / P. E. Bourne / J. Bouwman / A. J. Brookes / T. Clark / M. Crosas / I. Dillo / O. Dumon / S. Edmunds / C. T. Evelo / R. Finkers / A. Gonzalez-Beltran / A. J. G. Gray / P. Groth / C. Goble / J. S. Grethe / J, Heringa / P. A. C. 't Hoen / R. Hooft / T. Kuhn / R. Kok / J. Kok / S. J. Lusher / M. E. Martone / A. Mons / A. L. Packer / B. Persson / P. Rocca-Serra / M. Roos / R. van Schaik / S.-A. Sansone / E. Schultes / T. Sengstag / T. Slater / G. Strawn / M. A. Swertz / M. Thompson / J. van der Lei / E. van Mulligen / J. Velterop / A. Waagmeester / P. Wittenburg / K. Wolstencroft / J. Zhao / B. Mons, The FAIR Guiding Principles for Scientific Data Management and Stewardship, Scientific Data 3/1 (2016), https://doi.org/10.1038/sdata.2016.18 (last access 08.03.2024).

Zhao (2023): F. Zhao, A Systematic Review of Wikidata in Digital Humanities Projects, Digital Scholarship in the Humanities 38/2 (2023), 852–874, https://doi.org/10.1093/llc/fqac083 (last access 08.03.2024).

Author Contact Information[204]

Dr. Sarah Middle
Visiting Fellow, Classical Studies
The Open University
E-Mail: sarah.middle@open.ac.uk

204 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.

DIGITAL CLASSICS ONLINE

GIS analysis and the abandoned settlements in Pausanias' *Periegesis*: a methodological proposal

Marco Cornaglia

Abstract: The *Periegesis* of Pausanias, an invaluable account detailing the state of Greece in the 2nd century AD, represents an ideal example of a classical text suitable for a complex geospatial representation, due to the sheer amount and precision of the author's descriptions. More importantly, Pausanias' accounts of specific features, mostly identified by archaeological research, often adhere to precise literary and cultural criteria. This approach by the author may lead to several layers of analysis of the text's content. The distribution of abandoned settlements, in particular, may be indicative not only of the actual state of Roman-era Greece and the demographic evolution of the accounted regions, but also of the cultural goals of the author. This article focuses on the possibility of adopting a *GIS*-based methodology in the analysis and evaluation of the content of a classical text, starting with its digitization and subsequently isolating specific elements related to a particular phenomenon. Such features, in turn, underwent a process of textual analysis and data systematization, thus enabling the evaluation of conclusions potentially useful for the literary research.

Introduction

The rise of digital humanities has seen a strengthening of the ties between several fields of research and different types of information systems. More than mere tools, they have been rightly labeled as "humanities of the digital realm", focused on the study and representation of the human experience.[1] An invaluable asset in the use of digital tools for literary research is doubtlessly represented by their ability to rapidly perform complex searches through a vast amount of data. Specialized lexica and thesauri, many of which freely available online, represent an obvious example in this sense.[2] The creation of similar databases is in turn an ideal first step in the implementation of tools able to process the gathered data, highlight relations between single components from a repository, and elaborate scientifically valid conclusions. A promising field in textual data retrieval and analysis is represented by the adjustment of literary databases to a proper geographical context;[3] an approach that proved particularly versatile in the geographic representation of classical antiquity. A visual reconstruction of the spatial features of a text does indeed allow to adopt and explore a different approach in the study of a particular literary material, and to appreciate features such as specific traits of an ancient author, the chronological context, etc. *GIS* mapping tools represent the most effective choice for the creation of this particu-

1 Scheuermann / Kroeze (2017).

2 A common flaw in the creation of these digital libraries is represented, however, by the lack of a critical apparatus, particularly necessary, for example, for the proper use of Greek and Latin texts; Damon (2016).

3 Franco Moretti's seminal work in analyzing the applications of a literary atlas adequately illustrates the potential of investigating the social and cultural implications behind both a single writer's literary geography and a comparison between the approaches pursued by different authors; Moretti (1997). See Taylor / Gregory (2022) for a case study regarding the analysis of a literary digital atlas; for an assessment of the digitization of literary maps, Cooper et al. (2017).

lar type of database due to their ability of both storing a large amount of complex data and combining them in different ways in order to evaluate particular features from a case study.[4] Indeed, the demand for digital maps has recently soared due to the versatility of spatial representations, which are capable of intuitively illustrating the outcomes of a research and generating further insights from basic inquiries.[5]

The potential of geographic databases in classical research has been enthusiastically exploited. This is hardly surprising, as this field ideally combines historical, archaeological, and literary research with a cross-disciplinary approach. The foundation for much of these analyses is provided by Richard Talbert's two-volume *Barrington Atlas of the Greek and Roman World*[6], a seminal text in the study of classical antiquity. Most notably, it provides the identifications for the *Pleiades* digital atlas, a database focused on the classical world and quite possibly the most well-known case of a digital atlas with precise ancient literary references.[7] This was used in turn for the creation of further similar projects, such as the *DARE* (*Digital Atlas of the Roman Empire*), edited by the Lund University.[8] *Pleiades* was also a fundamental asset in the identification of ancient sites for projects such as *PELAGIOS* (*Pelagios Enable Linked Ancient Geodata In Open Systems*)[9] or *ToposText*[10]. The latter, edited by the Aikaterini Laskaridis Foundation, provides several tools for the research on classical texts, including a detailed digital atlas, and does offer a visual representation and a description of 'ancillary' elements related to the sites, such as, for example, the temples of an ancient settlement. However, not all these features are always taken into consideration. Moreover, the atlas, like the other aforementioned projects, usually provides no more than one identification, when several alternatives are often necessary when analyzing features such as those from an ancient landscape, whose location may very well be uncertain. All these sites do represent invaluable tools for the analysis of ancient topography; nevertheless, they all share a few limitations, as they do not generally provide multiple identifications for uncertain locations and are essentially committed to the analysis of specific areas in the 'broader' sense. They do indeed pin ancient locations on their geographic context, but fail to account in a satisfactory way for all the unidentified elements and for those that occasionally lied *inside* the described sites, essentially neglecting a more complete analysis of the features related to the single points of interest. These features may include, for example, a single temple that lied within a settlement or, even more specifically, the sculptures that were part of a temple; elements generally not systematically considered or analyzed by the aforementioned projects.

Pausanias' *Periegesis* offers a detailed description of the landscape and sites of Greece under Roman rule and is one of the most important sources in the definition of the ancient layout of the region. The aforementioned websites often refer to locations described by this author, albeit with the already described limitations. However, these platforms are not exclusively focused on Pausanias, nor on specific features such as the state – whether inhabited or deserted – of Greek settlements during the Roman era, as described by the author. The digital Pausanias project[11], on the other hand, is indeed focused on referring Pausanias' text to its proper landscape by converting the information from the *Periegesis* into

4 Refer to Martí Henneberg (2023) and Lünen / Travis (2013) for an analysis of historical *GIS* in particular. Bodenhamer et al. (2010), for an assessment on the current and foreseeable impact of the *GIS* on humanities in general.

5 Knowles / Hillier (2008); Travis et al. (2020).

6 Talbert / Bagnall (2000).

7 https://pleiades.stoa.org/ (last access 11.11.2024).

8 https://www.lunduniversity.lu.se/lup/publication/dcaf05b8-a13b-4024-a47b-325dbb68a035 (last access 11.11.2024).

9 http://pelagios-project.blogspot.com.es/ (last access 11.11.2024).

10 https://topostext.org/ (last access 11.11.2024).

11 https://gis.periegesis.org/ (last access 11.11.2024).

geodata through the *Recogito* open-source platform.[12] The resulting map has been implemented by using the *ArcGIS* software, and the results have been uploaded to the *ArcGIS* online cloud. However, the representation of Pausanias' sites, for which further possible developments are nevertheless contemplated, is by now still incomplete, with its points of interest associated with very simple data (such as the toponym, type, and a few tags) while the description of the sites is mostly provided by links to pages on the *Brill's New Pauly*, an encyclopedia ancient history and classical studies, published by *Brill.com*.[13] Furthermore, the project presents the same issues of the aforementioned databases, as both structured descriptions for 'accessory' sites and possible alternatives in identifying still uncertain areas are missing. As for the yet *unidentified* features they are, again, not subjected to a thorough analysis which may highlight their characteristics and possible identification. While this is, of course, comprehensible, an approximate clarification of such occurrences may be useful, if based on elements provided by the author, in defining the broader context of an area under exam. Finally, it is crucial to grant the possibility of constantly improving and updating the collected data according to the progress of the research. Considering the aforementioned projects and their issues, the following prerequisites emerge in defining a functional map based on ancient Greek and Roman texts:

- The map must not only identify the described sites, but also provide a description of *related* elements.
- As for sites of uncertain identification, it is necessary to provide all possible localizations, along with the respective bibliographic references.
- Outright unidentified sites should also be taken into consideration and positioned according to the details provided by the classical author, albeit with the necessary caution due to the hypothetical nature of such identifications, in order to help to better contextualize the reconstructed landscape.
- The considerations of an author regarding the conditions of a specific area should be reported. This information is particularly important, for example, regarding the possible state of ruin of a location at the time of its description.
- All data related to the sites should be be categorized according to a defined taxonomy. This can be particularly useful in conducting a targeted survey based on features such as era, identification status, type, etc.

As part of my PhD thesis, I have created a multimedia *GIS* atlas of Greece in the Roman era, based on the data retrieved from the analysis of the ten books of Pausanias' *Periegesis*. This atlas,[14] freely available online,[15] provides a representation of the sites described by the author, each one placed on the base of a critical analysis of the text. This work, carried out through the use of the *ArcGIS* software,

12 https://recogito.pelagios.org/ (last access 11.11.2024).

13 https://referenceworks.brillonline.com/browse/brill-s-new-pauly (last access 11.11.2024).

14 This atlas was presented at the International PhD Colloquium *Sistema Binario III Colloquio Internazionale dei Dottorandi. Sulle molteplici prospettive del viaggio: dimensione reale e virtuale*, organized by the University of Macerata from March 30th to April 1st 2021. The publication of the proceedings of the event is still ongoing; see https://studiumanistici.unimc.it/it/site-news/eventi/2021/sistema-binario-iii-colloquio-internazionale-dei-dottorandi (last access 11.11.2024).

15 http://purl.org//framelab/pausanias/atlas (last access 11.11.2024).

does not aim to the recreation of a strictly *historical* or *archaeological* context,[16] as much as to the digitization of Pausanias' assessments from a *literary* point of view. The map takes into account the perspective of the classical text, which (while amply vindicated by the archaeological research,[17] in stark contrast with earlier late 19th century – early 20th century century reservations and criticism[18]) nevertheless is not free from curious inconsistencies and omissions identified by the to archaeological research.[19]

Any discrepancies of the areas in exam with the literary atlas are noted in the data insertion, but the perspective adopted in creating the atlas is strictly that of the author. The aim of the project was in fact to assemble a reliable digitization of an ancient text, and to provide the tools to rapidly perform several forms of complex evaluations on the collected accounts as reported by their author. The *Periegesis* may represent an excellent case study in this sense. Not only the text is embedded with a vast number of descriptions organically tied to a geographic context, but the nature of the data itself, or rather their features and position within the framework of the text, are pliable to multiple literary interpretations regarding the author and the described setting. Little is known regarding Pausanias himself; his words from the *Periegesis* practically represent the only source of information regarding his background and the specific time period we can rely on.[20] More importantly, the text is indicative of the *cultural* background and motives of the author, an aspect that was taken into consideration and that emerged frequently in the creation of the atlas.

As Pausanias carries on his narration of the cities, sanctuaries, works of art and natural features disseminated throughout the area he explores (roughly corresponding to modern southern continental Greece[21]), his preferences and cultural purposes become increasingly evident, and they have been reflected in the setting up of the digital map. The author is notorious for his thoroughly meticulous descriptions of the main points of interests (such as large cities and famous sanctuaries) of Greece. Nevertheless, he is equally famous for his tendency to seek out for the lesser-known curiosities, myths, and historical anecdotes,[22] in a quest that inevitably draws him far from the main routes of his time,[23] in search of isolated or little-known sites. A particular feature of the narration that emerged during the

16 Of course, the archaeological state of the art was a necessary and invaluable source of information I constantly used for the correct identification of the sites from the *Periegesis*. But while their identification was indeed possible thanks to the consultation of the archaeological data, the choice of the sites itself, their taxonomy and the relative data were solely based on the accounts from the *Periegesis*; see *infra* pp. 55–59.

17 Pausanias arguably represents one of the ancient sources on which archaeologists and scholars rely the most. More than that, the author can be considered *de facto* a decisive influence on classical archaeological methodology; see Stewart (2013) and Hutton (2005), 3.

18 See Habicht (1985), 167–169; 221.

19 Monuments such as the Stoa of Attalus (Hutton [2008], 292) in Athens or the Nymphaeum of Herodes Atticus in Olympia (Alcock et al. [2003], 186) are in fact conspicuously absent from the *Periegesis'* account.

20 The accounts and chronologically certain events mentioned within the *Periegesis* point to a period that roughly spans from 120 to 180 AD. This would place Pausanias' activity during the rule of the Antonine emperors, with his birth taking place roughly around 115 AD; see Pretzler (2013), 23.

21 The reasons behind the choice of this specific zone are not entirely clear. One possible explanation is that the territory included in the *Periegesis* roughly coincides with the area of the Delphic Amphictyony. Delphi is indeed one of the main focal points in Pausanias' narrative, but in that case, regions like Opuntian Locris, Doris, Euboea and Thessaly should also be present, while they are conspicuously absent in the text; see Bearzot (1988). Pausanias' Greece might also coincide with the Greece of the city-states. This could explain, for example, the absence of regions such as Aetolia, characterized more by the presence of ἔθνη than cities; another possible explanation is represented by the distribution of the panhellenic sanctuaries, so crucial in Pausanias' descriptions; see Bultrighini (1990), 295; 300. Finally, the area described by Pausanias roughly coincides with the boundaries of the Roman province of Achaea; see Hutton (2005), 61.

22 See Bruit Zaidman (2003) and Goldhill (2001), 156.

23 See Pritchett's work on ancient Greek topography, for a focus on Pausanias and the ancient routes through Greece: Pritchett (1965).

creation of the atlas is the fact that Pausanias visited and described several abandoned settlements through Greece. Some of these were indeed of some historical and cultural importance, while others were definitely more obscure. This represents a measurable aspect of the *Periegesis*, consisting in a precise phenomenon with a discernible spatial distribution. It also proved to be a promising example of isolating and evaluating a set of features from a classical text, from which a series of considerations could be drawn regarding their location.

The state of factual abandonment[24] of several settlements becomes all the more remarkable when one takes into consideration the fact that relevant sites such as Mycenae (2,15,4–2,16,7), Tiryns (2,25,8), or Nauplia (2,38,2–2,38,3) were deserted by the age in which Pausanias wrote. The state of the ancient Greek towns in Pausanias' time may be tied to several different explanations. In many instances, it could be tempting to simply connect this phenomenon to the consequences of the Roman occupation. In describing the city of Megalopolis for example, Pausanias does explicitly refer to population movements in Greece together with the 'disaster' (συμφορά) of Roman rule. The author proceeds thus to draw a parallel between the case of the Arcadian city, founded in the aftermath of the Battle of Leuctra in 371 BC, and the population shifts that characterized the Roman dominion.[25] While Pausanias' attitude towards the Roman intervention in Greece is often critical,[26] however, it cannot be considered *wholly* negative.[27] Nor can it be taken into account for every case of settlement depopulation, as the trend of a general depopulation of Greece, while certainly ongoing during the Roman domination, had been underway at least since the Hellenistic period.[28]

Pausanias' account of Megalopolis (8,27,1–8) is indeed relevant in this respect, as an example of depopulation independent of Roman rule, and helps to illustrate the circumstances of a population decline. The foundation of the city, established in opposition to the hegemony of Sparta, required in fact the settlement of a first wave of inhabitants, which was carried out by transferring the population of

24 It is worth of note, nevertheless, that there are many instances in which the depopulation of a specific settlement does not imply a state of complete abandonment. The remains of a deserted town could still be used by the neighboring communities, especially for very specific cultural purposes; see more *infra* pp. 66–68.

25 […] ἡ δὲ Μεγάλη πόλις νεωτάτη πόλεών ἐστιν οὐ τῶν Ἀρκαδικῶν μόνον ἀλλὰ καὶ τῶν ἐν Ἕλλησι, πλὴν ὅσων κατὰ συμφορὰν ἀρχῆς τῆς Ῥωμαίων μεταβεβήκασιν οἰκήτορες (8,27,1); the Greek texts from the *Periegesis* and the respective translations in this article are drawn from the editions of the *Perseus Digital library*, edited by W. H. S. Jones and H. A. Ormerod; see Jones et al. (1918).

26 Emblematic in this sense is the case of Corinth, destroyed by Lucius Mummius during the Achaean War in 146 BC. Pausanias notes that, during his time, there were no longer any native inhabitants of Corinth, but only descendants of the colonists relocated to the area by the Romans following the city's 'refoundation' by Caesar (2,1,2).

27 The matter of Pausanias' attitude towards the Roman rule is complex, and still a subject of debate; see Hutton (2008) and Moggi (2002). In several cases, Pausanias shows open criticism towards the interference of *other* non-greek entities, as in the case of the Macedonian occupation of the Piraeus (1,26,3); see Bultrighini (1984) and Beschi / Musti (2013). See Bearzot (1992) for Pausanias' negative judgment of the role played by Macedonia in Greek history.

28 Mainland Greece represented in fact a noticeable exception in a broader context of population growth that involved the whole Mediterranean area during the Hellenistic period. The phenomenon, observable at least from the 3rd century BC, manifested itself with a decline in rural population; this trend is partly attributable to a shift of population towards larger settlements, as well as to the emergence of large land estates in the rural areas; see Alcock (1994), 178–179. The late second and third centuries AD, on the other hand, were characterized by a demographic drop in the Greco-Roman world, caused by pandemics such as the Antonine Plague (165–180 AD); see Scheidel (2004), 747. It is furthermore necessary to take into account the changing social conditions which had occurred in Greece by the time the *Periegesis* was written. By the 2nd century AD, many inhabitants of the Greek countryside (in particular, landowners), had begun to move to larger towns, to the detriment of small settlements. See for example Tomlinson (2014), 28 for a survey of the case of Argolis, and Rizakis (2014) for a survey on the changes in the economic relations between the ancient Greek city and the countryside during the early Roman empire.

the nearby towns, which were consequently deserted.[29] Pausanias' antiquarian attitudes, nevertheless, led the author to describe and locate the remains of these settlements in their respective geographical context.

This digital atlas may be used to evaluate such occurrences, and was conceived with the aim of overcoming the limitations and issues inherent in similar previous projects. Furthermore, it means to provide a replicable model in the geographic approach to the interpretation of classical texts, based on a taxonomy that may ensure not only the consultation of a complete database but also an easy and rapid collection of data. The scope of this paper is thus to outline a model of interpretive methodology based on the use of a *GIS* software, by illustrating how the creation of a georeferenced database based on the analysis of a classical text may allow to draw several conclusions about the author's goals and interests in relation to the area under examination.

First, I will elaborate on the methods applied in the creation and implementation of the digital atlas. I will thus take into account the methodological choices behind the definition of the taxonomy of the map, particularly in regard to the issues related to the creation of an updated list of toponyms for the identification of the sites. I will also list the specific features of these locations and how they have been annotated and categorized. The dissemination of abandoned settlements through second century AD Roman Greece will be successively examined, with a complete list of all these cases.[30] Following this, I will elaborate on a few considerations drawn from this list through the georeferencing of the collected data. I will therefore list all the causes of abandonment of the sites under examination, together with the regional distribution of the different cases. More importantly from the literary point of view, it is worth noticing that the area of an abandoned settlement may be indicated by the author by using different, but significant terms. Therefore, I will also scrutinize all the distinct expressions used by Pausanias to connotate a specific site, while evaluating their distribution in the described Greek regions.Both the terms used to indicate the type of settlement and those used to describe the specific state of ruin of these areas have been taken into account.

Since Pausanias was mainly guided by his educated interests, with a very strong emphasis on the antiquity and cultural relevance of the sites he visited, I shall also examine all the occurrences of elements of interest that could have led the author to a deserted settlement. I will thus proceed to provide several lists, detailing all the cases in which a particular area still hosted religious activities or was otherwise relevant in the eyes of Pausanias from a cultural or even mythological point of view.

Methods

Text analysis and taxonomy

At the very beginning in the creation of this atlas, by then the subject of my PhD thesis, I chose to use a critical edition of Pausanias' *Periegesis* in order to grant the utmost precision in the identification and location of the sites described in the text. To this end, the choice fell on the updated critical edition

29 See Drakopoulos (1997) for the case of Asea. The village was one of the settlements which suffered depopulation in the founding of Megalopolis. Other significant cases of depopulation and population movement in Pausanias are presented in the instance of the foundation of Nicopolis (7,18,8); see instead Purcell (1987) and Isager (2009) regarding the migrations of the Aetolians (8,24,11; 10,38,4). On the topic in general, also refer to Mackil (2004).

30 See also Alcock (1996), 145–49. This work certainly does not aim to be the first example of a study on the subject itself, but rather to present a case study on the use of digital mapping for an efficient representation, cataloging, and further stages of analysis of the material provided by Pausanias' *Periegesis*.

of the volumes published by the Fondazione Valla.[31] As for the *Perseus Digital Library*, possibly the most well-known online thesaurus,[32] it presented a series of issues that would have made more difficult the creation of the map. Several of the comments from the online library are in fact drawn from sources in many respects outdated, such as the W. H. S. Jones and H. A. Ormerod edition of the *Periegesis* (1918) or William Smith's Dictionary of Greek and Roman Geography (1854); James Frazer's comments from his 1898 edition of the *Periegesis* are also referenced.[33]

An indiscriminate use of these comments, even an indirect one, would have meant to adopt many identifications which are by now outdated, and which would have proved particularly problematic as they do not take into account, for example, the many changes in Greek toponymy through the 20th century.[34] The texts edited by the Fondazione Valla, on the other hand, offered a corpus of data vastly updated in regard to both the archaeological research and the changes in toponymy, together with a detailed critical apparatus and a rich bibliography which proved an excellent asset in the implementation of the *GIS* atlas. Nevertheless, a closer scrutiny proved the presence of a few outdated toponyms and identifications even in these texts. Indeed, several identifications in the Fondazione Valla volumes too were drawn, ultimately, from outdated sources such as Frazer's work.[35] In order to overcome this issue and try to correctly place the accounted locations on their proper position, I decided to also refer to the identifications provided by the *PANDEKTIS* website.[36] This platform, developed by the Institute of Neohellenic Research, the Institute of Byzantine Research and the Institute of Greek and Roman Antiquity, records the various changes in the toponymy of modern Greece and has enormously facilitated the identification of the current position of the sites mentioned in the *Periegesis*. A careful consultation of the classical text, following the aforementioned measures, eventually led to the recognition and methodical insertion[37] of the sites into the digital map.

