# KTÈMA

## CIVILISATIONS DE L'ORIENT, DE LA GRÈCE ET DE ROME ANTIQUES


### La rhétorique de la diplomatie en Grèce ancienne

Cinzia Bearzot, Laura Loddo — Introduction ..... 5
Laura Loddo — Political Exiles and Their Use of Diplomacy in Classical Greece ..... 7
Cinzia Bearzot — À propos du parallélisme entre deux discours d'ambassade à Sparte (Xén. *Hell.* V, 2, 12-19 et VI, 1, 4-16) ..... 23
Paolo A. Tuci — The Speeches of Theban Ambassadors in Greek Literature (404-362 B.C.) ..... 33
Francesca Gazzano — Greek Ambassadors and the Rhetoric of Supplication. Some Notes ..... 53
Nicholas D. Cross — The (Im)balance of Power. Demosthenes' Complex Case for an Alliance with the Megalopolitans ..... 71
Davide Amendola — *Presbeutikoi* and *Enteuktikoi Logoi* in Hellenistic Interstate Relations. Some Further Thoughts from an Epigraphical Perspective (c. 306-205 B.C.) ..... 87

### Sicile hellénistique et romaine

Jonathan R. W. Prag — *I.Sicily*, Open Scholarship, and the Epigraphic Landscape of Hellenistic/Roman Sicily ..... 107
Lorenzo Campagna — Trasformazioni urbanistiche in Sicilia alle origini della *Provincia*. Riflessioni sul ruolo di Roma ..... 123
Cristina Soraci — Cultes et politique dans la Sicile du Ier siècle av. J.-C. Le cas de la Vénus Érycine et de la Cérès d'Henna ..... 145

### *Varia*

Edith Foster — Athens' Political Failures in the Central Chapters of Book 4 of Thucydides ..... 163
Walter Lapini — Note interpretative e testuali alla col. XXII del Papiro di Derveni ..... 171
Sylvain Perrot — La place de la musique dans la politique culturelle de Téos dans la première moitié du IIe siècle avant notre ère ..... 179
Gianpaolo Urso — Catilina *legatus*. Considerazioni su un discusso frammento sallustiano ..... 197
Anne Jacquemin — La dédicace aux Muses d'Eurydice fille de Sirrhas (Plutarque, *Sur l'éducation des enfants*, 20) ..... 209



**N° 44** **STRASBOURG** **2019**

PRESSES UNIVERSITAIRES DE STRASBOURG

# *I.Sicily*, Open Scholarship, and the Epigraphic Landscape of Hellenistic/Roman Sicily

Résumé–. Cet article présente l'évolution du corpus numérique *I.Sicily* des inscriptions de la Sicile antique dans le contexte des débats actuels sur l'érudition en libre accès. L'évolution du projet *I.Sicily* est située dans le cadre des publications d'épigraphie sicilienne et dans celui des pratiques de publication en humanités. La manière dont le projet *I.Sicily* aborde les pratiques liées aux données ouvertes et les défis que posent celles-ci est examinée à travers deux exemples : le premier offre un bref aperçu de la culture épigraphique sicilienne dans l'Antiquité, le second considère les possibilités de recherche collaborative et progressive offertes par de tels ensembles de données et corpus.

Mots-clés–. Sicile antique, épigraphie, érudition en accès libre, humanités numériques

Abstract–. This paper presents the development of the *I.Sicily* digital corpus of the inscriptions of ancient Sicily within the context of current debates about open scholarship. The evolution of the *I.Sicily* project is situated within the framework of Sicilian epigraphic publication and in turn within the framework of humanities publication practices. The approach of the *I.Sicily* project to open data practices, and the challenges which these pose, is examined through two case studies: the first offering a brief survey of Sicilian epigraphic culture in antiquity; the second considering the possibilities such datasets and corpora offer for collaborative and progressive research.

Keywords–. Ancient Sicily, epigraphy, open scholarship, digital humanities

This paper is intended as a reflection upon the current progress of epigraphic study of Hellenistic/Roman Sicily, within the context of wider developments in digital scholarship.[1] Section I considers the challenges in studying Sicilian epigraphy and the role of open data in overcoming such challenges. Section II summarises the development of the *I.Sicily* digital corpus of Sicilian inscriptions as a response to these challenges. Section III then considers the merits and challenges in this context of open data, open access publication, collaborative study and the free availability of data for re-use. Given the author's directorship of the *I.Sicily* project, the argument presented here is clearly open to accusations of self-interest; it is, however, assumed that the overall, shared scholarly ambition of advancing our understanding of ancient Sicily and its epigraphic culture will take priority in the new framework of open scholarship.

