The Datacons Project: An Open-access Dataset of Late Roman Consular Dates

MARCO DOSI

ABSTRACT

The DataCons Project provides an open-access dataset of late Roman consular dating formulae spanning from CE 284 to 541. It aggregates consular materials from publications worldwide. Currently, the dataset contains over 4,800 documents, penned in three distinct scripts, originating from ten ancient regions, and categorised by material type and textual content. The present release of the dataset (2.0.0) focuses on the Latin and Greek documentation from CE 476 to 526, includes nearly 880 items in total, and pertains exclusively to papyri and inscriptions. Periodic updates are anticipated to expand its temporal coverage. The dataset is tailored to support interdisciplinary research involving consular material, and the project will soon be accompanied by the launch of an online relational database.

CORRESPONDING AUTHOR:
Marco Dosi
Department of History, King’s College London, GB; Digital Humanities Hub, University of London School of Advanced Study, GB
marco.dosi@kcl.ac.uk

KEYWORDS:
History; Political History; Cultural History; Databases; Epigraphy; Papyrology

TO CITE THIS ARTICLE:
**OVERVIEW**

**REPOSITORY LOCATION**

https://doi.org/10.5281/zenodo.10251711

DOI: 10.5281/zenodo.10251711

**CONTEXT**

This paper presents an open-access dataset comprising nearly 880 late Roman consular dating formulae drawn exclusively from papyri and inscriptions, and ranging from CE 476 to 526. Consular dating was widely used in the Roman world for centuries and applied to various ancient artefacts, including papyri, inscriptions, and other archaeological and literary materials. Its study remains crucial for modern scholars, enabling them to accurately date texts and extract vital information about the life, politics, religion, ideology, institutions, and functioning of the Roman Empire and its successor states (Bagnall, Cameron, Schwartz, & Worp 1987; Bagnall & Worp, 2004; Burgess, 1989; Dosi, 2022; Gonis, 2005a, 2005b; Henning, 1999; Meimaris, Kritikakou, & Bougia, 1992; Salway, 2008; Sguaitamatti, 2012).

The dataset presented here is a refined subset of a more extensive collection gathered in the context of the PhD project “Consular Dating and Consular Dissemination in Late Antiquity” (Dosi, 2022) based at King’s College London and funded by the AHRC (LAHP). This dataset now forms the foundation of the larger DataCons Project, which seeks to create a comprehensive, searchable open-access database of late Roman consular documents from CE 284 to 541. The imminent release of this database is aimed at bolstering research in late Roman history. Although the database will not initially support data downloads or extraction, the project ensures accessibility to its valuable dataset through the present release and future updates.

**METHODOLOGY**

**2.1. STEPS**

1. **Data Sourcing and Compilation:** The dataset is primarily sourced from a variety of major scholarly references, academic journals, epigraphic and papyrological corpora. Initial data points derive from the latest revision of Egyptian papyri in Bagnall and Worp (2004) and the epigraphic sources from Bagnall et al. (1987) and Meimaris et al. (1992). The corpus was updated through March 2022 with new discoveries, especially from AE, SEG, P.Oxy, ZPE, ICI and Marin, Gauthier, and Prévot (2010). Additionally, the dataset integrates substantial data from key online resources and databases.

---

1. Année épigraphique, Paris, 1889-.
2. Supplementum Epigraphicum Graecum. Leiden, 1923-.
3. The Oxyrhynchus Papyri, London, 1988-.
4. Zeitschrift für Papyrologie und Epigraphik, Köln, 1968-.
5. Inscriptiones Christianae Italieae VII s. antiquiores (nova serie). Bari, 1985-.
6. For an overview of the sources and their bibliographic abbreviations, see Dosi, 2022, pp. 11-14, 495, 500-505 and the accompanying documentation on Zenodo.
2. **Data Entry and Structuring**: Data were manually entered into an Excel document, organised in tabular form, and sorted by key attributes. This process included integrating hyperlinks to other databases’ records and images.