Shapefiles definition

Most of the data adopted for my digital map are currently in Italian. An English translation is underway, to ensure the accessibility of the collected records to as many users as possible, both scholars and non-experts. The collection of the sites from the *Periegesis* soon led to a vast amount of information, which required a proper categorization in order to organically explore the locations described by Pausanias as a complex of distinct elements. By doing so, both the numbers and distribution of the spatial data could be evaluated to draw useful conclusions for the research. In light of the interpretation of Pausanias' text, a precise taxonomy was defined, and seven distinct categories were chosen for the purpose of cataloging all the elements described by the author in the ten books of the *Periegesis*. Each

31 The texts cover all the books of the *Periegesis*: Attica (2013), Corinth and Argolis (2008), Laconia (2008), Messenia (2010), Elis and Olympia (in two volumes, 2007 and 2013 respectively), Achaea (2008), Arcadia (2007), Boeotia (2012), and Phocis and Delphi (2017).

32 http://www.perseus.tufts.edu/hopper/ (last access 11.11.2024).

33 See Jones et al. (1918), Smith (1854) and Frazer (1898).

34 Indeed, the process continued through all of the 20th century, with repercussions as late as into the 21st century. On the matter of the Hellenization of modern Greek toponyms, see Kyramargiou et al. (2020); Dimitropoulos et al. (2020). For the name changes as part of more recent administrative changes in Greece, see Akrivopoulou et al. (2012).

35 The example of the hydrography names of the Achaea region as reported in the critical editions may help to better illustrate the issues related to this specific topic. In the cases of the rivers Phoenix, Miganita, and Buraicus, the ancient names were indeed recovered and are currently in use. However, within the critical comments from the Fondazione Valla volumes, they are still referred to as Salmeniko, Gaidaropniktes and Ladopotamou respectively; see Moggi / Osanna (2008), 317; 333.

36 https://pandektis.ekt.gr/ (last access 11.11.2024).

37 In the case of my project, the insertion of the shapefiles corresponding to the sites from the *Periegesis* was carried on manually.

typology was represented on the digital map through the insertion of specific layers[38], associated to distinct shapefiles[39] in order to recreate the layout of the ancient landscape. The sdefined layers are:

- Cities: a category comprehensive of all the settlements still inhabited during Pausanias' age.
- Forests: this layer specifically includes sacred groves (ἄλση), often associated in turn with monuments or sanctuaries.
- Locations: a layer designed in order to include more generic spatial features within Pausanias' *Periegesis*. Therefore, the layer includes sites of historical battles, unspecified locations, etc. The abandoned settlements identified during the creation of the digital map were initially ascribed to this specific typology.
- Monuments: all the isolated monuments, either commemorative or cult-related, encountered by Pausanias, which stood *outside* of the broader context of ancient settlements or sanctuaries.
- Mountains: the mountains described within the text, which often hosted additional elements.
- Rivers: the modern course of Greek hydrography, rather than a reconstruction of the ancient river routes, was recreated for the digital map, as Pausanias often describes several smaller watercourses for which it was not possible to find evidence of their state in antiquity.
- Sanctuaries: a complete list of all the accounted isolated sanctuaries, not including those part of a broader context, such as shrines located on the top of mountains or in the frame of ancient settlements.

Additionally, it was necessary to define what features would have been examined for the single sites associated to these layers. A *GIS* software provides several ways to view such data; in particular, by listing them inside an attribute table. A further step required thus the definition of the layout of such a table, consisting in:

- Name: the toponym of a specific site.
- Source: the proper reference to Pausanias' text.
- Data: a brief description of the selected shapefile. In accordance with the literary approach of this digital map to the *Periegesis*, the description is represented by a paraphrase of Pausanias' testimony rather than an objective comment, which is instead delegated to other fields (see below).
- Era: a chronological contextualization has been provided to better define the distribution of sites of interest through the centuries.
- Typology: the type of layer to which each shapefile is related is indicated within this field.
- Location: the identification of a specific site within the context of modern Greece. In those cases where it was not possible to determine the area of a particular element, this very fea-

38 In the context of a digital atlas, a layer represents a database of geometric elements disseminated through the map. It may be considered a macro-category of associated elements from the map itself. As the name implies, it visually appears as a layer of contents; several layers may be swapped, activated or deactivated for a better comprehension of a digital atlas.

39 A shapefile is a vector format for geographic data, ideally corresponding to the elements inserted in the map itself.

ture was labeled as 'unidentified', but nevertheless inserted inside the map according to the indications provided by Pausanias. In all the cases where there are instead multiple possible identifications, the most acceptable one is applied and recorded first in a list of proposals (a choice reflected by the placement of the shapefile on the map). This approach aims to provide a comprehensive overview of the state of the identifications of the *Periegesis* sites, thus overcoming the limitations of other digital atlases, which in similar cases usually present a *single* identification. Every proposed identification is supported by a bibliographic reference.

All of these features may be viewed by selecting the attribute table option. A more effective and immediate method, however, would be to directly select the shapefiles on the map. This action will open a pop-up, with all the site features accurately listed. Every dataset created for the representation of these elements is a *point* layer. The only exception is represented, of course, by the rivers, for which it was necessary to adopt a *polyline* type layer in order to recreate their course. While the procedure was originally carried out by using the open source *QGIS*[40] mapping system, it was eventually decided to use a different software. The project required, in fact, an application that would allow a faster completion of the *Periegesis* map, not to mention the association of complex data with the localized sites. It was also necessary to use a program capable of uploading the digital atlas online in order to grant a further implementation by providing more advanced tools for the search of contents, while eventually ensuring its dissemination. To this end, it was decided to transfer the layers of the already created map within *ArcGIS Pro*[41]. This proprietary software, provided by Esri[42], has indeed been a reliable standard for the creation, statistical analysis, and sharing of complex maps for a long time.

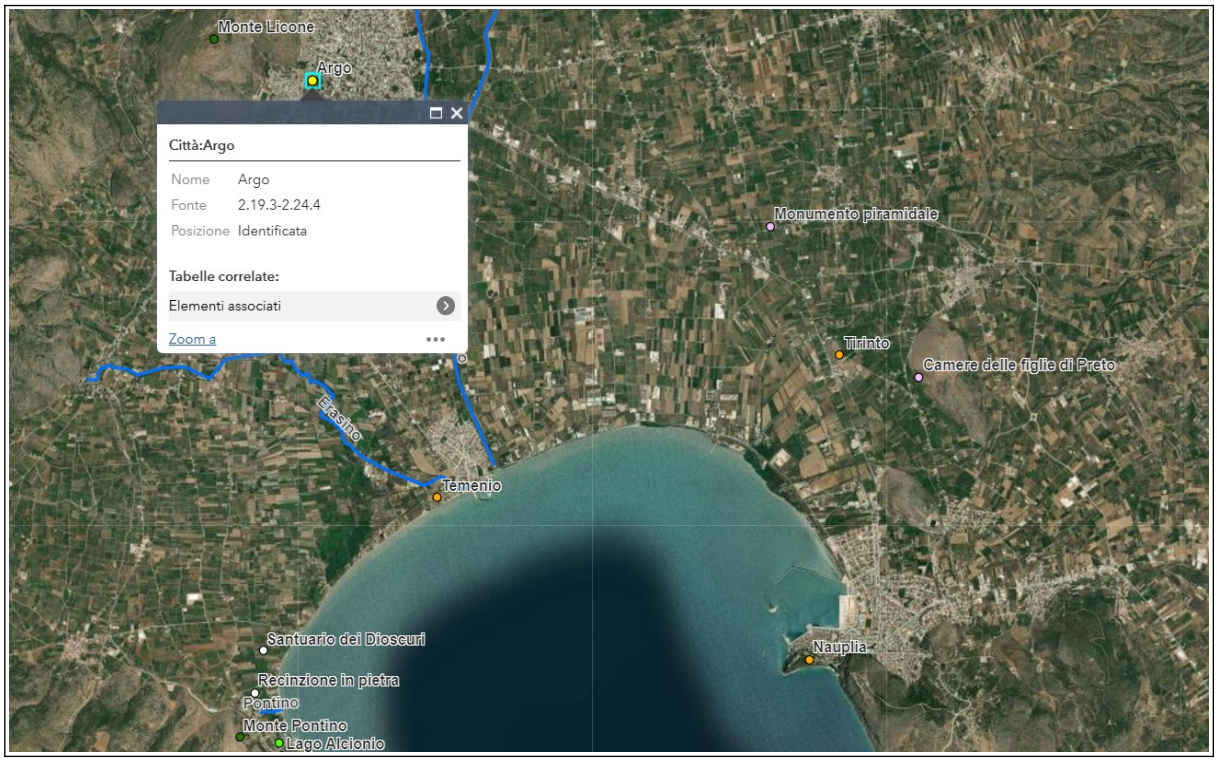


Fig. 1: A screenshot from the digital atlas. It is possible to view several shapefiles arranged to recreate a precise context from Pausanias' *Periegesis*. By clicking on a specific shapefile (in this case corresponding to the city of Argos), it is possible to view a list of related data. The features visible in relation to the 'city' layer are, from top to bottom, layer name, shapefile name, source, and state of identification, respectively. From the bottom of the pop-up it is possible to select a list of related elements; see *infra* pp. 59–60.

40 https://www.qgis.org/it/site/ (last access 11.11.2024); see Graser et al. (2017) for a survey on the *QGIS* software; Geddes / Gregory (2014) for a focus on the applications of *GIS* software to the historical research.

41 https://www.arcgis.com/index.html (last access 11.11.2024).

42 https://www.esri.com/en-us/home (last access 11.11.2024).

While the sites described by the *Periegesis* had been successfully digitized inside the map, it was necessary, as aforementioned, to represent also all the elements contained *within* them, or otherwise related to them. This is the case, for example, for all the single buildings, monuments or generic features of a described settlement. Another case are the works of art that were once displayed within specific contexts such as ancient sanctuaries (whose detailed description represents a well-known trait of Pausanias' exposition[43]), or the single shrines located inside forests or on the top of mountains. To this end, a series of Excel tables were compiled with all the data related to these elements. Again, just like in the case of the layers from the *GIS* map, the data were associated with a series of specific features.

- Shapefile name: the shapefile of the site to which a certain related element is associated.
- Name: the name of the analyzed element itself.
- Source: the reference from Pausanias' text.
- Location: the precise position of a feature inside the broader context of an ancient site.
- Status: the feature identification, either certain, uncertain or not possible.
- Era: in this case, too, a chronological contextualization was provided for a specific element.
- Proposed identification(s): again, each identification is tied to a bibliographic reference.
- Description: a succinct description, following Pausanias' account.
- Typology: the exact type of analyzed element, specified in order to facilitate the search through the gathered data.

These tables were eventually loaded into the *ArcGIS Pro* map as *CSV* files, and exported to the project's geodatabase. A subsequent phase consisted thus in the association of the records embedded inside these tables to the respective geographic locations, already represented through the shapefiles. To this end, a *Create Relationship Class* spatial operation was launched between the tables and the single layers. In order to relate these records, a common field, or *key*, was selected; in this case, the key consisted in the name of a site. This information is in fact present in both the map layers and the *CSV* table. The specific relationship selected to link these records was the 'one to many' cardinality.

43 See Arafat (1992), 387. Typical in Pausanias are accurate and critical descriptions of monuments and works of art, which often lead him to dwell on the style, author, or even materials of the described elements; see Pretzler (2013), 109.

Argo	Statue di Zeus e Artemide	2.19.7	Complesso di Apollo Lykios	Non identificato	Non specificata	<Null>	Xoana lignei su colonna di Zeus e Artemide.	Statua
Argo	Tombe di Lino figlio di Apollo...	2.19.8	Complesso di Apollo Lykios	Non identificato	Non specificata	<Null>	Duplice tomba legata con il tempio di Apollo Lykios.	Testimonianza mitologica
Argo	Simulacro di Apollo Agyieus	2.19.8	Complesso di Apollo Lykios	Non identificato	Non specificata	<Null>	Simulacro aniconico di Apollo, che segnalava forse l...	Statua
Argo	Altare di Zeus Hyetios	2.20.1	Complesso di Apollo Lykios	Non identificato	Non specificata	<Null>	Altare presso il quale gli alleati di Polinice avrebber...	Altare
Argo	Statua di Creuga	2.20.1	Complesso di Apollo Lykios	Non identificato	Non specificata	<Null>	Statua del pugile Creuga di Epidamno, presso il te...	Statua
Argo	Trofeo della vittoria sui Corinzi	2.20.1	Complesso di Apollo Lykios	Non identificato	Non specificata	<Null>	Monumento possibilmente dedicato in occasione d...	Monumento commemorativo
Argo	Statua di Zeus Meilichios	2.20.1	Complesso di Apollo Lykios	Non identificato	Non specificata	<Null>	Statua seduta in marmo bianco, opera di uno dei d...	Statua
Argo	Rilievo di Cleobi e Bitone	2.20.2	Complesso di Apollo Lykios	Non identificato	Non specificata	<Null>	Sculture a rilievo di Cleobi e Bitone nell'atto di trasc...	Statua
Argo	Santuario di Zeus Nemeo	2.20.3	Agorà	Incerto	Classica	Fondazione sul...	Tempio con una statua di bronzo, opera di Lisippo.	Edificio sacro
Argo	Tomba di Foroneo	2.20.3	Agorà	Non identificato	Non specificata	<Null>	Monumento nei pressi del tempio di Zeus Nemeo.	Testimonianza mitologica
Argo	Tempio di Tyche	2.20.3	Agorà	Non identificato	Non specificata	<Null>	Tempio in cui Palamede avrebbe dedicato i dadi ch...	Edificio sacro
Argo	Tomba della Menade Corea	2.20.4	Agorà	Non identificato	Non specificata	<Null>	Tomba eroica in ricordo della spedizione di Dioniso...	Testimonianza mitologica
Argo	Santuario delle Ore	2.20.5	Presso l'Agorà	Non identificato	Non specificata	<Null>	Santuario dedicato alle Ore nel ruolo di nutrici di Er...	Edificio sacro
Argo	Cenotafio degli Argivi	2.20.6	Agorà	Non identificato	Non specificata	<Null>	Monumento dedicato agli Argivi che si ritieneva fos...	Testimonianza mitologica
Argo	Tomba di Danao	2.20.6	Agorà	Non identificato	Non specificata	<Null>	Tomba eroica che, secondo Strabone (VIII 371), ave...	Testimonianza mitologica
Argo	Santuario di Zeus Soter	2.20.6	Agorà	Non identificato	Non specificata	<Null>	Santuario verosimilmente sull'agorà (Musti e Torell...	Edificio sacro
Argo	Santuario per le celebrazioni d...	2.20.6	Presso l'Agorà	Non identificato	Non specificata	<Null>	Una delle rarissime installazioni note di Adone, giu...	Edificio sacro
Argo	Santuario del Cefiso	2.20.6	Presso l'Agorà	Incerto	Non specificata	Terrazza ad est d...	Santuario dedicato al fiume Cefiso, di una tipologia...	Edificio sacro
Argo	Testa di Medusa in pietra	2.20.7	Presso l'Agorà	Non identificato	Non specificata	<Null>	Monumento in pietra attribuito ai Ciclopi.	Testimonianza mitologica

Fig. 2: Screenshot of a table with related features from the digital atlas. From left to right it is possible to view the specific characteristics associated with each element; in this case, the name of the shapefile, name of the element, source, exact position, identification status, period, proposed identification, description and type of the element respectively.

Once completed, the operation successfully tied the elements from the table to their proper position. By clicking on a specific shapefile on the map, it is thus possible not only to evaluate all the characteristics of the site itself listed inside a pop-up, but also to open a second pop-up where all the related elements are listed, each one in turn associated to its features categorized inside a proper table.

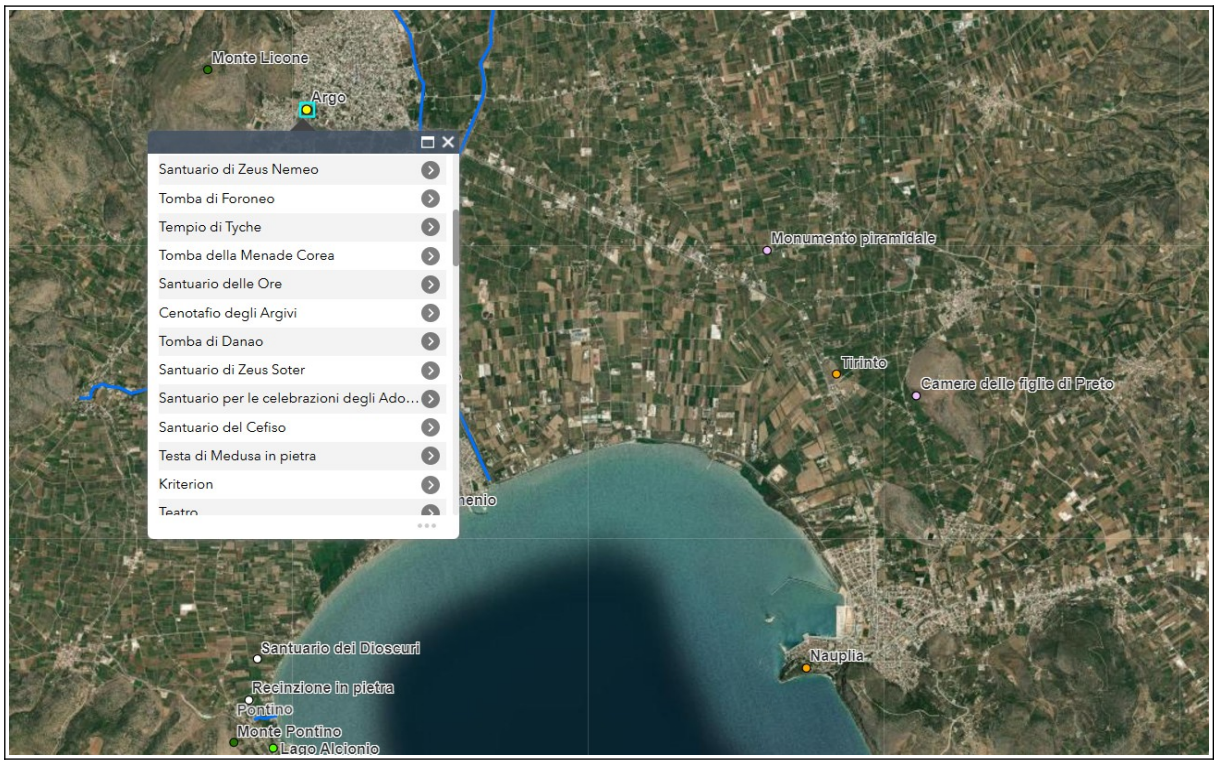


Fig. 3: A screenshot from the digital atlas, focused on a specific area. By expanding the related elements list it is possible to view all the elements associated with a single site. In this case, the picture shows a series of elements associated to the city of Argos, listed inside a pop-up.

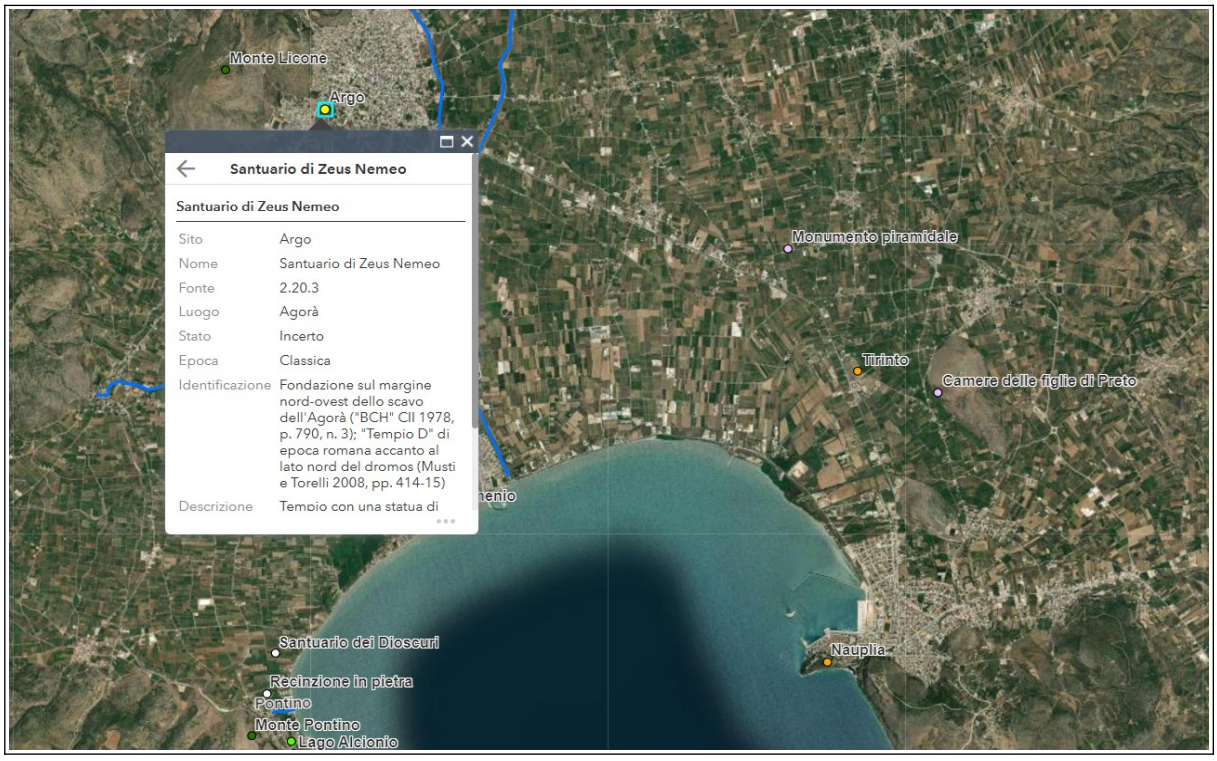


Fig. 4: A screenshot of the expanded table with related elements; by clicking on a specific feature from the list, it is possible to view all the data associated to an 'accessory' element: namely, the temple of Nemean Zeus in Argos (2,20,3). In this case, from top to bottom, the shapefile of which the element is part of, the name of the element, the source, the position, the state of identification, the time period, a few hypotheses, and a brief description.

This feature of the atlas aims to provide a further level of precision to the representation of an ancient landscape often lacking in other thesauri and digital maps.[44] Once the map was completed, the collected data amounted to 2791 elements from the *Periegesis* (868 shapefiles and 1923 related elements). The next step involved its uploading to an online platform. To this end, the use of the *ArcGIS* software allows to upload a digital map to *ArcGIS Online*, a cloud platform provided by ESRI for content sharing. Additionally, *ArcGIS Online* allows to embed a created map inside web apps which facilitate the dissemination and consultation of complex geodatabases. In this case, the *ArcGIS Web App Builder* function was chosen to create a web app that would convey the digital map containing the data from the *Periegesis*. The *ArcGIS Web App Builder* also allows to include within a web app several distinct *widgets*; commands that enable different kinds of intervention on the digital map and its contents. In this case, the 'search' widget was inserted, which allows possible users to type specific keywords, thus performing a targeted and in-depth search among the digitized contents of the *Periegesis* within the map. The analysis and digital representation of the case study, composed of all the elements accounted for in the text, eventually allowed the creation of a further, preliminary map focused on the abandoned settlements encountered by Pausanias. This in turn was eventually used as the basis for the creation of five new maps detailing specific features related to this aspect of the *Periegesis*, which provided the means to investigate further case studies.