(1) The paper began life as a presentation at a workshop on Hellenistic Sicily, in Paris, April 2018, at the invitation of Prof. Sylvie Pittia. As such, it offered an overview of Sicilian epigraphic culture (see §3.1 below), and I am grateful to Prof. Pittia for the encouragement to use it now to explore the issues addressed here.

## I. SICILIAN EPIGRAPHY AND THE CHALLENGE OF OPEN DATA

A familiar challenge presented by the Sicilian epigraphic evidence is that of using the epigraphic evidence as a primary source for the writing of history, especially in the absence of parallel literary sources.[2] Any such efforts are dependent upon the accessibility of material and the existence of, or facility to undertake works of synthesis. Until recently the existing *corpora* for the island were extremely dated, and consequently omitted huge amounts of material and more recent scholarship.[3] At the same time, it remains the case that not all the major (let alone many minor) collections of inscriptions on the island have been properly studied, catalogued and published (the most obvious omission, currently the focus of study, remains the collection of the Museo P.Orsi, Siracusa). The impact of this for those working, for example, on the historical linguistics of the island is severe, and the responses adopted in such scholarship exemplify the situation: on the one hand, such analyses 'are destined to remain little more than hypotheses, until a comprehensive list of all epigraphic texts from ancient Sicily is assembled'; on the other, a scholar may well attempt to build their own working list in order to meet this lack.[4] Such a list, however, while representing a very substantial undertaking, in reality remains of extremely limited use and accessibility to future scholars and while serving to underpin the specific analysis and conclusions of the work at hand, all too often is essentially 'lost' to future scholarship.[5] Arguably even more problematic, from a more strictly scientific perspective, such work is effectively uncheckable because it cannot be easily reproduced, since the underlying data cannot be examined without first attempting to repeat the original highly laborious task of data gathering. To quote from a recent paper on the underlying principles:[6]

> […] there is growing recognition that data availability is a critical feature of an efficient, progressive, and ultimately self-correcting scientific ecosystem that generates credible findings. A minimum level of credibility we would expect of all published findings is that they can be reproduced when the reported analyses are repeated upon the raw data: a concept referred to as computational or analytic reproducibility. However, assessment of analytic reproducibility is typically precluded by poor access to research data. Furthermore, even when data are shared, inadequate documentation and formatting can render them unusable.

This may be a debate that is currently taking place primarily in the domain of the sciences, but in practice almost every principle enunciated in the preceding sentences should be equally applicable to humanities scholarship dependent upon a dataset such as a material or textual corpus. As humanities scholarship moves into the digital age, this debate becomes ever more pertinent.

The present author can hardly claim to be innocent in this discussion. My own first attempts to sketch the epigraphic landscape of Hellenistic and Roman Sicily were themselves egregious examples of precisely this.[7] Furthermore, having built up the dataset to support such work, I subsequently maintained that (private) dataset in order to build a unique position within the research community, publishing further studies which exploited that data.[8] There is nothing

(2) See Manganaro 1988 and Prag 2007 for two rather different approaches to the problem.

(3) De Vido 1999 provides a survey.

(4) Tribulato 2012: 324; Mimbrera 2012: 24-44 §2.3.

(5) In the specific case of Mimbrera 2012, the hard-copy only publication, the organisation of the list, and the format of referencing would make it extremely onerous indeed to attempt to re-use the data; the fundamental purpose of the publication of the list, while itself commendable, therefore becomes that of demonstrating that the scholar did the work thoroughly (it is hardly unique in this respect).

(6) Hardwicke *et al* 2018: §1.

(7) Principally, Prag 2002 (listing the sources for the data underlying various quantitative assessments, but not the source data itself) and Prag 2007 (similar, only referencing individual examples within discussion).

(8) *E.g.* Prag 2008, 2010.

remarkable in that, *per se*, and it reflects widespread, indeed standard, practice in much of the humanities. But there are two very obvious flaws in such an approach – the first has already been outlined; the second is that such data is almost always effectively 'lost' when the researcher moves on, and the work becomes more of a dead-end than a stepping stone. The motivation for this paper, however, is a radical change in my own position, and the establishment of a new open access, online corpus for Sicilian epigraphy – the *I.Sicily* project.[9]

## II. *I.SICILY*, DIGITAL EPIGRAPHY, AND THE SICILIAN EPIGRAPHIC LANDSCAPE

In order to produce a study of the epigraphic culture of ancient Sicily, given the limitations of existing *corpora*, it was necessary to construct a new database of Sicilian inscriptions. This was originally undertaken between 2001 and 2004 in a privately maintained MS Access database, and subsequently updated (albeit irregularly). The data collected for this database was entirely 'metadata' (information about the inscribed object and its publication history), not the epigraphic texts themselves. The rationale for doing so was to create a consistent point of reference for the highly dispersed and repeatedly republished Sicilian epigraphic corpus, while systematically gathering particular categories of data in a standardised form in individual data fields for future analysis. At the time, there were three reasons for excluding the texts: textual analysis was not the focus of the initial study; databases did not handle multiple ancient languages and alphabets very easily; and the investment of time and resource was unrealistic.