3. **Data Normalisation and Enhancement**: Basic Excel operations were performed to normalise data, generate unique identifiers for each row, and separate the GPS coordinates into latitude and longitude columns. Texts are presented in original and restored forms following the Leiden convention, supplemented with photographs of the original texts or scholarly reconstructions and transcriptions.

4. **Documentation and Metadata Management**: Although the dataset does not include a critical apparatus for the dating of each document – which will be addressed in future publications – the accompanying documentation on Zenodo provides an overview of the rationale behind the dating of each document (when it differs from previous scholarly opinions), as well as a description of each column.

2.2. **SAMPLING STRATEGY**

The sampling objective was to comprehensively gather all material from CE 476–526. Tokens that were securely dated within this timeframe received priority during the collection process, ensuring a focus on information pertinent to each attribute for this specific period.

2.3. **QUALITY CONTROL**

The palaeography of each formula was confirmed using original documents, photographs, or scholarly editions to ensure accuracy. Attribute information underwent multiple rounds of verification by the author to maintain reliability. Additionally, findspot identification and GPS coordinates were validated using authoritative sources such as the Barrington Atlas (Talbert & Bagnall, 2000), Trismegistos Places (Verreth, 2008) and/or Pleiades (Elliott, 2021) ensuring geographical precision.

(3) **DATASET DESCRIPTION**

In version 2.0.0, the dataset comprises 39 data attributes, represented as columns. These include: a unique identifier; Possible Year(s); Year Assigned; Year Assigned (m.l.d. e.app.); Class; Carrier; Language; Type; Text; Evidence (i.e., where the evidence has been sourced, with the main bibliographic notes); Formula (full titulary); Formula (Full titulary, simplified); Formula (simplified); Day; Month; Indiction; Other supporting dating elements; Errors; Latitude; Longitude; Place; Region; Macro-Region, Image, and fifteen additional columns linking to external databases.10

The dataset comprises virtually all published material from CE 476 to 526, including 614 stone and 264 papyrus formulae. These records range from single to multi-year dates, like 461 or 482, or 4th to 6th c.

Consular dating typically involved two types of formulae: post-consular and consular formulae. Post-consular formulae, rendered as post consulatum + N. et N. (gen.) in Latin and μετὰ τὴν ὑπατείαν + N. καὶ N. (gen.) in Greek, dated events relative to the most recent consulship known at the place of writing. In contrast, consular formulae, which referenced the names of the consuls in office at the time of writing, were normally expressed as N. (et) N. consulibus or consulatu + N. (et) N. (gen.) in Latin, and as ὑπατείας + N. καὶ N. (gen.) or (ἐν) ὑπατείᾳ + N. καὶ N. in Greek. Most texts primarily provide the dating protocol, namely the consular or post-consular date, day and month designations, and occasionally other dating systems like the indication, and the place of writing.11

---

10 For a detailed description of each column and the rationale behind it, please refer to the accompanying README file on Zenodo.

In detail, over 500 dates are consular, nearly 200 are post-consular, and the remainder are of uncertain type. They are mainly in Greek and Latin, with a few bilingual scripts. These formulae date six groups of texts: 1) funerary; 2) administrative; 3) legal/fiscal; 4) monumental; 5) military; 6) ‘other’ (mostly private), and 7) uncertain.

Maps and graphs presented in Figures 1, 2, 3, 4, 5 to 6 illustrate that the distribution of these texts varies significantly by region, time, language, and type, reflecting the cultural complexities and varying conditions of preservation. Notably, although the dataset covers 10 ancient regions and over 190 findspots, more than 50% of the evidence is from the Italian diocese, with 40% specifically from Rome.  

Finally, an analysis of 726 texts (about 80% of the total) reveals that over 70% of this sample shows abbreviations or variations in the titulary, misspellings, and other non-linguistic irregularities (i.e., omitted, wrong or conflicting dating elements within the formula).

12 Cf. the README file on Zenodo for a detailed explanation of the regional classification.
Figure 3 Top 10 most represented findspots and typological complexity of their datasets. Colour shows details about text type.