In defining the elements within abandoned towns, it was decided to consider only those features that Pausanias recognizes as inherent to the context of the settlements. The author is usually keen in separating certain monuments, sanctuaries, etc., from the settlements, inhabited or not, that stood nearby. Therefore, those elements that Pausanias places outside the boundaries of settlements or in their imme-

44 As discussed above, p. 52.

diate vicinity were not taken into consideration as part of these areas, and have been recreated on the map through seperate shapefiles.[45]

A similar clarification needs to be made in regard to the method used to catalog the accessory features associated with the elements described by Pausanias. In the case of cult statues located within deserted settlements, for example, it was decided to consider them as standalone entities unless the author specified otherwise. The statues, altars, works of art, etc., that Pausanias explicitly associates with the sanctuaries inside these settlements, on the other hand, have been linked to these contexts, and are listed in their respective attribute table. Finally, in some cases, multiple designations have been used by the author in describing the same element. In the case of Nonacris, for example, the village is indicated as both πόλισμα (8,17,6) and πόλις (8,27,3–4). It was decided, therefore, to reflect the literary approach of the *GIS* atlas and to indicate sites with multiple denominations by using a distinct category.

Results

Preliminary lists

The creation of the Pausanias' *Periegesis* georeferenced database and the implementation of further maps represented an opportunity to isolate and evaluate all those instances where the author mentions an abandoned settlement. As already stated, the element of novelty of this case study consists not only in the fact that the data may be immediately visualized and queried, but also in the fact that several more complex analyses of the text accounts and their own correlations may be easily and rapidly carried out. This in turn allowed to elaborate a few considerations based on the number, type and distribution of some of the most prominent features concerning the deserted settlements described by Pausanias.

45 The constructions identified by Pausanias as the 'Chambers of the daughters of Proetus', for example, lie immediately near Tiryns (2,25,9). However, Pausanias differentiate them from the ruins of the ancient city itself. Thus, a separate shapefile was used to describe the 'Chambers of the daughters of Proetus' as a distinct element not included in the description of Tiryns itself, nor in the analysis of its related features.

Fig. 5: View of the 83 deserted settlements in Pausanias' *Periegesis*. This map consists in a single layer, obtained from the preliminary *Periegesis* map. These data, in turn, were used to create further layers and maps detailing the features inherent to this particular case study from Pausanias' text. Note the absence in the chart of any element located in the area of Attica; see infra p. 63.

Pausanias' fascination for ruins and deserted (but often culturally relevant) sites is of course tied to the author's personal interests and background, which naturally led him to seek out elements characterized by an intrinsic archaism and with a link with the local tradition (or both[46]). His narrative itineraries from one main settlement to the next one often lack a logical spatial linearity, providing instead an occasion for detours in search of local curiosities. These explorations either relied on the accounts of guides, local hypotheses, or more educated sources drawn from the literary tradition, with a strong preference for Homer.[47] It is also possible that Pausanias could at times deliberately choose areas on which he was not well informed, while using the main settlements as a starting point for his explorations.[48] A quick survey of the map of the deserted sites allows to appreciate a series of features regarding the location of these areas. It is immediately evident that the highest concentration of these elements lies in the area of Arcadia, with a cluster of shapefiles related to the phenomenon lying immediately around the city of Megalopolis. The presence of so many (30) abandoned sites disseminated

46 See Bruit Zaidman ed. (2003) and Goldhill (2001), 156 for Pausanias' search for archaicism. This tendency by the author does not represent an isolated case in the broader picture of 2nd century AD literature; on the contrary, it is fully compatible with the style and canons of the Second Sophistic movement; see Anderson (2005), and Alcock et al. (2003), 63–67, for a comparison between Pausanias and Longinus.

47 For the use of Homer in Pausanias, see Bacher (1919).

48 See in this regard Hutton (2005), 83–174, who theorizes the existence of a basic structure for Pausanias' travels. The author postulates an organization of the explorations in the *Periegesis* according to a 'radial plan', with a central point from which various itineraries branched out towards destinations that Pausanias considered worthy of attention; see also Pretzler (2004), 203.

through Arcadia is compatible both with the history of the region, which had suffered a marked depopulation, and with the aims of Pausanias. The area was in fact related to a rich corpus of archaic traditions, which naturally piqued the interest of the author.[49]

As for Megalopolis in particular, with the presence of 16 abandoned settlements within a radius of roughly 14 kilometers from the city, this situation is confirmed by the local history, too.[50] The second region by number of deserted settlements is instead Laconia, with 13 abandoned towns, mostly disseminated along the southern coast, in an area partially corresponding to the territory of the Eleutherolakones.[51] Boeotia instead presents 10 elements, concentrated in particular in the area between the Lake Copais and the Asopus river. Next are the areas of Argolis and Corinthia, characterized by a rich historical and cultural past which certainly interested Pausanias,[52] Achaea and Elis (8 elements in each case), Messenia and Phocis (both with 3 elements). Curiously, there is no evidence of abandoned settlements in Attica.[53] This is explainable in part by the Roman policies, which had privileged the area, especially during the Antonine rule.[54] Furthermore, Attica represented in many aspects an area not 'archaic' enough for the tastes of Pausanias.[55]

Case studies

The use of a *GIS* software also made it possible to highlight and distinguish the various causes for a settlement abandonment (or destruction). In most cases (48 elements), the reasons for the decline of a specific settlement are not reported. Naturally, this does not mean that the reasons are utterly unknown to the research, but simply that Pausanias was not able (or did not deem necessary) to elaborate on the specific circumstances for each case. The abandonment of a total of 17 elements, on the other hand, was attributed by Pausanias to a predetermined depopulation for the purpose of populating a new city. The phenomenon, once again, is particularly evident in Arcadia, with a marked concentration around the area of Megalopolis. This is not surprising since, as mentioned earlier, the establishment of Megalopolis itself was an event that required the depopulation of nearby villages.

49 See Berchet (1986).

50 See Drakopoulos (1997).

51 See Chrimes (1949), 435–6.

52 In the case of Argolis in particular, I must point out that the city of Mases was not considered part of these lists, as it was not, strictly speaking, an 'abandoned' settlement as much as a town repurposed as Hermione's seaport (2,36,2); the same goes for the city of Samicum (5,5,3), which was a fortress by the Roman age (see Strab. 8,3,19).

53 Tab. 1: List of the abandoned settlements from the *Periegesis*; see in the data repository: https://doi.org/10.11588/data/FO1QAT (last access 16.10.2024).

54 The area was notoriously interested by the philhellenic policies of Hadrian, which had by then recently created the league of the *Panhellenion*; see Anderson (2005), 3.

55 On the other hand, the Roman policies in Attica can be considered as an additional element of 'modernity' and foreign interference, which might have made the region less appealing to the eyes of Pausanias; see Beschi / Musti (2013), XLVIII–LI.

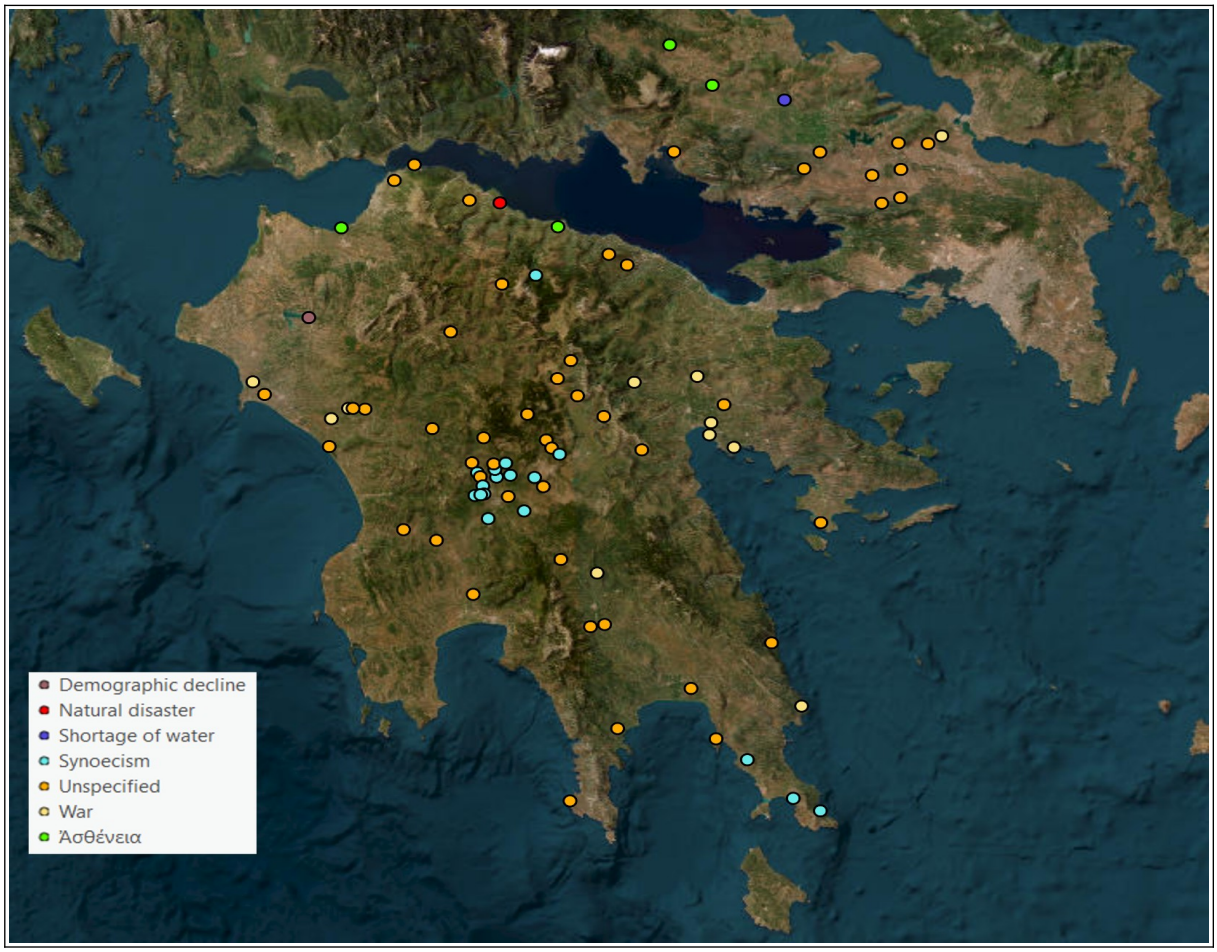


Fig. 6: Map of the deserted settlements from the *Periegesis*, differentiated by color based on the causes of their abandonment. It is possible to view, in particular, an evident cluster of settlements (in cyan), in the area of Megalopolis in Arcadia, that were abandoned following a synoecism.

Of the 12 settlements destroyed and never inhabited again, the destruction is never attributed by Pausanias to the interference of the Roman presence (in the famous case of Corinth, the depopulated city was eventually reoccupied). The responsibility is instead attributed to the Argives (5 elements)[56], the Eleans (3 elements)[57], the Thebans, the Achaeans and the Spartans (1 element in each case)[58], and even to natural phenomena (1 element)[59]. The abandonment of 4 settlements is also curiously attributed to 'ἀσθένεια', which may be translated as 'weakness'.[60] Finally, one settlement was apparently abandoned due to lack of water, and another due to a marked demographic decline.[61]

56 The Argives were responsible for the depopulation of Mycenae (2,16,5), Nauplia (4,24,4), Orneae (2,25,6), Tiryns (2,25,8), and Asine (2,36,4–5). In cases where the inhabitants of a city were forcibly relocated to another settlement as a result of a conflict (such as in the case of Tiryns, 2,25,8) and not as part of a 'voluntary' synoecism (as in the case of Megalopolis), the event has been considered a case of depopulation due to war.

57 The Eleans, on the other hand, depopulated Scillus (5,6,4), Pisa (6,22,4), and Dyspontium (6,22,4).

58 Mycalessus was depopulated by the Thebans (1,23,3), Sellasia by the Achaeans (3,10,7), and Zarax by the Spartans (3,24,1).

59 The city of Helice, in Achaea, was indeed destroyed by an earthquake in 373 BC, and no remains were visible in Pausanias' time (7,24,6).

60 See Alcock (1996), 148.

61 Tab. 2: List of the causes of abandonment of the *Periegesis'* deserted settlements; see in the data repository: https://doi.org/10.11588/data/FO1QAT (last access 16.10.2024).

In 61 cases, Pausanias refers to the abandoned settlements as πόλεις. The slightly more generic term πόλισμα, on the other hand, is used in 13 cases. There are, moreover, only 3 occurrences for κώμη, and 1 occurrence for the terms δῆμος and περίοικος respectively.[62]

As already stated, the list is inclusive of the cases in which multiple terms (such as πόλις and πόλισμα) have been used on different occasions in referring to a single settlement. In 12 cases Pausanias simply does not refer with a specific term to the settlements he visited.

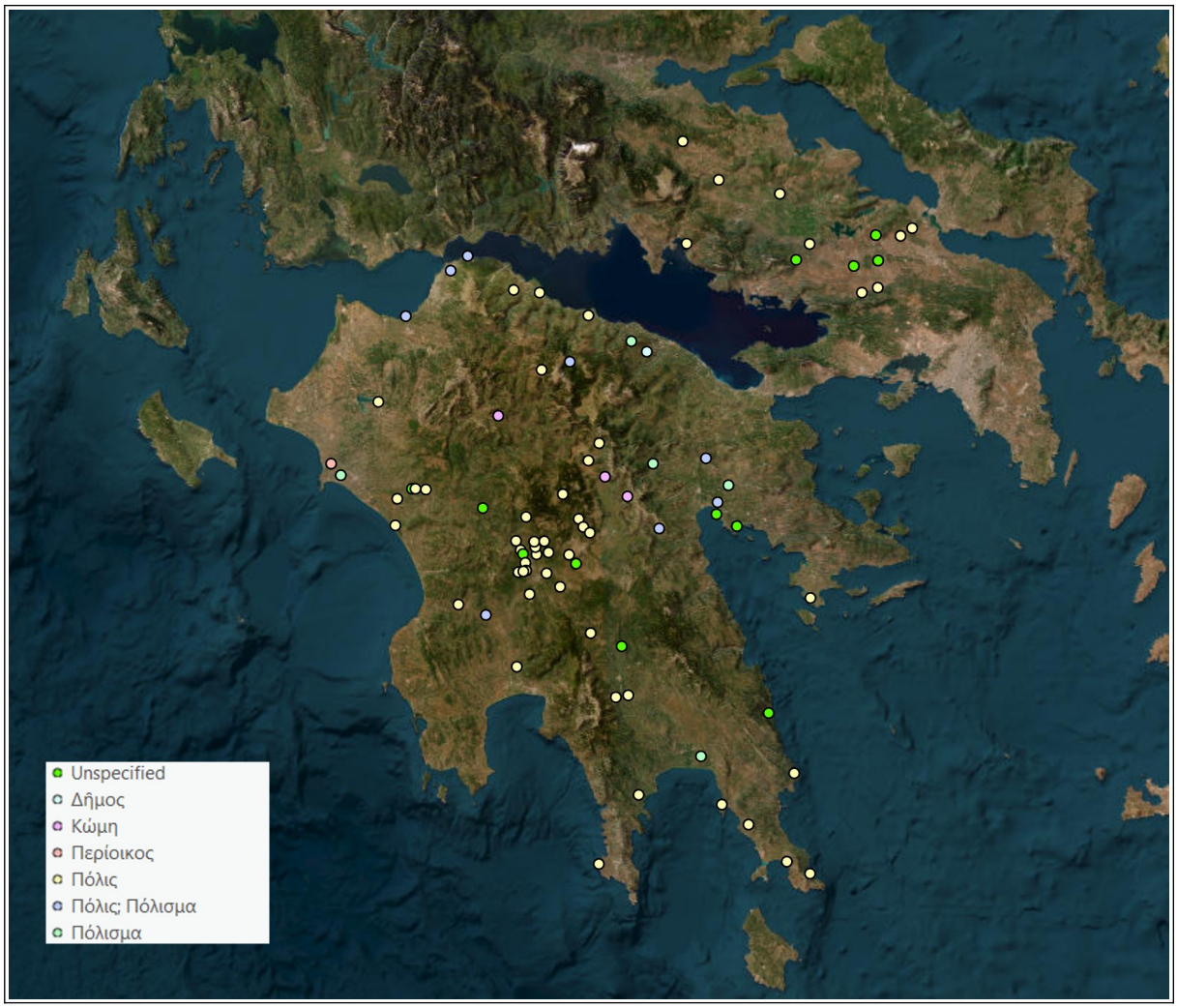


Fig. 7: Map of the deserted towns from the *Periegesis*, differentiated by color based on the term used by Pausanias in referring to the type of settlement. While the use of the term πόλις is evidently prominent and essentially equally distributed throughout all Greece, it is possible to appreciate the use of alternative terms for the area of Argolis.

62 Tab. 3: List of the denominations of the abandoned settlements from the *Periegesis*; see in the data repository: https://doi.org/10.11588/data/FO1QAT (last access 16.10.2024).

Another map was created to evaluate the distribution of the terms used by Pausanias in referring to the state of ruin of the abandoned settlements.

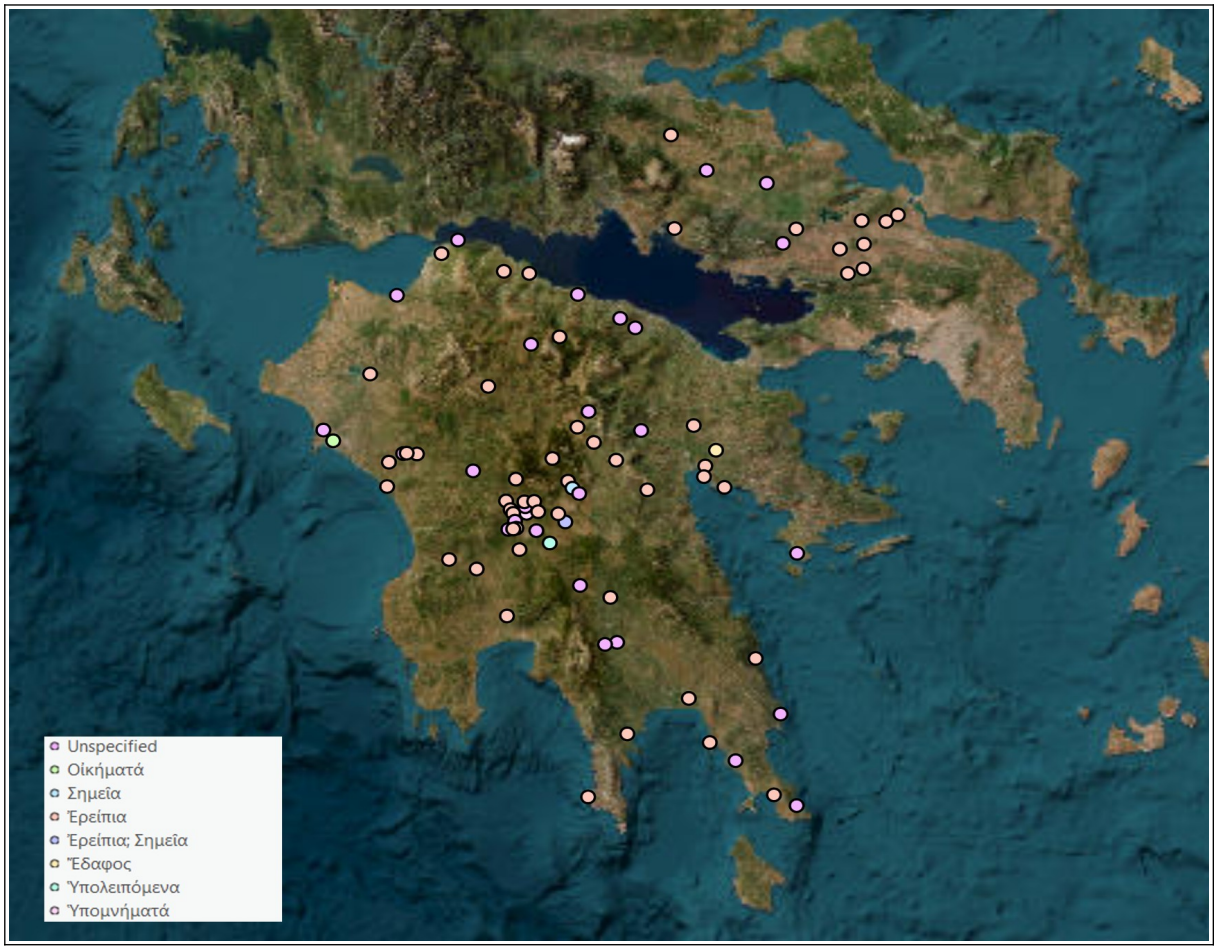


Fig. 8: Map of the abandoned towns from the *Periegesis*, differentiated by the specific term used in referring to their respective state of ruin. The name ἐρείπια is used most frequently, with noticeable clusters visible in Arcadia and southeastern Boeotia.

In referring to 26 settlements, the author does not use any specific denomination when describing deserted cities, while he reports 52 times the term ἐρείπια, which is undoubtedly the term most frequently adopted. The term σημεῖα (2 cases) follows up, with ἔδαφος, οἰκήματα, ὑπολειπόμενα and ὑπομνήματα ensuing (1 occurrence in each case)[63].

A peculiar aspect of the abandoned settlements scattered throughout Greece in the 2nd century AD is the continuity of worship observed in several cases.[64] While describing the city of Potniae, for example, Pausanias clearly specifies that the settlement was in ruins. At the same time, however, he describes in detail the presence of a sacred grove in this site, dedicated to Demeter and Kore (9,8,1). The grove in turn must have certainly been tied to rituals associated with the cult of Demeter held in the area.[65] In several other cases, an entire temple still stands within an otherwise deserted settlement. This is the case with the ruins of Asine, where a sanctuary of Apollo Pythaeus was still visible in the Roman

63 Tab. 4: List of the terms used in referring to the ruined settlements from the *Periegesis*; see in the data repository: https://doi.org/10.11588/data/FO1QAT (last access 16.10.2024).

64 Only in the cases of the abandoned settlements of Bryseae (3,20,3) and Potniae (9,8,1) Pausanias provides a few details in regard to the continuity of worship.

65 διαβεβηκότι δὲ ἤδη τὸν Ἀσωπὸν καὶ τῆς πόλεως δέκα μάλιστα ἀφεστηκότι σταδίους Ποτνιῶν ἐστιν ἐρείπια καὶ ἐν αὐτοῖς ἄλσος Δήμητρος καὶ Κόρης. τὰ δὲ ἀγάλματα ἐν τῷ ποταμῷ τῷ παρὰ τὰς Ποτνιὰς ** τὰς θεὰς ὀνομάζουσιν. ἐν χρόνῳ δὲ εἰρημένῳ δρῶσι καὶ ἄλλα ὁπόσα καθέστηκέ σφισι καὶ ἐς τὰ μέγαρα καλούμενα ἀφιᾶσιν ὗς τῶν νεογνῶν· τοὺς δὲ ὗς τούτους ἐς τὴν ἐπιοῦσαν τοῦ ἔτους ὥραν ἐν Δωδώνῃ φασὶν ἐπὶ ** λόγῳ τῷδε ἄλλος πού τις πεισθήσεται. (9,8,1).

era, spared by the Argives after leveling the rest of the city (2,36,4–5) in 740 BC. It is clear, therefore, that the religious element ensured a continuity of use for various sites, a fact duly noted by Pausanias in his *Periegesis*. The *Periegesis* atlas allows the analysis of all the cases in which Pausanias has indeed noted the presence of temples within otherwise uninhabited areas. In total, the author mentions the presence of 21 temples within as many abandoned settlements.

Fig. 9: Map of the distribution of sanctuaries and worship areas in general in abandoned settlements according to the accounts of the *Periegesis*. Once again, the area of Arcadia presents the highest concentration of elements; Phocis, on the other hand (in the northwest), is evidently not represented by any shapefile on the map.

Once again, the highest number of occurrences is found in the region of Arcadia with 7 elements. Following Arcadia are Laconia (5 elements), Elis, Argolis (3 elements for each region), Boeotia (2 elements) and Messenia (1 element). It is possible to supplement this data with sites that, although not occupied by proper sanctuaries, are nevertheless indicated by Pausanias as sacred. Accordingly, the Poseidium area in Achaea, abandoned over the centuries but still clearly identified by Pausanias as sacred to Poseidon, may be included (7,27,8). Similar cases are represented by the aforementioned area of Potniae and by Bryseae (3,20,3), where sacred groves existed, and where Pausanias observed the survival of religious practices despite the abandonment of the ancient settlements. It is also possible to add the spot where a statue of Hermes still stood in Acacesium (8,36,10), and the sacred cave of Asclepium in Cyphanta (3,24,2). Following these considerations, the total number of abandoned areas with

significant cultic activities recorded by Pausanias rises to a total of 26 elements (8 in Arcadia, 7 in Laconia, 3 in Argolis, Boeotia and Elis respectively, and 1 in both Achaea and Messenia).[66]

It is also interesting to note that most of these sanctuaries are located inland. In my opinion, this can be interpreted as a further confirmation of Pausanias' search for archaism. The phenomenon could be in fact explained by the author's cultural interests, which likely led him to prioritize the exploration of relatively less accessible, inland areas. These locations in fact could have been the ones that, in the author's view, were most likely to preserve ancient monuments and traditions, or at least the less known ones. On the other hand, the data may confirm the fact that Pausanias seemingly didn't rely much on sea routes, preferring instead to travel by land.[67] As for the deities to whom these temples were dedicated, the use of *GIS* software has allowed for the creation of a specific list.[68] A further table may be thus drawn from this list, detailing the number of attestations for each single deity worshiped in the abandoned settlements.[69]

Determining the actual state of preservation of the sites visited by Pausanias is quite more difficult. The author does on occasion specify the cases in which a temple was in ruins or even unfinished. With the exception of the temple of Apollo near Hysiae in Boeotia,[70] the ruins of the temple of Artemis in Oresthasium,[71] and the temple of Athena at Phrixa,[72] however, there is no explicit mention of the possible state of ruin of these structures, thus allowing to presume that they were still standing in the Roman era. The description of cult statues and ongoing religious activities may indeed support this presumption.

As mentioned, there are also several mythological traditions that Pausanias refers to in connection with some of the abandoned settlements. The presence of mythological backgrounds drawn from ancient tradition served indeed as both an incentive for the author to personally visit certain sites and a guide in his itineraries through 2nd century Greece. This was especially true in the case of Homer, who served as a privileged source of information for the author of the *Periegesis*.[73] In this case as well, the *GIS* software was used in order to highlight all the sites in Roman-era Greece that had some connection to a mythological tradition reported by Pausanias, with a total of 63 abandoned settlements. Again it is not surprising to see the prominence of Arcadia, with 25 elements. The region must have represented a privileged target for Pausanias' investigations, as the archaic area of the Peloponnese *par excellence*.[74]

---

66 Tab. 5: List of the abandoned settlements with cultic activities from the *Periegesis*; see in the data repository: https://doi.org/10.11588/data/FO1QAT (last access 16.10.2024).

67 See Pretzler (2004), 203.

68 Tab. 6: List of the religious elements associated with the abandoned settlements; see in the data repository: https://doi.org/10.11588/data/FO1QAT (last access 16.10.2024).

69 Tab. 7: List of the specific deities attested in reference the abandoned settlements in the *Periegesis*; see in the data repository: https://doi.org/10.11588/data/FO1QAT (last access 16.10.2024).

70 γῆς δὲ τῆς Πλαταιίδος ἐν τῷ Κιθαιρῶνι ὀλίγον τῆς εὐθείας ἐκτραπεῖσιν ἐς δεξιὰ Ὑσιῶν καὶ Ἐρυθρῶν ἐρείπιά ἐστι. πόλεις δέ ποτε τῶν Βοιωτῶν ἦσαν, καὶ νῦν ἔτι ἐν τοῖς ἐρειπίοις τῶν Ὑσιῶν ναός ἐστιν Ἀπόλλωνος ἡμίεργος καὶ φρέαρ ἱερόν· πάλαι δὲ ἐκ τοῦ φρέατος κατὰ τὸν Βοιωτῶν λόγον ἐμαντεύοντο πίνοντες (9,2,1).

71 μετὰ δὲ Αἰμονιὰς ἐν δεξιᾷ τῆς ὁδοῦ πόλεώς ἐστιν Ὀρεσθασίου καὶ ἄλλα ὑπολειπόμενα ἐς μνήμην καὶ Ἀρτέμιδος ἱεροῦ κίονες ἔτι: ἐπίκλησις δὲ Ἰέρεια τῇ Ἀρτέμιδί (8,44,2).

72 ἐν ταύτῃ τῇ χώρᾳ λόφος ἐστὶν ἀνήκων ἐς ὀξύ, ἐπὶ δὲ αὐτῷ πόλεως Φρίξας ἐρείπια καὶ Ἀθηνᾶς ἐστιν ἐπίκλησιν Κυδωνίας ναός. οὗτος μὲν οὐ τὰ πάντα ἐστὶ σῶς, βωμὸς δὲ καὶ ἐς ἐμὲ ἔτι […] (6,21,6).

73 19 of the uninhabited towns described by Pausanias are indeed mentioned by Homer: Erythrae, Harma (Il. 2,495), Mycalessus (Il. 2,495), Scolus (Il. 2,495), Glisas, Medeon (Il. 2,500), Onchestus (Il. 2,505), Aspledon, Orchomenus (Il. 2,510), Tiryns (Il. 2,555), Asine (Il. 2,560), Mycenae (Il 2,565), Donussa (Il. 2,570), Helice (Il. 2,575), Bryseae, Pharis (Il. 2,580), Helos (Il. 2,580), Arene, Dorium, (Il. 2,590) and Thuria/Antheia (Il. 9,150).

74 See Hejnc (1961).

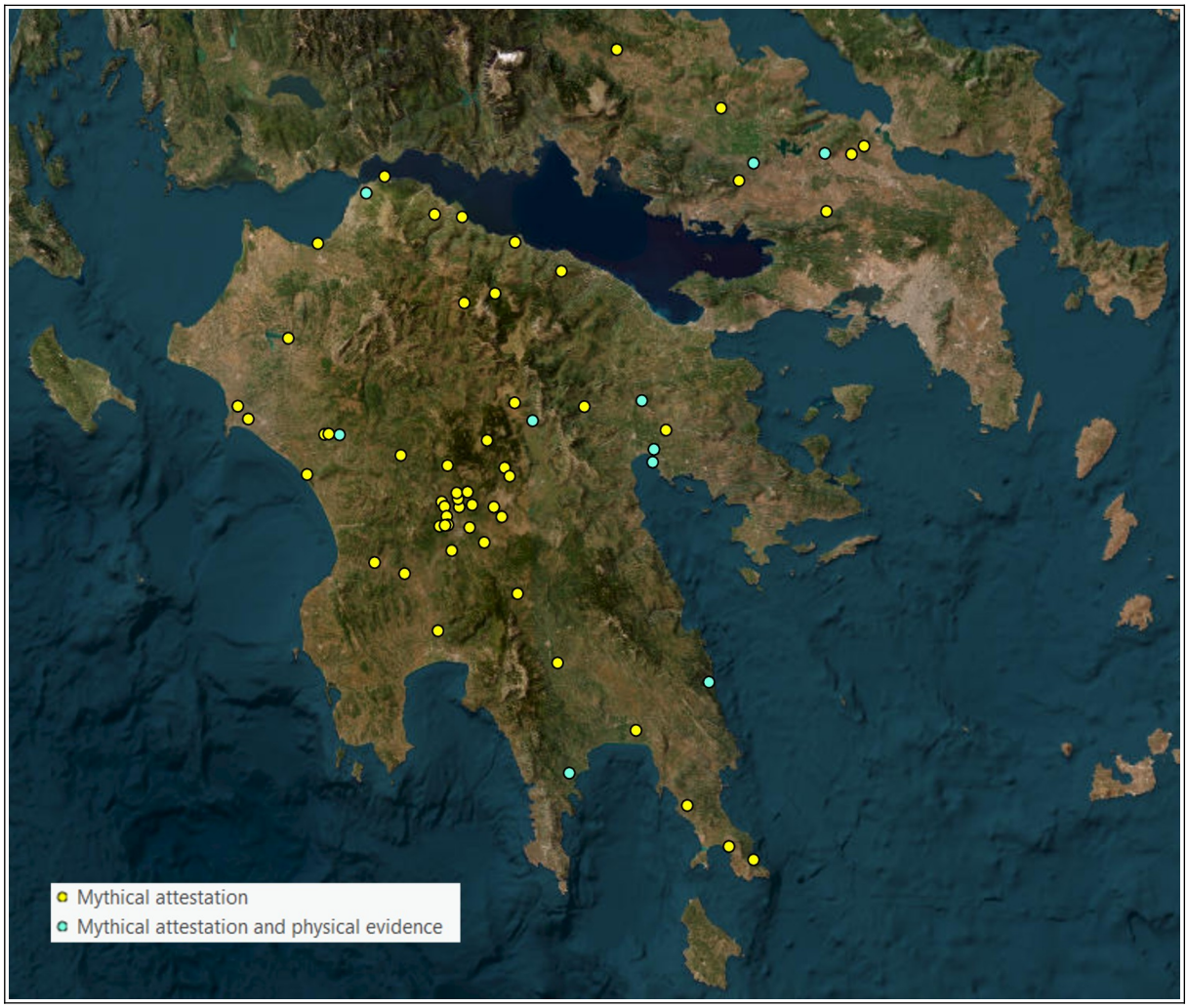


Fig. 10: Chart of the distribution of all the abandoned settlements tied to the mythical tradition visited and described by Pausanias in his *Periegesis*.

Specifically, a well-defined cluster of mythological attestations can be identified in the very heart of this region, in the area of the Maenalus, Ostracina and Thaumasius mountains. Next are the regions of Elis and Laconia (each with 8 elements). Elis in particular presents a second defined cluster, easily detectable in the area of Olympia / Pisa.[75] Next are Achaea (7 elements), Boeotia (6 elements), Argolis (5 elements) Messenia (3 elements), and Phocis (1 element).[76]

From this list, it is possible to further define data that may allow to evaluate those abandoned settlements which not only were linked to a specific myth, but also retained tangible elements recognized by local tradition as connected to mythic events. These may correspond to anthropic elements that were considered to have been created by the protagonists of local myths, or even to features of the landscape, nevertheless ascribed to a similar nature. Following these considerations, the picture offered by the processed data changes significantly, with 3 elements found in Argolis, 2 in Boeotia and 1 in Laconia, Elis, Achaea and Arcadia respectively. In 14 cases, furthermore, there is a coincidence of mythic attestations and areas of worship.

75 For a survey on the area of Olympia / Pisa and its historical and mythical background, see Maddoli / Saladino (2013), 364–365; Ekroth (1999).

76 Tab. 8: List of the mythic attestations for the abandoned settlements from the *Periegesis*; see in the data repository: https://doi.org/10.11588/data/FO1QAT (last access 16.10.2024).

Conclusions

In light of the collected data, the use of *GIS* software has proven to be an extremely versatile tool in the representation and analysis of a literary text from classical antiquity, with results beyond the simple spatial recreation. The ability of these applications to store a vast amount of data on one hand, and to associate them to their proper location on the other have been crucial not only in gathering the material provided by Pausanias, but also in contextualizing it, aiming thus to facilitate the assessment of this case study. The possibility of visually representing and georeferencing these data allows for the evaluation not only of the distribution of certain elements but also of their own specific features, with the purpose of scientifically elaborating valid results. As the research progresses and new elements relative to the landscape of the *Periegesis* emerge, this *GIS*-based methodology can be constantly employed to update the collected data.

The overall collection of material gathered during the preliminary digitization of the *Periegesis* has thus allowed the reprocessing of the data into new cartography. A careful cataloging and representation of the terminology used by Pausanias and the causes of abandonment of the ancient settlements reported by the author have also effectively facilitated the assessment of features related to the history of the sites under examination. The degree of precision in the analysis of the text is reinforced by the choice to establish a relation between the mapped areas and the associated elements. This allows for a more detailed evaluation of data that usually escape similar mapping projects of the classical world. The resulting atlas has furthermore enabled the representation of various features associated not only with the phenomenon of the abandonment of settlements but also with the persistence of cultural memories and activities tied to these sites.

As the *Periegesis* atlas itself may allow the production of new cartography, the six additional maps, detailing various features related to the presence of abandoned settlements, can be used simultaneously to better investigate the contents of the digitized corpus. The use of *GIS* maps allows in fact to assemble several layers at the same time, thus appreciating several distinct features of the same area. Moreover, the capability of this tool to upload and share a project online not only makes the work available for research, either on this specific text or on this area in general, but also allows the data to be scrutinized by experts and collaborators. This in turn can lead to the creation of a community capable of managing and further enhancing the collected database.

Sources

Online sources and digital corpora

ArcGIS = http://www.esri.com/software/arcgis (last access 11.11.2024).

Brill's New Pauly = https://referenceworks.brillonline.com/browse/brill-s-new-pauly (last access 11.11.2024).

Data repository = https://doi.org/10.11588/data/FO1QAT (last access 11.11.2024).

Digital Atlas of The Roman Empire = https://www.lunduniversity.lu.se/lup/publication/dcaf05b8-a13b-4024-a47b-325dbb68a035 (last access 11.11.2024).

Digital Pausanias = https://gis.periegesis.org/ (last access 11.11.2024).

Pandektis = https://pandektis.ekt.gr/ (last access 11.11.2024).

Pelagios network = https://pelagios.org/ (last access 11.11.2024).

Perseus Digital Library = http://www.perseus.tufts.edu/hopper/ (last access 11.11.2024).

Pleiades – The Stoa Consortium = https://pleiades.stoa.org/ (last access 11.11.2024).

Recogito = https://recogito.pelagios.org/ (last access 11.11.2024).

Sistema Binario III Colloquio Internazionale dei Dottorandi. Sulle Molteplici Prospettive del Viaggio: Dimensione Reale e Virtuale = https://studiumanistici.unimc.it/it/site-news/eventi/2021/sistema-binario-iii-colloquio-internazionale-dei-dottorandi (last access 11.11.2024).

ToposText = https://topostext.org/ (last access 11.11.2024).

QGIS = https://qgis.org/en/site/ (last access 11.11.2024).

ESRI = https://www.esri.com/en-us/home (last access 11.11.2024).

Text editions

Beschi / Musti (2013): L. Beschi / D. Musti (eds.), Pausania, Guida della Grecia. Libro I. Attica, Milan 2013.

Frazer (1898): J. G. Frazer (ed.), Pausanias's Description of Greece, 6 vols., London / New York 1898.

Jones et al. (1918): W. H. S. Jones / Litt. D. / H. A. Ormerod (eds.), Pausanias Description of Greece with an English Translation in 4 Volumes, London 1918.

Maddoli / Saladino (2013): G. Maddoli / V. Saladino (ed.), Pausania, Guida della Grecia. Libro VI. L'Elide e Olimpia, Milan 2013.

Moggi / Osanna (2008): M. Moggi / M. Osanna (ed.), Pausania, Guida della Grecia. Libro VII. L'Acaia, Milan 2008.

References

Akrivopoulou et al. (2012): C. Akrivopoulou / G. Dimitropoulos / S. G. Koutnatzis, The 'Kallikratis Program'. The Influence of International and European Policies on the Reforms of Greek Local Government, Istituzioni del Federalismo 3 (2012), 653–693.

Alcock (1994): S. E. Alcock, Breaking Up the Hellenistic World. Survey and Society, in: I. Morris (ed.), Classical Greece. Ancient Histories and Modern Archaeologies, Cambridge 1994, 171–190.

Alcock (1996): S. E. Alcock, Graecia Capta. The Landscapes of Roman Greece, Cambridge 1996.

Alcock et al. (2003): S. E. Alcock / J. F. Cherry / J. Elsner, Pausanias. Travel and Memory in Roman Greece, New York 2003.

Anderson (2005): G. Anderson, The Second Sophistic. A Cultural Phenomenon in the Roman Empire, London 2005.

Arafat (1992): K. W. Arafat, Pausanias' Attitude to Antiquities, The Annual of the British School at Athens 87 (1992), 387–409.

Bacher (1919): G. Bacher, De Pausaniae studiis Homericis, Halle 1919.

Bearzot (1988): C. Bearzot, La Grecia di Pausania. Geografia e cultura nella definizione del concetto di Ἑλλάς, in: M. Sordi (ed.), Geografia e storiografia nel mondo classico, Milan 1988, 90–112.

Bearzot (1992): C. Bearzot, Storia e storiografia ellenistica in Pausania il Periegeta, Venice 1992.

Berchet (1986): J.C. Berchet, Et in Arcadia ego!, Romantisme 51 (1986), 85–104.

Bodenhamer et al. (2010): D. J. Bodenhamer / J. Corrigan / T. M Harris, The Spatial Humanities. GIS and the Future of Humanities Scholarship, Bloomington 2010.

Bruit Zaidman (2003): L. Bruit Zaidman, La Notion d'Archaion dans la Périégèse de Pausanias, in: G. Lachenaud / D. Longrée (eds.), Grecs et Romains aux Prises avec l'Histoire. Représentations, Récits et Idéologie, Colloque de Nantes et Angers 12–15 septembre 2001, Rennes 2003, 21–30.

Bultrighini (1984): U. Bultrighini, Pausania 1,26,3 e la liberazione del Pireo, Rivista di Filologia e di Istruzione Classica 114 (1984), 54–62.

Bultrighini (1990): U. Bultrighini, La Grecia Descritta da Pausania. Trattazione Diretta e Trattazione Indiretta, Rivista di Filologia e di Istruzione Classica 118 (1990), 282–305.

Chrimes (1949): K. M. T. Chrimes, Ancient Sparta. A Re-examination of the Evidence, Manchester 1949.

Cooper et al. (2017): D. Cooper / C. Donaldson / P. Murrieta Flores, Literary Mapping in the Digital Age, New York 2017.

Damon (2016): C. Damon, Beyond variants: Some Digital Desiderata for the Critical Apparatus of Ancient Greek and Latin texts, in: M. Driscoll / E. Pierazzo (eds.), Digital Scholarly Editing. Theories and Practices, Cambridge 2016, 201–218.

Dimitropoulos et al. (2020): D. Dimitropoulos / E. Kyramargiou / G. Papakondylis, New Names – New Map: Renaming Settlements in Greece, 1831–2011, in: Αλλάζοντας τον χάρτη. Ζητήματα μετονομασιών στη Μεσόγειο, 19ος–20ός αιώνας, eds. Εθνικό Ίδρυμα Ερευνών, Ινστιτούτο Ιστορικών Ερευνών, Athens 2020, 19–28.

Drakopoulos (1997): E. C. Drakopoulos, Asea in Arcadia. A Case of Site Abandonment and Population Movement, L'Antiquité Classique 66 (1997), 301–309.

Ekroth (1999): G. Ekroth, Pausanias and the Sacrificial Rituals of Greek Hero-cults, in: R. Hägg (ed.), Ancient Greek Hero Cult: Proceedings of the Fifth International Seminar on Ancient Greek Cult, organized by the Department of Classical Archaeology and Ancient History (1999), Göteborg University, 21–23 April 1995, 145–158.

Geddes / Gregory (2014): N. Geddes / A. Gregory, Toward Spatial Humanities: Historical GIS and Spatial History, Bloomington 2014.

Goldhill (2001): S. Goldhill, Being Greek Under Rome: Cultural Identity, the Second Sophistic and the Development of Empire, Cambridge 2001.

Graser et al. (2017): A. Graser / B. Mearns / A. Mandel / V. Olaya / A. Bruy, QGIS: Becoming a GIS Power User, Birmingham 2017.

Habicht (1985): C. Habicht, An Ancient Baedeker and His Critics: Pausanias' Guide to Greece, Proceedings of the American Philosophical Society 129/2 (1985), 220–224.

Hejnc (1961): J. Hejnic, Pausanias the Perieget and the Archaic History of Arcadia, Prague 1961.

Hutton (2005): W. Hutton, Describing Greece. Landscape and Literature in the *Periegesis* of Pausanias, Cambridge 2005.

Isager (2009): I. Jacob, Destruction or Depopulation of Cities in Pausanias. Nikopolis, Aetolia, and Epirus, Proceedings of the Danish Institute at Athens 6 (2009), 201–215.

Johnson (2018): O. Johnson, Ruins: Classical Theater and Broken Memory, Ann Arbor 2018.

Knowles / Hillier (2008): A. K. Knowles / A. Hillier, Placing History. How Maps, Spatial Data and GIS Are Changing Historical Scholarship, Redlands 2008.

Kyramargiou et al. (2020): E. Kyramargiou / Y. Papakondylis / F. Scalora / D. Dimitropoulos, Changing the Map in Greece and Italy. Place Name Changes in the Nineteenth Century, in: The Historical Review / La Revue Historique Section of Neohellenic Research, Institute of Historical Research 17 (2020), 205–250.

Lünen / Travis (2013): A. Lünen / C. Travis, History and GIS. Epistemologies, Considerations and Reflections (1st ed.), Dordrecht 2013.

Mackil (2004): E. Mackil, Wandering Cities: Alternatives to Catastrophe in the Greek Polis, in: American Journal of Archaeology 108/4 (2004), 493–516.

Martí Henneberg (2023): J. Martí Henneberg, Creative Ways to Apply Historical GIS. Promoting Research and Teaching about Europe, Cham 2023.

Moggi (2002): M. Moggi, Pausania e Roma (Nota di lettura a VIII 27,1), Gerión 20/1 (2002), 435–449.

Moretti (1997): F. Moretti, Atlante del Romanzo Europeo: 1800–1900, Turin 1997.

Purcell (1987): N. Purcell, The Nicopolitan Synoecism and Roman Urban Policy, in: E. Chrysos (ed.), Nicopolis 1, Proceedings of the Ist Internat. Symposium on Nicopolis, Preveza, 23–29 Sept. 1987, 71–90.