Figure 4 Regional, linguistic and quantitative complexity of the evidence. Colour shows details about Language (formula). The evidence sourced from Hispaniae, North Africa, and Britannia is exclusively in Latin.

Figure 5 Chronological distribution by decade. Colour shows details about macro-region.
3.1. OBJECT NAME
DataCons Dataset version 2.0.0

3.2. FORMAT NAMES AND VERSIONS
CSV UTF-8

3.3. CREATION DATES
Start date (v. 1.0.0): 2017-10-1; End date (v. 2.0.0): 2023-12-2

3.4. DATASET CREATORS
Marco Dosi

3.5. LANGUAGE
Ancient Greek, Latin, English

3.6. LICENSE
CC BY-SA Int. 4.0

3.7. REPOSITORY NAME
Zenodo

3.8. PUBLICATION DATE
2023-12-02

(4) REUSE POTENTIAL
The dataset offers an in-depth exploration of the late Roman world, with a particular focus on the consulship as both an institution and a dating system. Scholars using this resource can delve into a wide range of themes, including the socio-cultural and political roles of the consulship, its historical significance and decline, and its varied adoption as a dating system across societal levels and regions. The dataset facilitates detailed analysis of movement dynamics, spatial and geographical considerations, and the perception and management of time in the late Roman period. It also sheds light on aspects of late Roman communications and bureaucratic operations, contributing to a comprehensive understanding of these complex topics. Additionally, the dataset is instrumental in examining the geopolitical dimensions of consular dissemination and their implications for West-East relations during periods like CE 476–526.
By centralising material formerly dispersed across numerous publications, the dataset supports both quantitative and qualitative research, aligning with the digital trend in epigraphy, papyrology and related fields like palaeography and linguistics (Berti, 2019; Bodel, 2012; Liuzzo, 2019; Stokes, 2009, 2015; Xella & Zamora, 2018). Integration with platforms such as Trismegistos (Depauw & Gheldof, 2014), Papy.info, Pleiades (Elliott, 2021), and similar projects will further enrich existing resources and deepen our understanding of specific texts, periods and regions. The ability to closely examine the full text of each document via hyperlinks facilitates a detailed analysis of linguistic shifts across different sources.

(5) LIMITATIONS

The current dataset acknowledges certain areas in need of improvement. One primary shortcoming is the omission of detailed discussions on the dating of the formulae. This is obviously a major limitation, but at this stage the dataset aims to function as a checklist more than a critical edition of the entire corpus. Additionally, while the dataset offers valuable insights into the demographics of these regions and periods, it currently lacks a detailed, structured categorisation of the individuals referenced.

As mentioned, the dataset provides only the original text of the dating protocol rather than the full text of each document, and these texts have not been fully encoded. To mitigate these constraints, the dataset includes hyperlinks to external databases where complete and encoded texts are available.

I am committed to continually enhancing the dataset and addressing these shortcomings in future updates. In the meantime, users should remain cognisant of these limitations, which might restrict the scope of their analysis.

ACKNOWLEDGEMENTS

I would like to express my profound gratitude to Peter Heather for his steadfast support in the DataCons Project, and to Nikolaos Gonis for his invaluable insights and discussions on consular papyri. My thanks also extend to the JOHD reviewers for their significant contributions to enhancing the first draft of this article. Additionally, I am deeply thankful to the scholars and researchers whose comprehensive surveys and publications have laid the groundwork for this dataset. The project would have faced considerable delays without their meticulous and pioneering work, which has been instrumental in both the completion and publication of this dataset.

FUNDING INFORMATION

AHRC London Arts and Humanities Partnership Research Studentship (1883425).

COMPETING INTERESTS

The author has served as peer-reviewer for the journal but has not taken part in the peer reviewing process of this article.

AUTHOR AFFILIATIONS

Marco Dosi

Department of History, King’s College London, GB; Digital Humanities Hub, University of London School of Advanced Study, GB