Pretzler (2004): M. Pretzler, Turning Travel into Text: Pausanias at Work, Greece & Rome 51 (2004), 199–216.

Pritchett (1965): W. K. Pritchett, Studies in Ancient Greek Topography, University of California Press, Berkeley 1965.

Richter / Johnson (2017): D. S. Richter / W. A. Johnson, The Oxford Handbook of the Second Sophistic, Oxford 2017.

Rizakis (2014): A. D. Rizakis, Town and Country in Early Imperial Greece, in: J. Bintliff (ed.), Recent Developments in the Long-term Archaeology of Greece (Pharos Supplement), Leuven 2014, 241–267.

Scheidel (2004): W. Scheidel, Demographic and Economic Development in the Ancient Mediterranean World,: Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft 160/4 (2004), 743–757.

Scheuermann / Kroeze (2017): L. Scheuermann / J. Kroeze, Digital Humanities and Information Systems. Innovating Two Research Traditions, AMCIS 2017 Proceedings, AIS Electronic Library (AISeL) 2017.

Smith (1854): W. Smith (ed.), Dictionary of Greek and Roman Geography, London 1854.

Stewart (2013): D. R. Stewart, Most Worth Remembering: Pausanias, Analogy, and Classical Archaeology, Hesperia: The Journal of the American School of Classical Studies at Athens 82/2 (2013), 231–261.

Talbert / Bagnall (2000): R. J. A. Talbert / R. S. Bagnall. Barrington Atlas of the Greek and Roman World, Princeton 2000.

Taylor/Gregory (2022): J. E. Taylor/ I. N. Gregory, Deep Mapping the Literary Lake District: a Geographical Text Analysis, Lewisburg 2022.

Tomlinson (2014): R. A. Tomlinson, Argos and the Argolid: From the End of the Bronze Age to the Roman Occupation, New York 2014.

Travis et al. (2020): C. Travis / F. Ludlow / F. Gyuris, Historical Geography, GIScience and Textual analysis: Landscapes of Time and Place, Springer 2020.

Author contact information[77]

Marco Cornaglia, PhD
Alma Mater Studiorum Università degli Studi di Bologna
Via degli Ariani, 1
48121 Ravenna RA, Italy
E-Mail: Marco.cornaglia2@unibo.it

77 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.

DIGITAL CLASSICS ONLINE

Manuell-hermeneutische Forschung als Goldstandard: Zur Kalibrierung digitaler Analysetools in den Digital Humanities

Franziska Schropp, Michael Wittweiler, Thomas E. Konrad, Marie Revellio, Barbara Feichtinger

The development of digital text analysis tools is a sophisticated and challenging endeavor. Traditional hermeneutic intertextuality research has established itself as a highly effective benchmark for the meticulous calibration and validation of computer-assisted methods used in identifying quotations and allusions. A comprehensive collection of previous citation discoveries is termed a *gold standard*. This standard embodies the present state of scholarly knowledge and serves as an essential baseline for these digital tools. To ensure optimal and compatible outcomes within the Digital Humanities, these analytical tools are calibrated against this *gold standard* in order to apply them to less extensively studied works and author combinations. The efficacy of a *gold standard* as a benchmark for calibration largely hinges on the quality of the associated scholarly literature and the rigor of its compilation. Therefore, the *gold standard* itself must be the focus of ongoing critical evaluation. This article explores the intertextual comparison of Jerome's letters with the complete works of Virgil and Cicero, showcasing the *gold standard* as a pivotal reference in the calibration and evaluation of digital tools. It delves into the conceptual foundations and definitions necessary for a robust *gold standard*. It also establishes guiding principles for the compilation and creation of an effective *gold standard*, summarized under the VIATE principle. Furthermore, the article addresses the challenges encountered in transitioning from manual hermeneutic methods to digital research and introduces a digital-ready citation concept. This concept demonstrates the ambivalent simplification of hermeneutic classification strategies for intertextual references, illustrated through specific text passages. The dynamic productivity of methodological shifts from manual-hermeneutic approaches to digital analysis – and vice versa – is highlighted not only by the vast scalability of text corpora analyzable through digital routines but also through critical discussions of existing research paradigms and ongoing reflections on the concepts of citation and intertextuality.

Einleitung

Die Untersuchung von Beziehungen zwischen Texten ist seit der Antike fester Bestandteil literaturwissenschaftlicher Bemühungen.[1] Im prädigitalen Zeitalter waren Forschende auf ihre eigene Vertrautheit mit verschiedensten Texten angewiesen, um intertextuelle Referenzen zu erkennen; inzwischen unterstützen zahlreiche digitale Tools die Forschenden im Feld der Intertextualitätsforschung.[2] Diese reichen von (mehr oder weniger justierbaren) Suchmasken, etwa für die *Library of Latin Texts* (*LLT*), bis hin zu ausgefeilten Textvergleich-Tools, beispielsweise dem *Tesserae-Project*,[3] die es erlauben, digital verfügbare Texteditionen auf textuelle Gemeinsamkeiten zu untersuchen. Freilich bleiben die digitalen

1 Vgl. auch das Prinzip *imitatio et aemulatio*, dazu Penzenstadler (2006) und Zimbrich (2006).

2 Zur Intertextualitätsforschung im Zusammenhang mit den Digital Humanities s. insbesondere Coffee (2018).

Analysewerkzeuge effiziente *Hilfsmittel* – ein manuell-hermeneutischer Blick, um die Ergebnisse zu bestätigen und deren Qualität einzuschätzen, ist immer noch unerlässlich.

Der im Rahmen von Revellio[4] entwickelte und im aktuellen Projekt[5] von uns ausgebaute *mixed-methods*-Ansatz nutzt ebenfalls digitale Detektion von Text-Text-Übereinstimmungen und hermeneutische Ansätze, um in den Briefen des Hieronymus kulturelle Hybridisierungsprozesse der Spätantike, wie sie sich in Klassikerzitaten manifestieren, zu untersuchen. So ist nicht erst bei der finalen Beurteilung, ob die digital erhobenen *matches* auch tatsächlich intertextuelle Referenz darstellen, hermeneutisches *close reading* unverzichtbar. Bereits in der ersten Phase der (Weiter-)Entwicklung und Etablierung eines – von uns bewusst als *transparent* intendierten – digitalen Analyseinstrumentariums[6] ist die Bezugnahme auf hermeneutische Grundlagen zur Kontrolle, Beurteilung und bedarfsgerechten Justierung der Analyseroutinen zielführend.

So ziehen wir für die Kalibrierung und ‚Eichung' der digitalen Analyseverfahren als bewährten *Maßstab* die Ergebnisse der traditionell-hermeneutischen Intertextualitätsforschung heran. Wir erstellen hierfür eine möglichst lückenlose Sammlung sämtlicher bisheriger Zitatfunde und bereiten sie digital auf. Dieser *manuelle Goldstandard* repräsentiert somit den *status quo* des Forschungsfeldes, entlang dessen die digitalen Analysewerkzeuge fortwährend getestet und ausgerichtet werden können. Vor allem aber repräsentiert er unsere Mindestanforderung an die digitalen Tools: Alle bekannten Zitatfunde müssen auch von diesen ausgewiesen werden – darüber hinaus können sie neue digital generierte Zitatfunde ausgeben.

Der Wert eines solchen *manuellen Goldstandards* als Referenzpunkt für Kalibrierungsprozesse hängt freilich maßgeblich von der Qualität der einschlägigen Forschungsliteratur sowie der Akkuratesse seiner Erstellungsweise ab. Somit hat der *Goldstandard* selbst Gegenstand kritischer Reflexion zu sein.

Im Folgenden wird daher der *manuelle Goldstandard* als Referenzpunkt in der Kalibrierung und Evaluation digitaler Tools vorgestellt und am Beispiel des Vergleichs der Briefe des Hieronymus mit Vergils *Bucolica*, *Georgica* und *Aeneis* sowie mit Ciceros Gesamtwerk diskutiert. Nach konzeptionellen Grundlagen und Begriffsbestimmungen wird ein Lösungsansatz für einen digital operationalisierbaren Zitatbegriff eingeführt. Bevor die Herausforderungen thematisiert werden, die sich im Übergang von der manuellen zur digitalen Forschung stellen, werden leitende Kriterien für die Erstellung eines einsatzfähigen *manuellen Goldstandards* aufgestellt.

Grundsätzlich ist anzumerken, dass jenseits aller digitalen Einsatzmöglichkeiten der von uns erstellte *manuelle Goldstandard* zu Hieronymus' *Epistulae*–Vergil und Hieronymus' *Epistulae*–Cicero natürlich auch für die hermeneutische Intertextualitätsforschung ein Gewinn ist, da er eine überprüfte Zusammenschau *sämtlicher* bisher in der Forschung entdeckter und gesicherter Referenzen auf Basis von Wortübereinstimmungen bietet, die als umfassender Zitatpool für weiterführende Analysen und Interpretationen zur Verfügung steht.[7]

---

3 Für die *LLT* siehe https://www.brepols.net/series/LLT-O (letzter Zugriff 08.04.2024), für das *Tesserae-Project* siehe https://tesserae.caset.buffalo.edu/ (letzter Zugriff 08.04.2024).

4 Vgl. Revellio (2022).

5 *Zitieren als narrative Strategie. Eine digital-hermeneutische Untersuchung von Intertextualitätsphänomenen am Beispiel des Briefcorpus des Kirchenlehrers Hieronymus*; gefördert durch die DFG, Projekt-Nr. 382880410; https://gepris.dfg.de/gepris/projekt/382880410?context=projekt&task=showDetail&id=382880410& (letzter Zugriff 08.04.2024).

6 Zur digital-hermeneutischen Routine s. Schropp et al. (2024).

7 Die entsprechende Datensammlung *Expandable database of verified quotations and allusions in Jerome* ist abrufbar unter https://doi.org/10.11588/data/FVCULR (letzter Zugriff 29.10.2024). Zunächst bereitgestellt für Cicero und Vergil handelt es sich um *Work in Progress*: So wird die Datensammlung im Rahmen des Projektes um weitere antike Autoren erweitert und soll letztlich ein umfangreiches Nachschlagewerk zur Verwendung der Klassiker in den Werken des Hieronymus darstellen.

Konzeptionelle Grundlagen

Zu Begriff und Struktur

Die Verwendung des Begriffs *manueller Goldstandard* für die Sammlung der aus traditionell-hermeneutischer Lektüre bekannten Zitate und Anspielungen, die zur Evaluation der computergestützten Routinen und der daraus generierten Ergebnisse nutzbar gemacht werden können, geht auf Revellio zurück.[8] *Manuell* dient dabei als Gegenbegriff zu *digital*. Der Begriff *Goldstandard*, der ursprünglich aus dem Englischen stammt und zunächst einen vorgeschriebenen Reinheitsgrad für Gold angab, wurde später in übertragener Bedeutung auch allgemein für einen Referenzpunkt höchster Qualität benutzt.[9] In diesem Sinne verweist er darauf, dass der jeweilige Stand der Forschung als Maß gesetzt wird, an dem die neuen Verfahren ausgerichtet werden. Im Moment des Methodenwechsels wird dadurch an die Zitate und Anspielungen, die in der Forschung bereits anerkannt sind, angeknüpft. Grundsätzlich ist daher zu konstatieren, dass ein solcher *manueller Goldstandard* – entgegen seiner Benennung als Standard – freilich nicht *a priori* als objektiv gültig und unfehlbar betrachtet werden kann, umfasst er doch die Momentaufnahme eines spezifischen Forschungsstandes, der veränderlich ist und sich im Einzelnen mitunter durch divergente Anschauungen auszeichnet.

Das Prinzip eines solchen Maßstabes lehnt sich an das Konzept des *Goldstandards* im Kontext des *Natural Language Processing* an. Dort bezieht sich *Goldstandard* auf „einen manuell überprüften und annotierten Datensatz, der als Maßstab zur Evaluation von automatisierten Annotationssystemen wie PoS-Taggern oder NER-Tools verwendet wird, die auf Machine-Learning-Algorithmen basieren.“[10]

Im Rahmen unserer Arbeit wird ein Datensatz aus intertextuellen Bezügen zwischen Hieronymus und den Klassikern Vergil und Cicero erstellt. Als Annotation wird hier die Zuschreibung eines Zitats gemäß eines transparent definierten Zitatbegriffs verstanden. Indem die Ergebnisse der traditionell-hermeneutischen Forschung als Grundlage dienen, ist die Annotation ein Ergebnis manueller Recherche und, jedenfalls *prima facie*, durch Experten des Feldes überprüft.[11]

Ein operationalisierbarer Zitatbegriff

Zentral für die Kollationierung und Annotation des Datensatzes ist zunächst ein transparenter, eindeutiger Zitatbegriff. In der hermeneutischen Intertextualitätsforschung ist es üblich, nicht nur zwischen (wörtlichen) Zitaten und Referenzen allgemeinerer Natur zu differenzieren, sondern Übergänge von exakter Wortkongruenz bis hin zur rein semantischen Allusion mithilfe von Synonyma (häufig durch den Hinweis ‚cf.‘) nahezu stufenlos zu skalieren und mittels vielfältiger Begrifflichkeiten wie Anspielung, Paraphrase, Reminiszenz und ähnlichen unterschiedliche Grade eines Bezugs zweier Textstellen zueinander zu differenzieren und auszudrücken. Dabei bleibt häufig vage, was exakt unter dem einen oder anderen dieser Begriffe verstanden wird oder wo die Grenzen verlaufen. Wenn aber aus der Forschungstradition ein Referenzpunkt als Kalibrierungsinstrument für digital unterstützte, experimentelle Methoden geformt werden soll, müssen insbesondere diese fließenden Übergänge und Grauzonen durch exakte Einteilung klassifiziert werden.

Für die Erstellung des *manuellen Goldstandards* definieren wir – aus den hermeneutisch erhobenen Referenzen als kleinstem Zitatsignifikanten abgeleitet – diejenigen intertextuellen Referenzen als *Zi-*

8 Zum Begriff vgl. Revellio (2022), 23.

9 Vgl. OED s.v. gold standard.

10 forTEXT (2016). Dabei wird oft auf mindestens zwei unabhängig erstellte Annotationen zurückgegriffen, um ein gewisses Maß an Objektivität zu erhalten.

11 Vgl. zu den Unstimmigkeiten und Divergenzen in den Urteilen von Forschenden unten S. 84ff. *Potentiale und Grenzen des manuellen Goldstandards*.

tat, die sich durch mindestens zwei übereinstimmende Lemmata auszeichnen, die – innerhalb eines (Neben-)Satzes – in einem näheren Umfeld positioniert sind.[12]

Alle anderen intertextuellen Referenzen, die sich nicht auf der Ebene der Wörter widerspiegeln, sondern beispielsweise durch die Verwendung von Synonyma allein durch den semantischen Gehalt entstehen, gelten folglich nicht als ein solches Zitat. Sie werden mit der zusätzlichen Einschätzung *confer* ‚(cf.)' versehen und in der Folge im digitalen Suchprozess nicht beachtet.[13] Grundsätzlich und einschränkend ist zu dem von uns angewandten Zitatbegriff anzumerken, dass eine beliebige Übereinstimmung von mindestens zwei Lemmata noch kein Zitat ist. Dafür ist ein gewisser semantischer Mehrwert, ein Sinnstiftungsmoment der Bezugnahme selbst konstitutiv. Bei den durch die traditionell-hermeneutische Forschung bestätigten intertextuellen Referenzen kann diese Sinnstiftung als gesichert vorausgesetzt werden.[14] In digitalen Matchingverfahren müssen potenzielle Neufunde dagegen erst durch eine hermeneutische Beurteilung als tatsächliche – sinnhafte – Zitate abgesichert werden.[15]

Unser *mixed-method*-Prozess vollführt somit eine Bewegung vom hermeneutischen Lesen zur computerbasierten Detektion und wieder zum hermeneutischen Lesen zurück: Zuerst wird aus dem Bestand hermeneutisch gesicherter, sinnhafter Zitate ein digital operationalisierbares ‚Minimal-Zitat' (zweier übereinstimmender Lemmata) abgeleitet. Dieses Minimal-Zitat liegt der Ausformung des *manuellen Goldstandards* zugrunde. Zugleich findet es auch in den digitalen Suchroutinen Anwendung, deren Ergebnisse dann *potenzielle* Zitate in Form von (mindestens) zwei Lemmata-Kongruenzen darstellen. Dieser Output wird schließlich in einem hermeneutischen *close reading* geprüft, um die sinnstiftenden, tatsächlichen Zitate zu selektieren und dem Gesamtbestand der Referenzen zuzuführen.

Unser Ziel ist es dabei, die digitalen Routinen – in mehreren transparenten Verfahrensstufen – so effizient selektierend zu gestalten, dass die Zahl der *matches* überschaubar ist und somit tatsächlich einem *close reading* unterzogen werden kann.

Zur Verwendung des *manuellen Goldstandards* als Maßstab

Übliche Metriken zur Evaluation digitaler Instrumentarien sind *recall* und *precision*. *Precision* bezieht sich auf das Verhältnis *gefundener relevanter Ergebnisse* zur Gesamtzahl *generierter Funde*. *Recall* bezeichnet das Verhältnis der *gefundenen relevanten Ergebnisse* zur Gesamtzahl der *relevanten Ergebnisse*.[16] Die *relevanten Ergebnisse* sind im vorliegenden Fall die Zitate in den Briefen des Hieronymus (entsprechend dem angeführten Zitatbegriff). Die *generierten Funde* sind der gesamte Output des digitalen Tools, also sämtliche Textpaare mit mindestens zwei übereinstimmenden Lemmata. Ziel der Verbesserung des digitalen Prozesses ist es, die Anzahl der relevanten Ergebnisse maximal hoch, dabei aber die Gesamtzahl der generierten Funde eher gering zu halten – *precision* und *recall* sollen also beide möglichst hoch sein.

Die relevanten Ergebnisse sind dabei eine *a priori* zunächst unbekannte Größe, da das digitale Instrumentarium gerade darauf abzielt, bisher unentdeckte Zitate ausfindig zu machen. Zugleich ist durch

12 Vgl. zur digitalen Umsetzung und zu den Parametern auch Schropp et al. (2024).

13 Allerdings bleibt ihre Evidenz in der *confer*-Rubrik des *manuellen Goldstandards* gesichert, so dass sie in eine abschließende Gesamtschau *aller* Referenzen für die Neufassung des Forschungsstands, der durch die zu generierenden Neufunde erweitert wird, problemlos wieder eingegliedert werden können.

14 Für die digitale Routine hingegen ist diese Unterscheidung zentral: Den Großteil der Ergebnisse unseres Analysetools machen nämlich solche Textparallelen aus, die aus nur sehr wenigen geteilten Wörtern bestehen. Hier ist das hermeneutische *close reading* zentral, das den für die Ergebnisse der traditionell-hermeneutischen Forschung bereits geprüften semantischen Mehrwert zu- oder absprechen kann.

15 Zur hermeneutischen Absicherung digitaler Neufunde vgl. auch die Diskussionen in Revellio (2022), 189–276.

16 Vgl. hierzu Murphy (2007), 2–3 sowie Bird et al. (2010), dort Kap. 6.3.3. Zur Adaption auf die vorliegende Routine vgl. Revellio (2022), 131–133.

die traditionell-hermeneutische Forschung aber bereits eine – je nach Autorenpaarung verschiedene – Menge an intertextuellen Referenzen und damit an relevanten Ergebnissen bekannt, die im *manuellen Goldstandard* verzeichnet sind. Im Rahmen des *recalls* kommt also der konkretisierte *manuelle Goldstandard* zum Einsatz, der eine *Mindestmenge* aufzufindender Zitate und damit relevanter Ergebnisse angibt. Konkret bedeutet dies, dass das digitale Instrumentarium nur dann positiv evaluiert werden kann, wenn der Output, also die computergenerierten Funde, mindestens die nach dem zugrundeliegenden Zitatbegriff als Zitat klassifizierten Stellenpaare des *manuellen Goldstandards* umfasst.[17] Das heißt, dass alle Parallelstellen, die im Rahmen des *manuellen Goldstandards* entsprechend des Zitatbegriffs als Zitat klassifiziert sind, unter den digital generierten Funden erwartet werden. *Confer*-Referenzen hingegen können unter den digital generierten Funden auftauchen, werden aber nicht erwartet.

Während der *manuelle Goldstandard* damit zunächst als Evaluationskriterium dient, wird er durch wiederholtes Evaluieren letztlich auch zur Kalibrierung genutzt, indem in jedem Schritt der Erweiterung des digitalen Analyseinstrumentariums geprüft wird, inwieweit der *manuelle Goldstandard* von den Ergebnissen der digitalen Routine abgedeckt ist und davon ausgehend Optimierungsstrategien entwickelt werden können.

Praktische Umsetzung am Beispiel der Briefe des Kirchenlehrers Hieronymus

Verfahren

Um als Evaluationskriterium und Bezugsrahmen für die Entwicklung neuer digitaler Instrumente dienen zu können, muss der *manuelle Goldstandard* insofern *vollständig* sein, als er *alle durch traditionell-hermeneutische Forschung identifizierten* intertextuellen Bezüge umfasst.[18] Somit ist die sorgfältige und umfassende Sichtung der Forschungsliteratur – nach dem Vier-Augen-Prinzip – grundlegend.

Alle intertextuellen Referenzen, bei denen die Forschungsliteratur zwei konkrete Textpassagen angibt und deren Bezug aufeinander sich im Wortmaterial selbst spiegelt, werden anhand des oben genannten Zitatbegriffs transparent und eindeutig eingestuft und in einer Tabelle festgehalten. Um Forschungsdiskussionen über Divergenzen und Unstimmigkeiten in den Urteilen einzelner Forschender nicht aus den Augen zu verlieren,[19] werden diese ebenfalls in einer gesonderten Kommentarspalte erfasst. Diese hat zwar auf den Einsatz als Evaluations- und Kalibrierungsinstrument keinerlei Auswirkungen, für das spätere hermeneutische *close reading* des Outputs hingegen kann sie als wertvolle Sammlung von Forschungsdebatten dienen. Auf diese Art wird eine umfangreiche Übersicht erarbeitet, die *idealiter* sämtliche von der manuell-hermeneutischen Forschung aufgezeigten intertextuellen Referenzen mit dem jeweiligen Sekundärstellenbeleg verzeichnet. Der aufbereitete Datensatz, der nur manuell durch (Datenbank-)Recherche und Sichtung der gefundenen Literatur erstellt werden kann, enthält somit in kondensierter Form den vorhandenen Forschungsstand.

17 Ein weiteres Ziel besteht darin, die Zahl der darüber hinausgehenden computergenerierten Funde, unter denen man neben bislang unentdeckten Zitaten auch ein Vielfaches mehr „Nieten“ findet, auf Seiten dieser „Nieten“ eher gering zu halten, um das *close reading* noch bewältigen zu können, vgl. zu diesen beiden Optimierungsregeln näher Revellio (2022), 130–133.

18 Vollständigkeit bedeutet hier also nicht, dass der *manuelle Goldstandard* bereits *alle überhaupt auffindbaren* Zitate und Allusionen beinhaltet. Dies kann nie final postuliert werden und würde die Suche nach weiteren Zitaten und somit auch jede Entwicklung neuer Detektionsinstrumente obsolet werden lassen.

19 Zu den Unstimmigkeiten und Divergenzen in den Urteilen verschiedener Forschender vgl. unten S. 84ff. *Potentiale und Grenzen des manuellen Goldstandards*.

Jede einzelne ausgewiesene Parallelstelle zu hinterfragen und nach eigenständiger Prüfung zu bestätigen oder zu verwerfen, würde die Erstellung des *manuellen Goldstandards* zu einem gewaltigen Unterfangen machen, und ist im engen Rahmen einer basalen Vorarbeit für digitale Untersuchungen schlichtweg nicht zu leisten. Es ist also notwendig, sich – stets im Wissen um die potenzielle Fehlbarkeit – zumindest weitgehend auf die einschlägige Literatur zu verlassen. Insbesondere für die weitere Nutzung als Kalibrierungselement im Rahmen der (Weiter-)Entwicklung eines digitalen Instrumentariums sollten die Ergebnisse der traditionell-hermeneutischen Forschung *verlässlich* sein, was ein angemessenes Maß an Vertrauen in die Arbeit der Forschenden voraussetzt.

Bei den von mehreren Forschenden einstimmig identifizierten Zitaten ist die Übertragung in den *manuellen Goldstandard* unkompliziert. Insbesondere das Vorliegen einschlägiger Vorarbeiten wie kritische Textausgaben mit Similienapparat, Indexbände und einschlägige Monographien vereinfacht in dieser Hinsicht den Zugang zum Forschungsstand, da der Mehrwert solcher Werke gerade darin liegt, Forschungsdiskussionen zu sichten und die Urteile abwägend zusammenzufassen. Somit nehmen sie bereits einen wichtigen Schritt in der Qualitätsprüfung vor.

Bei den nicht einstimmig identifizierten oder kontrovers bewerteten Zitaten ist die Übertragung dagegen erschwert. Doch die Tatsache, dass Unstimmigkeiten in der Regel in der Forschung diskutiert werden, bestätigt und rechtfertigt ein grundsätzliches Vertrauen in die kritische und präzise Arbeit der Forschenden. Die auftretenden Divergenzen in der Beurteilung des Zitatwerts[20] selbst sind bei der Erstellung des *manuellen Goldstandards* letztlich sogar hilfreich, weil sie jene (potenziellen) Referenzen markieren, die der nochmaligen Überprüfung und finalen Kategorisierung bedürfen.

Corpus

Die Werke des Kirchenlehrers Hieronymus und insbesondere sein stark rezipiertes Briefcorpus bieten sich als Testfeld in besonderem Maße an: Hochgebildet und höchst vertraut mit den heidnischen Klassikern, zugleich aber auch Christ, spiegeln seine Werke die Spannung zwischen heidnischer Kultur und deren christlicher Ablehnung.[21] Am deutlichsten kommen die Differenzlinien in der Traumerzählung zum Ausdruck, in der Hieronymus – von Gott für seine Lektüre der klassischen Literatur gestraft – schwört, der heidnischen Literatur zu entsagen.[22] Entgegen diesem Versprechen finden sich in seinen Werken freilich weiterhin zahlreiche Bezüge zu den Klassikern, was einerseits von seinen Zeitgenossen heftig kritisiert wurde, andererseits aber bis in die Gegenwart seine Werke für die Intertextualitätsforschung besonders interessant gemacht und zu reicher Forschungsliteratur geführt hat.

Folgende drei Werke sind *in puncto* Zitatforschung im Falle des Hieronymus in besonderem Maße einschlägig: Der Similienapparat der Textedition von Hilberg, der zugehörige Indexband von Kamptner sowie Hagendahls Untersuchung über die Beziehung der Kirchenväter zu den Klassikern, deren zweiter Teil sich allein mit Hieronymus befasst.[23] Ausgehend von der listenartigen Zusammenstellung Kamptners und der Monographie Hagendahls werden bei der Erstellung des *manuellen Goldstandards* für das *Matching* Cicero–*Epistulae* bzw. Vergil–*Epistulae* alle in diesen Werken aufgeführten intertextuellen Referenzen wie Zitate, Paraphrasen und Allusionen gesammelt. Intertextualitätsphänomene allgemeinerer Art sind nicht aufgenommen, da die angestrebte digitale Zitatanalyse ausschließlich Wortübereinstimmungen detektiert.[24]

---

20 Vgl. dazu unten S. 84ff. *Potentiale und Grenzen des manuellen Goldstandards*.

21 Vgl. für eine beispielhafte Detailuntersuchung dieses Spannungsfeldes Feichtinger (2021).

22 Hier. epist. 22,30. Vgl. dazu Feichtinger (1997).

23 Dies sind die Bände Kamptner (1996) und Hagendahl (1958). Vgl. auch Revellio (2022), 23 zu Kamptner (1996) und Hagendahl (1958) als Grundlage.

24 So wird zum Beispiel die Übernahme von allgemeinen Topoi oder Motiven nicht als intertextuelle Referenz aufgenommen, sofern sie sich auf Ebene der Wortwahl nicht durch mindestens zwei übereinstimmende Lemmata niederschlägt

Zusätzlich zu den genannten Werken werden Einzeluntersuchungen mit dem Erscheinungsjahr von 1958 oder jünger berücksichtigt: Es wird davon ausgegangen, dass der Forschungsstand bis zum Jahr 1958 durch Luebeck (1872) und Hagendahl (1958) hinreichend und in einer gewissen Einheitlichkeit dokumentiert ist.[25] Für Cicero und Vergil wurden daher noch insgesamt 43 Einzelbeiträge, die seit 1958 erschienen sind und sich mit Reminiszenzen an Cicero und Vergil innerhalb der Briefe des Hieronymus befassen, ergänzend hinzugezogen.[26]

Wird jeder Vers oder Prosaparagraph einzeln gelistet, ergibt sich hieraus konkret eine Gesamtzahl von 448 intertextuellen Bezügen in den Briefen des Hieronymus auf Vergils *Bucolica*, *Georgica* und *Aeneis* sowie eine Gesamtzahl von 198 intertextuellen Bezügen in den Briefen des Hieronymus auf eines von Ciceros Werken.[27]

Im Falle des Hieronymus wird das Zusammentragen des Forschungsstandes dadurch erleichtert, dass dank der detaillierten Monographie Hagendahls sowie insbesondere des Indexbands von Kamptner bei der Erstellung des *manuellen Goldstandards* auf ein solides und recht einheitliches Grundmaterial zurückgegriffen werden kann. Nicht für jeden Autor ist der Similienapparat der entsprechenden Textausgabe derart umfangreich.

Forschungskontext

Während die Ausgangslage bei Hieronymus oder auch beispielsweise Augustinus durch einschlägige, umfängliche Monographien und listenartige Darstellungen günstig ist,[28] sieht dies beispielsweise bei Ambrosius oder Cyprian anders aus. Umfassende Studien, die eine Vielzahl an Bezügen zu unterschiedlichsten Prätexten beleuchten, sind ebenfalls nicht selbstverständlich.

Analoge Quantitäts- und Qualitätsunterschiede wie bei den Zieltexten finden sich auch bei den antiken Quellentexten: Während Bezüge der spätantiken Schriftsteller auf zentrale Autoren der lateinischen Literatur wie z.B. Cicero und Vergil vielfach untersucht und diskutiert werden, rücken andere heidnische Prätexte stärker in den Hintergrund und werden weniger beachtet. Ähnlich divergente Forschungsstände sind natürlich auch für die Intertextualitäts- und Zitatanalyse zwischen antiken heidnischen Autoren zu gewärtigen. Entsprechend gestaltet sich die Erstellung des *manuellen Goldstandards* je nach Paa-

---

(vgl. zum Kriterium auch S. 79 *Ein operationalisierbarer Zitatbegriff*). In epist. 46,12,3 spricht Hieronymus – um einen konkreten Fall zu nennen – den *messor*, *arator* und *uinitor* an, die „neben den *pastores* zur Standardbesetzung rustikaler Literatur“ gehören und deren „Affinität zum Gesang“ ein „Topos der bukolischen Dichtung ist“ (Feichtinger [2020], 241). Für solche Phänomene werden in der Forschungsliteratur selbstverständlich Belege angeführt, im zitierten Aufsatz beispielsweise Verg. ecl. 3,42; 10,32 und 2,10 sowie die *Ilias latina*. Allein die Tatsache, dass mehrere Textstellen zur Stärkung des Arguments und zur Illustration eines Topos' oder Motivs angeführt werden, zeigt aber die deutlich andere Qualität solcher Bezüge. Für die Entwicklung des digitalen Analyseinstrumentariums sind sie zunächst zu allgemein und unspezifisch, als dass sie für die Kalibrierung herangezogen werden könnten.

25 Zur Kritik der Unvollständigkeit am Indexband von Kamptner vgl. Rebenich (1999). Aus diesem Grunde werden Einzelbeiträge ab dem Jahr 1958 und nicht erst ab dem Jahr 1996 berücksichtigt.

26 Diese sind Abbott (1979), Adkin (1992a), Adkin (1992b), Adkin (1992c), Adkin (1993), Adkin (1997a), Adkin (1997b), Adkin (2000), Adkin (2006), Adkin (2011), Adkin (2013), Adkin (2018), Adkin (2019), Antin (1970), Ascone (2019), Audano (2019), Banniard (1988), Bauer (1975), Capone (2018), Clark (2006), Courcelle (1976), Feichtinger (2020), Hagendahl (1974), Hamblenne (1996), Iannaccone (1963), Jaureguizar (1970), Jeanjean (2004), Jeanjean (2015), Jeanjean (2018), Laurence (2004), Layton (2002), Marolla (2017), McGill (2007), Neuhausen (1984), Poinsotte (1992), Préaux (1958), Revellio (2022), Ronnenberg (2015), Rzepkowski (2011), Schneider (2003), Trillitzsch (1965), Tsartsidis (2020) und Viscido (1980). Gesichtet wurden weitaus mehr Beiträge, aufgeführt sind dagegen nur diejenigen, aus denen Material entnommen werden konnte, in denen also intertextuelle Bezüge genannt und besprochen werden.

27 Vgl. den zugehörigen Datensatz *Expandable database of verified quotations and allusions in Jerome*, abrufbar unter https://doi.org/10.11588/data/FVCULR (letzter Zugriff 29.10.2024).

28 Für Augustinus ist beispielsweise Hagendahl (1967) einschlägig, darüber hinaus O'Donnell (1980) und speziell zu Vergil außerdem MacCormack (1998).

rung von Ziel- und Quellentext sehr unterschiedlich und eignet sich mehr oder weniger zur Kalibrierung neuer Analyseverfahren.

Einen Spezialfall stellen dabei Paarungen von Quellen- und Zieltext dar, für die in der traditionell-hermeneutischen Forschung keinerlei intertextuelle Referenzen belegt sind. Bei einer solchen defizitären Ausgangslage lässt sich selbstverständlich kein *manueller Goldstandard* erstellen. Zur Evaluierung und Kalibrierung eines digitalen Analyseinstruments ist eine solche Paarung also nicht geeignet. Dennoch kann ein anhand anderer Autoren kalibriertes und evaluiertes Tool möglicherweise gewinnbringend auf eine solche Paarung angewandt werden, um mithilfe digitaler Methoden neue, bisher unerkannte Zitate zu detektieren.

Der Goldstandard des *manuellen Goldstandards*

Der *manuelle Goldstandard* stellt zum einen ein Nachschlagewerk aller bereits bekannter intertextueller Bezüge dar und ist andererseits zentraler Referenzrahmen für die Evaluation digitaler Tools zur Intertextualitätsforschung. Um beide Aufgaben befriedigend erfüllen zu können, erscheinen Erarbeitungsstandards erforderlich, die – insbesondere angesichts von Paarungen divergenter Quell- und Zieltexte recht unterschiedlicher Herausforderungen – als Gütekriterien der Qualitätssicherung dienen. Folgende Merkmale zeichnen einen qualitätsvollen *manuellen Goldstandard* aus:

- **Vollständigkeit**: Der *manuelle Goldstandard* umfasst alle in der Forschung anerkannten intertextuellen Referenzen.
- **Interoperabilität**: Der *manuelle Goldstandard* wird in einem gängigen Datenformat angeboten.
- **Adaptierbarkeit**: Der *manuelle Goldstandard* lässt sich entsprechend neuen Erkenntnissen der Forschung anpassen.
- **Transparenz**: Der Zitatbegriff wird offengelegt. Auf Debatten und Widersprüchlichkeiten innerhalb der Forschungsliteratur wird verwiesen.
- **Eindeutigkeit**: Die Klassifizierung der Parallelstellen ist konsistent und ausreichend bestimmt.

= VIATE-Prinzip

Diese fünf Grundprinzipien des VIATE-Konzepts können projektübergreifend als Leitfaden dienen, um die Nutzbarkeit der Datensätze eines *Goldstandards* sicherzustellen und ihre Wiederverwendbarkeit zu verbessern. Hiermit tragen sie dazu bei, dem Format des *Goldstandards* eine verlässliche und einheitliche Struktur zu geben, wodurch unmittelbar ein Mehrwert für die Forschung entsteht, insofern das Kuratieren des *Goldstandards* kein Selbstzweck ist, sondern er durch bereitstellende Dokumentation maßgeblich zum Wissen selbst beitragen kann.

Potenziale und Grenzen des *manuellen Goldstandards*

Die Zusammenführung des Zitatbefundes aus der Forschungsliteratur erscheint auf den ersten Blick ein banaler und in seiner Ergebnisqualität nur von der Verfügbarkeit der einschlägigen Literatur und der Sorgfalt des Kopisten abhängiger Sammelvorgang zu sein. Allerdings zeigt sich, dass die Nutzung hermeneutisch ermittelter Forschungsergebnisse für eine digitale Anwendung, die wie oben erörtert ei-

ne binäre Klassifizierung erfordert, also ausschließlich zwischen Zitat und Nicht-Zitat differenziert, doch größere Herausforderungen mit sich bringt.

Dies stellt Forschende vor vielfältige Herausforderungen im Hinblick auf Operationalisierungsprozesse. Denn selbst wenn ein umfangreicher Zitatbefund vorliegt, ist dieser oft – in mehrfacher Hinsicht – uneindeutig: Der Forschungsstand, der über Monographien, Kataloge und (Einzel-)Beiträge verteilt sein kann, bietet teils uneinheitliche und widersprüchliche Urteile, was die Klassifizierung der Parallelstellen beim Sammeln des Materials für den *manuellen Goldstandard* erschwert. Solche Unstimmigkeiten betreffen dabei nicht nur die Klassifizierung, sondern auch die basale Einschätzung, ob überhaupt eine Bezugnahme auf einen bestimmten Prätext vorliegt. Die Notwendigkeit der Operationalisierung mit ihren strengen Klassifizierungsfragen wirft hier freilich auch theoretische Herausforderungen der Zitatforschung im Allgemeinen auf und trägt durch die eingehende Diskussion und Schärfung des Zitatbegriffs zu Erkenntnisgewinn über Methodengrenzen hinweg bei.[29]

Während die Anwendung eines operationalisierbaren Zitatbegriffs zur Vereinheitlichung der hermeneutischen Klassifizierungen führt, bedeutet sie auch eine gewisse Einengung des Zugangs, der im offenen Diskurs grundsätzlich aus verschiedenen Blickwinkeln erfolgen kann. Die folgenden Beispiele zeigen auf, in welchen Fällen der vorgestellte Zitatbegriff im Rahmen der Erstellung des *manuellen Goldstandards* zu einer solchen – ambivalenten – Simplifizierung führen kann. Exemplarisch können so Chancen, aber auch Grenzen dieser Reduktion diskutiert werden.

Beispiel I: Wortreihenfolge

Dass die traditionell-hermeneutische Forschung auch bei formgleicher Übereinstimmung von mehreren Worten zu dem Urteil kommen kann, dass es sich *nicht* um eine eindeutige, wörtliche Referenz handelt, oder sich zumindest in ihrer Einschätzung uneins ist, zeigt folgendes Beispiel:[30]

> *quaerenti et tectis urbis sine fine furenti / infelix simulacrum atque ipsius umbra Creusae / uisa mihi ante oculos et nota maior imago. / obstipui, steteruntque comae et* uox faucibus haesit.
> (Verg. Aen. 2,771–774)
>
> „Während ich suchte, die Häuser der Stadt ohne Ende durchraste, / da erschien mir der Schatten Krëusas selbst vor den Augen, / ihr unglückliches Bild, das größer war, als ich sie kannte. / Starr war ich da, die Haare gesträubt, und mir stockte die Stimme."

> *obstipuit sancta et grauis femina alienum habitum in nepte conspiciens, mater gaudio stabat adtonita. utraque uerum non credere, quod uerum esse cupiebant.* haesit uox faucibus *et inter ruborem atque pallorem metumque ac laetitiam cogitationes uariae mutabantur.*
> (Hier. epist. 130,5,5)
>
> „Diese würdevolle und heilige Frau erstarrte, als sie das ungewohnte Gewand an der Enkelin erblickte, die Mutter stand still, vor Freude ergriffen. Beide konnten kaum glauben, dass es wahr geworden sei, was sie sich so lange gewünscht hatten. Die Worte blieben ihnen im Halse stecken und zwischen Erröten und Erbleichen, Furcht und Freude, wechselten sich verschiedene Gedanken ab."

29 Vgl. unten S. 92f.

30 Vgl. Quellenverzeichnis für die verwendeten Textausgaben und Übersetzungen. Die Hieronymus-Übersetzungen stammen von den Autor:innen.

Kamptner listet diese Parallelstelle mit der Bezeichnung „cf.“. Hagendahl zählt sie zu den „less obvious reminiscences“[31] und Jeanjean markiert sie in seiner Aufzählung der Zitate in der Korrespondenz des Hieronymus überhaupt nicht speziell.[32]

Nach dem eingangs definierten Zitatbegriff ist diese Referenz dagegen eindeutig als Zitat einzustufen: Es handelt sich um dieselben Wörter, lediglich das Prädikat rückt von der letzten auf die erste Stelle. Hier unterstützt also die Anwendung des vorgeschlagenen Zitatbegriffs eine klärende Präzisierung der uneinheitlichen hermeneutischen Ergebnisse.[33]

Beispiel II: Gattungsreferenz

Auch folgende intertextuelle Referenz ist mit dem Zitatbegriff des *manuellen Goldstandards* klar als Zitat einzustufen:

> *et* curuo *Saturni dente relictam / persequitur* uitem attondens *fingitque putando. / … / iam uinctae uites, iam falcem arbusta reponunt, / iam* canit *effectos extremus* uinitor *antes; / sollicitanda tamen tellus puluisque mouendus / et iam maturis metuendus Iuppiter uuis.*
> (Verg. georg. 2,406b–419)

> „stellt dem kahlen Weinstock nach mit dem krummen / Zahne des Saturnus und stutzt ihn und formt ihn durch Schneiden. / … / Schon sind die Reben gebunden, Gesträuch gibt Ruhe der Sichel, / weil die Reihen stehn, frohlockt schon der hinterste Winzer; / dennoch ist Erde zu lockern und aufzuhäufeln der Staub, und / reife Trauben sogar müssen Juppiter immer noch fürchten.“

> *quocumque te uerteris, arator stiuam tenens alleluia decantat, sudans messor psalmis se auocat et* curua adtondens uitem *falce* uinitor *aliquid Dauiticum* canit.
> (Hier. epist. 46,12,3)

> „Wohin du dich wendest, hält ein Landmann den Pflugsterz und singt das Halleluja, unterhält sich ein schwitzender Mäher durch das Singen von Psalmen und ein Weinbauer, der mit einem gekrümmten Winzermesser die Rebe beschneidet, singt etwas von David.“

---

31 Hagendahl (1958), 257. Hagendahl verweist zudem auf die zweite Parallelstelle Verg. Aen. 3,48 – diese deutet auch Hilberg im Similienapparat durch „etc.“ an, allerdings ohne konkrete Nennung der *Aeneis*-Passage.

32 Jeanjean (2018), 196–197. Dabei markiert Jeanjean Allusionen als solche und zeigt durch Kursivdruck in allen anderen Fällen lediglich die Veränderungen des Hieronymus gegenüber dem Originaltext an.

33 Unterschiedliche Auffassungen über den Graubereich zwischen direktem Zitat und Referenzen allgemeinerer Art zeigen sich auch in folgendem Beispiel.: *nempe abruptis turbata procellis / nocte natat caeca serus freta, quem super ingens / porta tonat caeli, et* scopulis inlisa reclamant / aequora*; nec miseri possunt reuocare parentes, / nec moritura super crudeli funere uirgo* (Verg. georg. 3,259b–263 „Spät ja in finsterer Nacht noch durchschwimmt den / Sund er, den Stürme durchtosen, und über ihm donnert des Himmels / riesiges Tor, im Widerhall brüllt, an die Klippen gepeitscht, die / Brandung; zurückrufen können die armen Eltern ihn nicht, auch / nicht die Jungfrau, die – grausam ist’s! – sterben wird über der Leiche“) vs. *totam circa insulam fremit insanum mare et sinuosis montibus* inlisum scopulis aequor reclamat*; nullo terra gramine uiret; nullis uernans campus densatur umbraculis; abruptae rupes quasi quondam horroris carcerem claudunt.* (Hier. epist. 3,4,4 „Um die ganze Insel braust das tosende Meer, und die Flut, an die Klippen gepeitscht, hallt in den Felsbuchten wider. Die Erde wird durch kein Gras grün; ohne Schatten bleibt das keimende Feld undurchdringlich hart; schroffe Felsen umschließen gleichsam einen Kerker des Schreckens“). Diese Parallelstelle kennzeichnet Kamptner als „cf.“, Hagendahl stellt ohne nähere Erläuterung beide Stellen nebeneinander und Jeanjean markiert die Formanpassungen durch Kursivdruck, ohne allerdings den ansonsten genutzten Hinweis auf eine Allusion hinzuzusetzen. Im *manuellen Goldstandard* ist diese Referenz hingegen angesichts der Tatsache, dass es sich allein um Formanpassungen (*inlisa* vs. *inlisum*; *reclamant* vs. *reclamat*; *aequora* vs. *aequor*) und Umstellungen der Wortreihenfolge handelt, aber dieselben vier Lemmata zum Einsatz kommen, als Zitat geführt.

Hieronymus ruft hier bukolische Idylle auf und überträgt diese auf Bethlehem (*uinitor*, *arator*, *messor*, *decantare*, *canere*): Die Passage ruft damit klar die literarische Gattung und bestimmte Prätexte auf, zu denen selbstverständlich auch die *Bucolica* gehören.[34] Bei Vergil über eine ganze Reihe von Versen verteilt, erscheint das Wortmaterial, das sich beide Passagen aufgrund der ähnlichen Thematik teilen, bei Hieronymus gehäuft in nur einem Satz. Mag die Parallelstelle bei einer hermeneutischen Sichtung daher mit allgemeineren Begrifflichkeiten wie „*Georgica*-Referenz“[35] oder ähnlich umschrieben werden, ist sie in den *manuellen Goldstandard* dennoch als Zitat aufgenommen: Schließlich umfasst sie pro Satzeinheit mehr als zwei (sogar formgleiche) Wörter in direkter Nachbarschaft.

Gegenüber den vagen Einstufungstendenzen der hermeneutischen Forschung begünstigt der *manuelle Goldstandard* aufgrund des zugrunde liegenden Zitatbegriffs abermals eine stärkere Betonung des Zitatcharakters. Während dies einerseits begrüßenswert ist, rückt zugleich aber die im Forschungsdiskurs eröffnete Andersartigkeit dieser Referenz in den Hintergrund. Die gänzlich andere Qualität dieses Fundes im Vergleich beispielsweise zum zuvor diskutierten Fall *uox faucibus haesit* vs. *haesit uox faucibus* muss erst durch eine hermeneutische Betrachtung wieder eingebracht werden.[36] Mit Blick auf größere Corpora ist diese Art der Vereinheitlichung und Vereinfachung aber eine, die gewinnbringend zu sein verspricht: In den meisten Fällen sind längere direkte Zitate von den Forschenden bereits manuell identifiziert worden. Das erwartete Potenzial des durch digitale Prozesse möglichen *distant reading* liegt gerade im Auffinden von Passagen wie dieser.

Beispiel III: Sprichwörtliches

Wenn schon im Fall von Gattungsreferenzen der Bezug auf einen eindeutig bestimmbaren Prätext diskutabel sein kann, so gilt dies in besonderem Maße auch für sprachliche Bilder, Sprichwörter und ähnliche in den alltäglichen Gebrauch übergegangene Formulierungen. In solchen Fällen ist es häufig schwierig, einen konkreten Urheber auszumachen und daraus ein Zitat abzuleiten. So übergeht beispielsweise Hagendahl die folgende Passage aus den Briefen des Hieronymus:[37]

> Nihil amantibus durum est, *nullus difficilis cupienti labor. respice, quanta Iacob pro Rachel pacta uxore sustineat. et seruiuit, inquit scriptura, Iacob pro Rachel annis septem. et erant in conspectu eius quasi pauci dies, quia amabat illam.*
> (Hier. epist. 22,40,1)

> „Nichts ist für einen Liebenden hart, keine Anstrengung ist schwierig für einen, der begehrt. Sieh, was Jakob auf sich nimmt für Rahel, die ihm zur Frau versprochen war. ‚Jakob diente also‘, so sagt die Schrift, ‚sieben Jahre um Rahel. Weil er sie liebte, kamen sie ihm wie wenige Tage vor‘.“

Adkin hingegen führt, Luebeck folgend, die hieronymianische Formulierung auf Ciceros *Orator* zurück, auch wenn sie von allgemeinem Charakter ist:[38]

> *magnum opus omnino et arduum, Brute, conamur, sed* nihil difficile amanti puto. *amo autem et semper amavi ingenium studia mores tuos.*
> (Cic. orat. 33)

34 Zur Diskussion dieser Parallelstelle vgl. Feichtinger (2020), 241.

35 Feichtinger (2020), 241.

36 Um die hermeneutischen Ergebnisse nicht aus dem Blick zu verlieren, enthält der *manuelle Goldstandard* eine Spalte mit Bemerkungen, in denen ggf. auf Forschungsdebatten verwiesen wird.

37 Dass Hagendahl die Passage nicht aufführt, ist vor allem vor dem Hintergrund bemerkenswert, dass bereits Luebeck sie mit Cic. orat. 33 in Verbindung brachte und Hagendahl eigentlich auf dem Werk Luebecks aufbaut.

38 Vgl. Luebeck (1872), 133; Adkin (1997a), 28.

> „Eine in jeder Hinsicht gewaltige, komplizierte Aufgabe versuchen wir da, Brutus! Aber nichts ist schwer für den, der liebt, glaube ich. Ich liebe aber – und habe auch immer geliebt – deine Begabung, deine Bestrebungen, deinen Charakter."

Mit dem Zitatbegriff, der dem *manuellen Goldstandard* zugrunde liegt, gilt die Referenz hingegen aufgrund der Wortübereinstimmungen sowie des semantischen Gehalts als Zitat. In diesem Fall forciert der *manuelle Goldstandard* also die Zitatpotenz der Referenz – dass es sich um eine Aussage von allgemeinem Charakter handelt, bleibt nebensächlich und wird lediglich mit Verweis auf die Forschungsliteratur vermerkt.[39] Auch in solchen Fällen besteht also eine Tendenz zur Inklusion verschiedener Intertextualitätsphänomene in einen einheitlichen Zitatbegriff, der für die mittels digitaler Instrumente generierten Neu-Zitate einen passenden Rahmen schafft.

Beispiel IV: ‚Window allusions'

Ein Beispiel ähnlicher Art, allerdings mit genanntem Urheber des Sprichworts, ist folgendes:

> *Omnium autem rerum nec aptius est quicquam ad opes tuendas ac tenendas quam diligi nec alienius quam timeri. praeclare enim Ennius* quem metuunt oderunt quem *quisque* odit perisse *expetit. multorum autem odiis nullas opes posse obsistere si antea fuit ignotum nuper est cognitum.*
> (Cic. off. 2,23)

> „Von allen diesen Gründen ist aber keiner mehr geeignet, Einfluss zu bekommen und zu erhalten, als die Zuneigung und keiner weniger als die Angst. Denn Ennius hat schon sehr treffend formuliert: ‚Wen man fürchtet, hasst man; wen man hasst, dem wünscht man den Tod.' Dass aber keine Macht dem Hass vieler Menschen widerstehen kann, hat man erst neulich erfahren, wenn es bisher noch unbekannt war."

> *antiqua sententia est:* quem metuit quis, odit; quem odit, perisse *cupit. unde et in nostris litteris, cum initia paruulorum in timore consistant, perfecta dilectio foras mittit timorem. non quaeris monachos tibi esse subiectos et ideo magis subiectos habes.*
> (Hier. epist. 82,3,2)

> „Ein alter Spruch besagt: Wen einer fürchtet, den hasst er; wen er hasst, dem wünscht er, dass er zugrunde geht. Daher auch in unseren Schriften: Da die Anfänge kleiner Dinge in Furcht bestehen, vertreibt vollkommene Zuneigung die Furcht. Du verlangst nicht, dass dir die Mönche untergeben sind, und deshalb hast Du mehr Untergebene."

In der Diktion des Hieronymus eine *antiqua sententia*, bei Cicero dagegen klar Ennius zugeschrieben, ist die sprichwörtliche Aussage von der Forschungsliteratur als Cicero-Zitat gelistet, wobei sowohl Luebeck als auch Hagendahl noch auf das Ennius-Fragment verweisen.

Die Frage nach dem Urheber einer solchen sprichwörtlichen Formulierung mag in der traditionell-hermeneutischen Forschung für die Zuschreibung einer Referenz von Bedeutung sein, ist für den *manuellen Goldstandard* und den dort angelegten Zitatbegriff aber zunächst irrelevant, da es im Rahmen des digitalen Analyseinstrumentariums allein um das Auffinden textueller Übereinstimmungen geht. Die Parallelstelle ist daher als Cicero-Zitat aufgeführt und wird somit unter den maschinell generierten Funden erwartet. Der *manuelle Goldstandard* simplifiziert die Einstufung also, da die Referenz schlicht unter der Menge der Zitate auftaucht. Die Offenheit der hermeneutischen Diskussion wird insofern eingebüßt, obschon der *manuelle Goldstandard* mittels Bemerkungen noch darauf hinweist.

39 Vgl. hierzu die Kommentarspalte der entsprechenden Zeile im *manuellen Goldstandard*.

Beispiel V: Quelltext

Bisweilen kommt die manuell-hermeneutische Forschung auch zu unterschiedlichen Urteilen über den Quelltext:[40]

> *Librum tuum, quem pro Theodosio principe prudenter ornateque conpositum transmisisti, libenter legi. et praecipue mihi in eo subdiuisio placuit;* cumque in primis partibus uincas alios, in paenultimis te ipsum superas. *sed et ipsum genus eloquii pressum est et nitidum et, cum Tulliana luceat puritate, crebrum est in sententiis.*
> (Hier. epist. 58,8,1)

> „Dein Buch, das du für den Kaiser Theodosius auf gelehrte und beredte Weise verfasst und mir zugesandt hast, habe ich mit großem Genuss gelesen. Vor allem hat mir daran dieses Kapitel gefallen; Während du in den ersten Abschnitten andere Schriftsteller überragst, übertriffst du in den vorletzten dich selbst. Aber auch deine Schreibweise ist knapp, klar und, obwohl sie in ciceronianischer Reinheit erstrahlt, reich an Gedanken."

Kamptner und Hagendahl führen Cic. de orat. 3,3 als Quelltext an:[41]

> *hic, ut saepe inter homines sapientissimos constare vidi, quamquam hoc Crasso, cum aliquid accuratius dixisset, semper fere contigisset, ut numquam dixisse melius putaretur, tamen omnium consensu* sic esse tum iudicatum, ceteros a Crasso semper omnis, illo autem die etiam ipsum a se <es>se superatum.
> (Cic. de orat. 3,3)

> „Crassus war es – und in dem Punkt sind sich, wie ich oft festgestellt habe, die klügsten Männer einig –, wenn er besonders große Sorgfalt auf eine Rede verwendete, zwar fast immer gelungen, dass man glaubte, er habe nie besser gesprochen; doch bei dieser Gelegenheit urteilten alle, Crassus habe immer alle anderen Redner übertroffen, an diesem Tag aber auch sich selbst."

Adkin lehnt eine Bezugnahme des Hieronymus auf *De oratore* hingegen ab und argumentiert für ein Selbstzitat aus dem Prolog für die Erläuterungen der Homilien des Origines zum Hohelied:[42] *Origenes, cum in cæteris libris* omnes vicerit, *in Cantico canticorum* ipse se vicit (hom. Orig. in cant. prol. [PL 23,1173] „Während Origenes in den übrigen Büchern andere übertraf, übertraf er im Hohelied sich selbst"). Diese Passage wiederum führt Adkin überzeugend auf Ciceros *Pro Marcello* zurück:[43]

> *ita quantum operibus tuis diuturnitas detrahet, tantum adferet laudibus.* et ceteros quidem omnis victores bellorum civilium iam antea aequitate et misericordia viceras: hodierno vero die te ipse vicisti. *Vereor ut hoc quod dicam perinde intellegi possit auditu atque ipse cogitans sentio: ipsam victoriam vicisse videris, cum ea quae illa erant adempta victis remisisti. nam cum ipsius victoriae iure omnes victi occidissemus, clementiae tuae iudicio conservati sumus. recte igitur unus invictus es a quo etiam ipsius victoriae condicio visque devicta est.*
> (Cic. Marcell. 12)

> „So wird der zunehmende zeitliche Abstand, was er deinen Werken nimmt, deinem Ruhme zuwachsen lassen. Und alle früheren Siege in Bürgerkriegen hattest du schon vorher durch deine

40 Weitere Beispiele, bei denen Uneinigkeit herrscht, sind Cic. Brut. 262/Selbstzitat–Hier. epist. 47,1,2; Cic. orat. 47/Cic. de orat. 1,202–Hier. epist. 52,8,1; Cic. orat. 33–Hier. epist. 22,40,1; Cic. orat. 70–Hier. epist. 127,7,3.

41 Hagendahl (1958), 191 und 286; Kamptner (1996), 537.

42 Vgl. Adkin (1997b). Diese Passage führt zwar auch Hagendahl (1974) 221 auf, allerdings wiederum als Bezugnahme auf Cic. de orat. 3,3.

43 Vgl. Adkin (1997b).

> Zurückhaltung und Mäßigung übertroffen – am heutigen Tage hingegen hast du über dich selber einen Sieg errungen. Vielleicht kann man aus den folgenden Worten nicht genau das heraushören, was ich mir vorstelle und denke: du hast offenbar den Sieg selbst besiegt, indem du den Besiegten, was schon dein war, zurückgabst. Denn nach dem Recht, das der Sieg gewährt, hätten wir, die Besiegten, allesamt sterben müssen; doch dein Gnadenurteil hat uns geschont. Es ist also recht, daß du allein unbesiegt geblieben bist: du hast ja sogar das Gesetz des Sieges und die schrankenlose Gewalt besiegt."

Wie im Fall der sprichwörtlichen Aussagen ist die Frage des eigentlichen Urhebers für den *manuellen Goldstandard* zweitrangig, beide Referenzen werden entsprechend dem angelegten Zitatbegriff als Zitate aufgeführt und daher auch im Output der digitalen Analyse erwartet. Erneut wird also die Zitatpotenz der beiden Referenzen verstärkt, während die eigentlich strittige Frage nach der Identifizierung des Quelltextes zunächst in den Hintergrund rückt. Für die hermeneutische Interpretation kann sie reaktiviert werden, da sie im *manuellen Goldstandard* als Zusatzinformation hinterlegt ist.

Beispiel VI: Umformulierung

Während die Wortreihenfolge für den angelegten Zitatbegriff zwar keine Rolle spielt, zeigt sich eine starke Abhängigkeit vom Grad der Umformulierungen. Dies sei im Folgenden an zwei Beispielen illustriert, von denen das erste im *manuellen Goldstandard* als Zitat geführt wird:

> *sed haec et vetera et a Graecis; Cato autem sic abiit e vita, ut causam moriendi nactum se esse gauderet. vetatvenim dominans ille in nobis deus iniussu hinc nos suo demigrare; cum vero causam iustam deus ipse dederit, ut tunc Socrati, nunc Catoni, saepe multis, ne ille me Dius Fidius vir sapiens laetus ex his tenebris in lucem illam excesserit, nec tamen ille vincla carceris ruperit – leges enim vetant –, sed tamquam a magistratu aut ab aliqua potestate legitima, sic a deo evocatus atque emissus exierit.* Tota enim philosophorum vita, ut ait idem, commentatio mortis est.
> (Cic. Tusc. 1,74–75)

> „Doch dies sind alte und griechische Dinge. Cato aber ist derart aus dem Leben abgeschieden, dass er sich freute, eine Ursache des Sterbens gefunden zu haben. Denn der in uns regierende Gott verbietet es, ohne seinen Befehl von hier aus wegzuwandern. Wenn aber die Gottheit selbst eine gerechte Ursache gegeben hat, wie damals dem Sokrates, später dem Cato, oftmals Vielen, dann wird, bei Gott, der Weise heiter aus dieser Finsternis in jenes Licht übergehen; er wird nicht die Ketten des Gefängnisses zerbrechen – denn dies verbieten die Gesetze –, sondern wie von der Behörde oder einer anderen gesetzlichen Macht, so wird er, von Gott aufgerufen und entlassen, aus dem Leben abscheiden. Das ganze Leben der Philosophen ist ja, wie derselbe Sokrates sagt, eine Bekümmerung um den Tod."

> Platonis sententia est omnem sapienti uitam meditationem esse mortis. *laudant hoc philosophi et in caelum ferunt, sed multo fortius apostolus: cotidie, inquit, morior per uestram gloriam. aliud est conari, aliud agere; aliud uiuere moriturum, aliud mori uicturum. ille moriturus ex gloria est; iste moritur semper ad gloriam.*
> (Hier. epist. 60,14,2)

> „Ein Spruch Platons besagt, dass das ganze Leben für einen Weisen eine stete Betrachtung des Todes sei. Dies loben die Philosophen und heben es zum Himmel, aber um vieles eindrucksvoller spricht der Apostel: ‚Täglich sterbe ich, so wahr ihr mein Ruhm seid'. Zu versuchen, ist das eine, zu tun aber etwas ganz anderes: zu leben, um zu sterben, ist das eine, zu sterben, um zu leben, das andere. Jener wird mit Ruhm sterben, dieser stirbt immer zum Ruhm hin."

Das Paar *tota enim philosophorum vita, ut ait idem, commentatio mortis est* und *Platonis sententia est omnem sapienti uitam meditationem esse mortis* drückt in beiden Fällen denselben, Sokrates zuge-

schriebenen Gedanken aus. Bei dieser indirekten Übernahme ist die Formulierung sehr ähnlich, aber nur ein einziges Wort ist tatsächlich formgleich in beiden Texten vorhanden (*mortis*). Eine Umformulierung (*ait idem* vs. *sententia est*) hat hier weitere Formanpassungen zur Folge (*vita* vs. *vitam*; *est* vs. *esse*); dazu werden Synonyma verwendet (*tota* vs. *omnem*; *philosophorum* vs. *sapienti*; *commentatio* vs. *meditationem*). Angesichts der drei von beiden Textstellen geteilten Lemmata (*mors*, *esse*, *vita*) wird die Referenz im *manuellen Goldstandard* als Zitat geführt.

Das folgende Beispiel hingegen wird im *manuellen Goldstandard* als *confer*-Referenz eingestuft:

> *atque isti quidem qui docent, cum causas in plura genera secuerunt, singulis generibus argumentorum copias suggerunt. quod etiam si ad instituendos adulescentulos magis aptum est ut simulac posita causa sit, habeant quo se referant, unde statim expedita possint argumenta depromere, tamen et* tardi ingenii est rivolos consectari, fontis rerum non videre*, et iam aetatis est ususque nostri a capite quod velimus arcessere et unde omnia manent videre.*
> (Cic. de orat. 2,117)

> „Und diese Pauker fügen, wenn sie die Rechtsfälle in mehrere Klassen aufgespalten haben, für die einzelnen Klassen eine Menge von Argumentationshilfen hinzu. Das ist wohl besser geeignet, um ganz junge Männer auszubilden, damit sie, sobald der Fall vorgelegt ist, etwas haben, nach dem sie sich richten und woher sie sofort fertige Argumente entnehmen können; dennoch verrät es Trägheit des Geistes, den Lauf kleiner Rinnsale zu verfolgen, die Quellen der Dinge aber nicht wahrzunehmen, und unser Alter und unsere Erfahrung fordern, dass wir das, was wir wollen, von seinem Ausgangspunkt herleiten und wahrnehmen, woraus alles hervorgeht."

> *unde necessitate conpulsus sum transferre libros, in quibus plus mali quam boni est, et hanc seruare mensuram, ut nec adderem quid nec demerem Graecamque fidem Latina integritate seruarem. quorum exemplaria a supra dicto fratre poteris mutuari, licet tibi Graeca sufficiant et* non debeas turbidos nostri ingenioli riuos quaerere, qui de ipsis fontibus bibis.
> (Hier. epist. 85,3,2)

> „Daher bin ich von der Notwendigkeit getrieben worden, die Bücher zu übertragen, in denen mehr Schlechtes als Gutes steht, und dieses Maß zu wahren, dass ich weder etwas hinzufüge noch wegnehme und den griechischen Glauben in unversehrter lateinischer Diktion bewahre. Deren Abschriften wirst du vom oben erwähnten Bruder entleihen können, wenngleich dir auch schon die griechischen genügen und du nicht die trüben Bäche unseres geringen Geistes suchen musst, der du direkt aus den Quellen trinkst."

Wechsel zum Diminutiv (*rivolos* vs. *rivos*, *ingenii* vs. *ingenioli*), syntaktische Umstellungen, die Verwendung einer unpersönlichen, allgemeingültigen Aussage in Bezug auf ein angesprochenes ‚Du', eine Verschiebung in der Genitivfunktion von *ingenii–ingenioli* und Veränderungen in der Wortwahl (*fontis videre* vs. *fontibus bibis*) machen den intertextuellen Bezug ‚unkenntlich' und erschweren zunächst die Detektion. Dabei gründet sich die Zitatfunktion insbesondere auf die Semantik, da lediglich *fons* und *non* in beiden Texten auftauchen, im Falle des Hieronymus aber zu weit auseinander stehen. Aus diesem Grund handelt es sich im Sinne des *manuellen Goldstandards* lediglich um eine *confer*-Referenz, die im digitalen Analyseprozess nicht gefunden werden muss.

Mit Blick auf das zweite Beispiel muss somit konstatiert werden, dass der angelegte Zitatbegriff und damit der *manuelle Goldstandard* hinter der Hermeneutik zurückbleiben. Während das erste Beispiel noch genügend Übereinstimmungen aufweist und damit als Zitat im *manuellen Goldstandard* geführt wird, fehlen dem zweiten diese, sodass es nicht mehr unter der Menge der Zitate auftaucht. Bis zu einem gewissen Grad an Umformulierung bleiben somit paraphrasierende Referenzen als Zitate erhalten und werden trotz ihrer Andersartigkeit nicht von langen direkten Zitaten unterschieden. Übersteigt die Umformulierung aber eine bestimmte Schwelle, wird das Zitat nur noch als *confer*-Referenz geführt. Bisweilen kommt es durch strenge Anwendung des Zitatbegriffs zu einer hermeneutisch kontraintuiti-

ven Bewertungsumkehr, infolge derer stark umformulierte Referenzen aufgrund der Übereinstimmung einer gewissen Anzahl von Lemmata als Zitate aufgenommen werden und umgekehrt weniger stark umformulierte Referenzen als *confer*-Referenzen unbeachtet gelassen werden. Um diese Problematik transparent zu halten, enthält unser *manueller Goldstandard* eine Kommentarspalte, in der auf den Grad der Umformulierung hingewiesen werden kann.

Beispiel VII: ‚Totalausfall'

Eine gänzlich verkürzte Art der Bezugnahme stellt folgende Formulierung des Hieronymus dar:

> *hoc enim et Tullius tuus et ante illum in breui et solo uolumine scribit Antonius, primam causae esse uictoriam diligenter causam, pro qua dicturus es, discere.*
> (Hier. epist. 49,1,1)

> „Denn auch dein Landsmann Cicero schrieb und schon vor ihm Antonius in einem kurzen und einzelnen Buch, der erste Teilsieg einer Rechtssache bestünde schon darin, den Fall, für den man sprechen werde, sorgfältig zu studieren."

Reduziert auf die bloße Erwähnung (*Tullius tuus*) wird auf die antike Quelle für das Wissen über das Rhetorikbüchlein des Antonius verwiesen.[44] Die Markierung ist hier knapp gehalten, der Bezug ist vielmehr inhaltlich und lässt sich auf der Wortebene selbst kaum festmachen. Mithin wird die Parallele im *manuellen Goldstandard* als *confer*-Referenz ausgewiesen. Als solche wird sie trotz der expliziten Markierung durch *Tullius tuus* nicht unter den maschinell generierten Funden erwartet.

Da dies selbstverständlich höchst unbefriedigend ist, weil eine deutlich markierte und damit klarerweise intendierte Bezugnahme vorliegt, wird als Konsequenz das digitale Analyseinstrumentarium von uns um einen Verfahrensschritt ergänzt, der den Zieltext – in diesem Fall also die Briefe des Hieronymus – nach Nennungen eines Prätextes oder Autors durchsucht. Die Arbeit mit dem hermeneutischen Zitatpool ist in diesem Fall doppelt effizient, weil sie eine Ergänzung des Analyseinstrumentariums initiiert, die über die Übereinstimmung von Wortmaterial hinausgeht.[45]

Die Beispiele zeigen, dass die Anwendung einer binären Zitatklassifikation des *manuellen Goldstandards*, die für die Operationalisierung in einem digitalen Analyseinstrument notwendig ist, sowohl zu Informationsgewinn als auch zu Informationsverlust führt. Während im Beispiel I die eindeutige Zitatdefinition des *manuellen Goldstandards* die etwas zögerliche Diskussion der manuell-hermeneutischen Forschung, ob es sich nun um ein Zitat handelt oder nicht, aufzulösen vermag und den Zitatcharakter unterstreicht, geht in den Beispielen II, III, IV und V trotz der teilweisen Betonung des Zitatcharakters auch stets eine Nuance der genauen Beschaffenheit der Referenz, insbesondere hinsichtlich des Prätextes, verloren. Im Fall von Beispiel VI stößt der *manuelle Goldstandard* an seine Grenzen, weil umfassende Umformulierungen, also die meisten der in der manuell-hermeneutischen Forschung mit „cf." markierten Zitate, dem zugrundeliegenden Zitatbegriff nicht mehr entsprechen. Da solche *confer*-Referenzen dennoch im *manuellen Goldstandard* vermerkt sind, dient dieser als umfassende Zusam-

44 Cic. de orat. 1,206: *„quid? si" inquit Crassus „quoniam ego, quo facilius vos apud me tenerem, vestrae potius obsecutus sum voluntati quam aut consuetudini aut naturae meae, petimus ab Antonio, ut ea quae continet neque adhuc protulit, ex quibus unum libellum sibi excidisse iam dudum questus est, explicet nobis et illa dicendi mysteria enuntiet?"* „„Um euch leichter bei mir zu behalten', erwiderte Crassus, ‚habe ich mich schon eher nach eurem Wunsch als nach meiner Gewohnheit und Natur gerichtet. Wie wäre es da, wenn wir den Antonius bäten, er möge uns das, was er für sich behält und bis jetzt nicht ans Tageslicht gebracht hat und von dem ihm, wie er schon längst geklagt hat, ein kleines Buch entschlüpft ist, erklären und diese Geheimnisse der Redekunst verkünden?'"

45 Einen Ansatz hierzu stellt der *Auctoritas*-Filter in Schropp et al. (2024) dar, der in den Verlauf des Textabgleichs eingebunden ist und auf diese Art sogleich den konkreten Quelltext zu detektieren vermag, allerdings leicht als eigenständiger Prozess herausgelöst werden kann.

menschau aller in der Forschungsliteratur angeführten Zitate für die hermeneutische Forschung als ein kompaktes Nachschlagewerk.[46]

Den (verhältnismäßig geringen) qualitativen Einbußen stehen die deutlichen Vorteile des Methodenwechsels ins Digitale in Hinblick auf die quantitativen Möglichkeiten gegenüber: Sobald anhand des *manuellen Goldstandards* einmal eine Grundlage für ein solches Zitatdetektionsinstrumentarium gelegt wurde, können innerhalb kurzer Zeit viele intertextuelle Referenzen in sehr großen Textmengen gefunden werden, wie es durch einen rein hermeneutischen Vergleich nicht möglich wäre.

Und selbst wenn dabei von der digitalen Suchroutine unerwartete Ergebnisse produziert werden, da diese eben nicht mit dem philologischen Blick von Forschenden an die Texte herangeht, sollte dies nicht vorschnell als Mangel oder Nachteil abgeurteilt werden. Gerade diese Irritationen können höchst produktiv dazu herausfordern, das eigene Verständnis von Intertextualität stets neu zu reflektieren.

Fazit

Für die Intertextualitätsforschung stellt der *manuelle Goldstandard* ein Hilfsmittel dar, das einen bestehenden Forschungsstand für digitale Routinen nutzbar macht und damit als Referenzinstrument für Kalibrierungsprozesse dienen kann. Das Abbilden des Forschungsstandes geht dabei, wie gezeigt worden ist, über eine reine Sammeltätigkeit hinaus: Stets ist ein kritisches Prüfen der teils widersprüchlichen, teils uneinheitlichen, teils zu knappen Forschungsliteratur vonnöten.

So kann die Auseinandersetzung mit den Problemlagen auch zu einer Reflexion innerhalb der traditionell-hermeneutischen Forschung anregen, da zentrale Begriffe der traditionell-hermeneutischen Forschung im Kontext digital gestützter Methoden einem neuen kritischen Blick unterzogen werden und dabei weiter geschärft werden können.

Auch über diese Reflexion hinaus liegt für die traditionell-hermeneutische Forschung ein Gewinn in der Erstellung des *Goldstandards*: Er bündelt *sämtliche* bisher von der traditionell-hermeneutischen Forschung aufgezeigten intertextuellen Referenzen eines bestimmten Autoren-/Werke-Paares und stellt damit ein nützliches Nachschlagewerk für weiterführende Analysen und Interpretationen zur Klassikerrezeption dar. Letztlich wird sich der *Goldstandard* durch das kontinuierliche Einpflegen von (digitalen) Neufunden zu einem dynamischen Gesamtzitatverzeichnis entwickeln, welches gerade aufgrund der entwickelten VIATE-Prinzipien den Qualitätsanforderungen der Forschung entspricht.

46 Die komprimierte Darstellung vermag zudem Differenzen zwischen einzelnen Autoren (etwa in der Klassikerrezeption) in besonderer Deutlichkeit aufzuzeigen. Dies sei beispielhaft knapp anhand der Referenzen in den *Epistulae* des Hieronymus auf Vergil bzw. Cicero illustriert. Hieronymus zitiert Vergil häufiger explizit und über eine längere Strecke hinweg als Cicero. Obwohl der Dichter selten genannt ist, sind die Referenzen eindeutig identifizierbar und in der Forschung weitgehend unbestritten. Die Bezüge auf Cicero sind hingegen meist verknappt oder paraphrasierend. Häufig handelt es sich dabei um sprichwörtliche Aussagen oder Übernahmen aus eigentlich anderen Quelltexten. Der Grad der Abwandlung in der Formulierung ist bei Cicero durchschnittlich höher als bei Vergil, von dem oft ganze Verse übernommen werden. Ob dies nun allgemeine Merkmale von Bezügen auf Poesie oder Prosa sind oder nicht, lässt sich allein anhand Vergil und Cicero nicht beantworten. Es scheint naheliegend, dass Poesie durch die Metrik und damit einen vertrauten Rhythmus sowie die dichterische Ausdrucksweise bei der Einbettung in einen Prosatext einen höheren Wiedererkennungswert haben könnte als die ebenfalls prosaischen Texte Ciceros. Erhärten lässt sich eine solche Vermutung jedoch erst, wenn sich ähnliche Beobachtungen auch bei anderen Autoren und Quelltexten bestätigen.

Quellenverzeichnis

Textausgaben / Übersetzungen

Clark (1983): A. C. Clark, M. Tulli Ciceronis orationes II, Oxford 1983.

Conte (2019): G. B. Conte (Hrsg.), P. Vergilius Maro, Aeneis. Editio Altera, Berlin / Boston 2019.

Dückers (2009): P. Dückers (Hrsg.), Ambrosius, De virginibus. Über die Jungfrauen. Übersetzt und eingeleitet von P. Drückers, Fontes Christiani 81, Turnhout 2009.

Fuhrmann (1997): M. Fuhrmann (Hrsg.), Marcus Tullius Cicero, Die Prozessreden. Band II. Lateinisch – deutsch, Zürich / Düsseldorf 1997.

Gigon (1998): O. Gigon (Hrsg.), Marcus Tullius Cicero, Tusculanae disputationes. Gespräche in Tusculum. Lateinisch – deutsch, Düsseldorf / Zürich 1998.

Hilberg (1996): I. Hilberg (Hrsg.), S. Eusebii Hieronymi Epistulae, 3 Bde., Wien 1996.

Höhmann (2002): B. Höhmann (Hrsg.), Der Amos-Kommentar des Eusebius Hieronymus. Einleitung, Text, Übersetzung, Kommentar, Münster 2002.

Holzberg (2015): N. Holzberg (Hrsg.), Publius Vergilius Maro, Aeneis. Lateinisch – deutsch, Berlin / Boston 2015.

Holzberg (2016): N. Holzberg (Hrsg.), Publius Vergilius Maro, Bucolica Hirtengedichte / Georgica Landwirtschaft. Lateinisch – deutsch, Berlin / Boston 2016.

Kumaniecki (1969): K. Kumaniecki (Hrsg.), M. Tullius Cicero, De oratore, Leipzig 1969.

Kytzler (1998): B. Kytzler (Hrsg.), Marcus Tullius Cicero, Orator. Lateinisch – deutsch, Düsseldorf / Zürich 1998.

Migne (1883): J.-P. Migne (Hrsg.), Sancti Eusebii Hieronymi Stridonensis presbyteri opera omnia. Tomus 2, PL Series Latina 23/2, Paris 1883.

Nickel (2008): R. Nickel (Hrsg.), Marcus Tullius Cicero, De officiis. Vom pflichtgemäßen Handeln. Lateinisch – deutsch, Düsseldorf 2008.

Nüßlein (2007): T. Nüßlein (Hrsg.), Marcus Tullius Cicero, De oratore. Über den Redner. Lateinisch – deutsch, Düsseldorf 2007.

Ottaviano / Conte (2013): S. Ottaviano / G. B. Conte (Hrsg.), P. Vergilius Maro, Bucolica et Georgica, Berlin / Boston 2013.

Pohlenz (1982): M. Pohlenz (Hrsg.), M. Tullius Cicero, Tusculanae disputationes, Stuttgart 1982.

Winterbottom (1994): M. Winterbottom (Hrsg.), M. Tulli Ciceronis. De officiis, Oxford 1994.

Westman (1980): R. Westman (Hrsg.), M. Tullius Cicero. Orator, Leipzig 1980.

Literaturverzeichnis

Abbott (1979): K. M. Abbott, Satira and Satiricus in Late Latin, Illinois Classical Studies 4 (1979), 192–199.

Adkin (1992a): N. Adkin, Hieronymus Ciceronianus: the Catilinarians in Jerome, Latomus 51 (1992), 408–420.

Adkin (1992b): N. Adkin, Some features of Jerome's compositional technique in the Libellus de virginitate servanda (Epist. 22), Philologus 136 (1992), 234–255.

Adkin (1992c): N. Adkin, Istae Sunt, Quae Solent Dicere: Three Roman Vignettes in Jerome's Libellus De Virginitate Servanda (Epist. 22), Museum Helveticum: Schweizerische Zeitschrift für Klassische Altertumswissenschaft = Revue Suisse pour l'Étude de l'Antiquité Classique 49 (1992), 131–140.

Adkin (1993): N. Adkin, Jerome, Epist. 22,28,5: „a Fat Old Man", Bollettino dei Classici / Accademia Nazionale dei Lincei 14 (1993), 142–149.

Adkin (1997a): N. Adkin, Cicero's „Orator" and Jerome, VChr 51/1 (1997), 25–39.

Adkin (1997b): N. Adkin, Cicero, Pro Marcello 12 and Jerome, Philologus 141/1 (1997), 137–140.

Adkin (2000): N. Adkin, Two further echoes of Sallust's „Histories" in Jerome („Vita Hilarionis" 22,3 and 30,2), VetChr 37 (2000), 209–215.

Adkin (2006): N. Adkin, Hieronymus eunuchinus, GIF 58/2 (2006), 327–334.

Adkin (2011): N. Adkin, Some alleged echoes of Apuleius in Jerome, CPh 106/1 (2011), 66–75.

Adkin (2013): N. Adkin, Cicero's „Pro Milone" and Jerome, Euphrosyne 41 (2013), 367–374.

Adkin (2018): N. Adkin, Horace, *Carm*. 2.17.5 and Quintilian, Inst. 6 Prooem. in Jerome, Prometheus 7 (2018), 202–208.

Adkin (2019): N. Adkin, Cicero's Pro Sexto Roscio and Jerome, MH 76/1 (2019), 88–95.

Antin (1970): P. Antin, Jérôme antique et chrétien, REAug 16 (1970), 35–46.

Ascone (2019): L. L. Ascone, Origenes fortissime respondit: Gerolamo, Paolino di Nola e Origene in Hier. Epist. 85, RFIC 147/2 (2019), 417–440.

Audano (2019): S. Audano, Pisonianum vitium: intertestualità ciceroniana e polemica teologica in Girolamo (Ep. 69,2), Sileno 45/1–2 (2019), 1–15.

Banniard (1988): M. Banniard, Jérôme et l'elegantia d'après le De optimo genere interpretandi, in: Y.-M. Duval (Hrsg.), Jérôme entre l'occident et l'orient. XVIe centenaire du départ de saint Jérôme de Rome et de son installation a Bethléem. Actes du Colloque de Chantilly (Septembre 1986), Paris 1988, 305–322.

Bauer (1975): J. B. Bauer, Hieronymus und Ovid, Grazer Beiträge: Zeitschrift für die klassische Altertumswissenschaft 4 (1975), 13–19.

Bird et al. (2010): S. Bird / E. Klein / E. Loper, Natural language processing with Python. Analyzing text with the natural language toolkit (E-Book), Beijing 2010: https://www.nltk.org/book/ (abgerufen am 08.04.2024).

Capone (2018): A. Capone, Stoici, qui in nostro dogmati in plerisque concordant: Gerolamo e lo stoicism, Adamantius 24 (2018), 435–450.

Clark (2006): E. Clark, Dissuading from Marriage. Jerome and the Asceticization of satire, in: W. Smith (Hrsg.), Satiric advice on women and marriage. From Plautus to Chaucer, Ann Arbor (Mich.) 2006, 154–181.

Coffee (2018): N. Coffee, An Agenda for the Study of Intertextuality, TAPA 148/1 (2018), 205–223: https://doi.org/10.1353/apa.2018.0008 (abgerufen am 08.04.2024).

Courcelle (1976): P. Courcelle, Les lecteurs de l'Énéide devant les grandes invasions germaniques, Romanobarbarica 1 (1976), 25–56.

Feichtinger (1997): B. Feichtinger, Nec vero sopor ille fuerat aut vania somnia… (Hier. ep. 22,30,6). Überlegungen zum geträumten Selbst des Hieronymus, RÉAug 43 (1997), 41–61.

Feichtinger (2020): B. Feichtinger, Bukolisches Idyll in Betlehem: Zur kulturellen Hybridität von Hieronymus Epistula 46, Thersites 11 (2020), 218–246.

Feichtinger (2021): B. Feichtinger, Quid facit cum psalterio Horatius? (Hier. ep. 22,29,7): Untersuchung zu Hieronymus' Umgang mit klassischen und biblischen Referenzen am Beispiel von Epistula 3 ad Rufinum, VChr 75/4 (2021), 389–454.

forTEXT (2016): forTEXT, Goldstandard: https://fortext.net/ueber-fortext/glossar/goldstandard (abgerufen am 08.04.2024), veröffentlicht unter der Creative Commons Lizenz CC BY-SA 3.0 DE Deed.

Hagendahl (1958): H. Hagendahl, Latin fathers and the classics. A study on the apologists, Jerome and other Christian writers, Studia Graeca et Latina Gothoburgensia 6, Göteborg 1958.

Hagendahl (1967): H. Hagendahl, Augustine and the Latin classics, Studia Graeca et Latina Gothoburgensia 20, 2 Bde., Göteborg 1967.

Hagendahl (1974): H. Hagendahl, Jerome and the Latin Classics, VChr 28/3 (1974), 216–227.

Hamblenne (1996): P. Hamblenne, L'exemplum formel dans l'oeuvre conservé de Jérôme, Augustinianum 36/1 (1996), 93–146.

Iannacone (1963): S. Iannaccone, Sull'uso Degli Scritti Filosofici Di Cicerone Da Parte Di S. Girolamo, Giornale Italiano di Filologia, 17 (1964), 329–341.

Jaureguizar (1970): E. Jaureguizar, San Jerónimo Y Los Clásicos, Perficit: revista de estudios humanísticos 2/39–40 (1970), 458–477.

Jeanjean (2004): B. Jeanjean, „Quand il ne reste plus que le droit de gémir". Jérôme lecteur de Cicéron et de Sénèque le Père, in: B. Gain et al. (Hrsg.), „Chartae caritatis". Études de patristique et d'antiquité tardive en hommage à Yves-Marie Duval, Paris 2004, 385–399.

Jeanjean (2015): B. Jeanjean, Les métamorphoses de la figure de Didon chez saint Jérôme: quand l'emprunt aux auteurs latins profanes se fait detournement, in: A. Canellis et al. (Hrsg.): Caritatis scripta: mélanges de littérature et de patristique offerts à Patrick Laurence, Paris 2015, 112–124.

Jeanjean (2018): B. Jeanjean, Que fait Horace à côté du psautier? Virgile à côté des Evangiles? (Hieron. epist. 22,29): les citations poétiques profanes dans les Lettres de saint Jérôme, in: G. Müller (Hrsg.), Zwischen Alltagskommunikation und literarischer Identitätsbildung: Studien zur lateinischen Epistolographie in Spätantike und Frühmittelalter, Roma Aeterna 7, Stuttgart 2018, 169–199.

Kamptner (1996): M. Kamptner, Sancti Eusebii Hieronymi Epistulae. Pars IV: Epistularum Indices et Addenda, Wien 1996.

Laurence (2004): P. Laurence, Lettres sur la mort d'une fille. Seruius Sulpicius Rufus et saint Jérôme, in: L. Nadjo / É. Gavoille (Hrsg.), Epistulae antiquae. 3, Actes du IIIe colloque international „L'épistolaire antique et ses prolongements européens" (Université François-Rabelais, Tours, 25–27 septembre 2002), Paris 2004, 345–364.

Layton (2002): R. Layton, Plagiarism and lay patronage of ascetic scholarship. Jerome, Ambrose, and Rufinus, JECS 10/4 (2002), 489–522.

Luebeck (1872): E. Luebeck, Hieronymus quos nouerit scriptores et ex quibus hauserit, Leipzig 1872.

MacCormack (1998): S. MacCormack, The shadows of poetry: Vergil in the mind of Augustine, Transformation of the Classical Heritage 26, Berkeley 1998.

Marolla (2017): G. Marolla, Presenza di classici in Girolamo, epist. 123 a Geruchia, VetChr 54 (2017), 127–141.

McGill (2007): S. McGill, Virgil, Christianity, and the „Cento Probae", in: J. Scourfield (Hrsg.), Texts and culture in Late antiquity. Inheritance, authority, and change, Swansea 2007, 173–194.

Murphy (2007): K. P. Murphy, Performance evaluation of binary classifiers: https://www.cs.ubc.ca/~murphyk/Teaching/CS340-Fall07/reading/rocHandout.pdf (abgerufen am 08.04.2024).

Neuhausen (1984): K. A. A. Neuhausen, Hieronymus, Seneca und Theophrasts Schrift über die Freundschaft, in: E. Dassmann / K. Thraede (Hrsg.), Vivarium. Festschrift Theodor Klauser zum 90. Geburtstag, Münster 1984, 257–286.

O'Donnell (1980): J. J. O'Donnell, Augustine's classical readings, RÉAug 15 (1980), 144–175.

OED: Oxford English Dictionary, s.v. gold standard, n., sense 1: https://doi.org/10.1093/OED/6267693828 (abgerufen am 08.04.2024).

OED: Oxford English Dictionary, s.v. gold standard, n., sense 3: https://doi.org/10.1093/OED/1022694340 (abgerufen am 08.04.2024).

Penzenstadler (2006): F. Penzenstadler, s.v. Imitatio, in: H. Cancik / H. Schneider / M. Landfester (Hrsg.), Der Neue Pauly: http://dx.doi.org/10.1163/1574-9347_dnp_e1406130 (abgerufen am 08.04.2024).

Rebenich (1999): S. Rebenich, Rez. von Hieronymus: Briefe. Hrsg. von I. Hilberg. 3 Bde. Wien: Österreichische Akademie der Wissenschaften 1996, Gymnasium 106/1 (1999), 75–78.

Revellio (2022): M. Revellio, Zitate der Aeneis in den Briefen des Hieronymus. Eine digitale Intertextualitätsanalyse zur Untersuchung kultureller Transformationsprozesse, Berlin / Boston 2022: https://doi.org/10.1515/9783110760828 (abgerufen am 08.04.2024).

Poinsotte (1992): J-M. Poinsotte, Hieronymus Poeta (Ep. 108, § 33), in: J. Fontaine/ L. Holtz (Hrsg.), De Tertullien Aux Mozarabes: Mélanges Offerts À Jacques Fontaine, Membre De L'institut, À L'occasion De Son 70e Anniversaire Par Ses Élèves, Amis Et Collègues, I: Antiquité Tardive Et Christianisme Ancien: Iiie–Vie Siècles, Paris 1992, 211–221.

Préaux (1958): J. Préaux, Procédés d'invention d'un sobriquet par saint Jérôme, Latomus 17/4 (1958), 659–664.

Ronnenberg (2015): K. Ronnenberg, Mythos bei Hieronymus. Zur christlichen Transformation paganer Erzählungen in der Spätantike, Hermes Einzelschriften 108, Stuttgart 2015.

Rzepkowski (2011): K. Rzepkowski, Jerome, „palliatae", and „veteres comici". A note to Hier. Ep. 57.5.5 („De optimo genere interpretandi"), Palamedes 6 (2011), 109–113.

Schneider (2003): C. Schneider, Lactance, Jérôme et les recueils de déclamations pseudo-quintiliens, in: J.-Y. Guillaumin / S. Ratti (Hrsg.), Autour de Lactance. Hommages à Pierre Monat, Paris 2003, 63–76.

Schropp et al. (2024): F. Schropp / T. E. Konrad / M. Revellio / B. Feichtinger, Towards a differentiated digital-hermeneutic analysis tool for the detection of ‚short' quotations using the example of the Church Father Jerome, Digital Humanities Quarterly 18/3 (2024).

Trillitzsch (1965): W. Trillitzsch, Hieronymus und Seneca, Mittellateinisches Jahrbuch 2 (1965), 42–54.

Tsartsidis (2020): T. Tsartsidis, Jerome, Ep. 53.7 and the centonist Proba, CQ 70/1 (2020), 453–458.

Viscido (1980): L. Viscido, Appunti Sull'invettiva in Girolamo, Vigiliae Christianae: A Review of Early Christian Life and Language 34 (1980), 386–394.

Zimbrich (2006): U. Zimbrich, s. v. Mimesis, in: H. Cancik / H. Schneider / M. Landfester (Hrsg.), Der Neue Pauly: http://dx.doi.org/10.1163/1574-9347_dnp_e804920 (abgerufen am 08.04.2024).

Autoreninformation[47]

Franziska Schropp
Universität Konstanz
Universitätsstraße 10
78464 Konstanz
E-Mail: franziska.schropp@uni-konstanz.de

Michael Wittweiler
Universität Konstanz
Universitätsstraße 10
78464 Konstanz
E-Mail: michael.wittweiler@uni-konstanz.de

Thomas E. Konrad
Universität Konstanz
Universitätsstraße 10
78464 Konstanz
E-Mail: thomas.eugen.konrad@uni-konstanz.de

Dr. Marie Revellio
Universität Konstanz
Universitätsstraße 10
78464 Konstanz
E-Mail: marie.revellio@uni-konstanz.de

Prof. Dr. Barbara Feichtinger
Universität Konstanz
Universitätsstraße 10
78464 Konstanz
E-Mail: barbara.feichtinger@uni-konstanz.de

47 Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autor:innen. Alle Inhalte dieses Beitrages unterstehen, soweit nicht anders gekennzeichnet, der Lizenz CC BY 4.0.