The compilation of the original database therefore facilitated the necessary degree of certainty about the reliability of the data (i.e. lack of duplication of texts), and provided the data necessary for a first set of studies on the epigraphic habit of the island. At this point it became possible to generate basic data on the overall numbers, geographical, chronological, linguistic, typological and material distribution of the stone inscriptions of the island in antiquity.[10] However, the published discussion, which briefly summarised the basic methods and listed the primary published sources of the underlying data, focused upon a series of data visualisations (maps and charts) and did not make the raw data available. (A revised version of the same analysis was submitted as part of a doctoral thesis, but at the time there was no requirement for the accompanying data to be submitted or retained for the thesis either.[11]) Consequently, readers had to take the analysis (or at least the security of its foundations) on trust. However, there is a more fundamental point to consider, which is that, theoretically at least, such a dataset has considerable further potential for future analysis and expansion,[12] but any such action remains completely dependent upon the whims of the original researcher (I shall return to this point below). At the same time, the analytical potential of such a dataset was limited by the lack of the epigraphic text (and imagery also, but that remains outside this particular discussion).

Over time, it became clear that this database offered the basis for the creation of a comprehensive new corpus.[13] Plain text databases in individual languages, such as the PHI database of Greek inscriptions have been steadily evolving since the later 20th century, and these have led scholars

(9) The project data is at <http://sicily.classics.ox.ac.uk/> (accessed 03.02.2019); documentation on the project can be found at < https://isicily.org/> (accessed 26.05.2019).

(10) Prag 2002.

(11) Prag 2004.

(12) A minor example would be the specific analysis first attempted for the Hellenistic/Republican period in Prag 2003.

(13) A full discussion of the creation of *I.Sicily*, including more technical information can be found in Prag, Chartrand, and Cummings 2017; Prag and Chartrand 2018.

and developers to confront the challenges of how to render complex epigraphic texts, in diverse languages and alphabets, in machine readable (and so readily amenable to search and analysis) digital format.[14] The crucial shift in potential, however, came with the development of open source standards in text-encoding (TEI EpiDoc in XML), which not only enabled the encoding of information about the text itself, but also of supporting information or metadata, i.e. information about the text-bearing object, its discovery and subsequent study, etc.[15] One further development deserving of note is the Trismegistos project, which aims, among other things, to provide a published unique identifier for every epigraphic document from the ancient Mediterranean, which constitutes an important additional step towards resolution of the problem of duplication and individuation of texts with a complex publication history.[16]

These developments, together, mean that it is now possible to do several things in one place, which previously had been dispersed, and to do so at a level that is entirely comparable to any standard edition of an individual inscription (and so to talk about corpora and editions, and not simply 'databases').[17] *I.Sicily* now contains (or will contain) all the information for individual inscriptions, from text and material description to commentary and bibliography, that might be expected in any traditional paper edition. The difference is that the data within that edition is increasingly standardised and encoded in ways that make it machine readable and so available for analysis, thereby enabling analysis across the entire corpus in a multitude of different ways. Furthermore, all the data now housed within the corpus is made openly and freely available, and can readily be downloaded (in this case, under a CC-BY 4.0 license, which permits free copying, redistribution, remixing, transformation, or further development; but requires the user to give appropriate credit, provide a link to the license, and indicate if changes were made).[18]

However, the adoption of new models of digital publication poses some fundamental challenges. Perhaps the most basic of these is a simple question of acceptance and citation. At present, scholars of classical epigraphy freely and frequently use the various existing online text databases, and increasingly use the reference numbers of, e.g., *EDCS*, in order to reference an inscription (although doing so begs some serious questions, since the underlying source of the online text is often very far from clear).[19] At the same time, many scholars remain very unclear how to cite digital publications,[20] and more understandably remain cautious about both the quality and the longevity of digital publications. A book is assumed to stand on a library shelf for many years; but will a digital resource remain stable and accessible in comparable fashion? The concerns regarding quality are arguably irrational, given that a great many paper publications are not in fact subject to any meaningful form of peer review (and the concept of open review has yet to take hold in much of the humanities); at the same time, the failure of the discipline, by and large, to embrace digital publications and, e.g., subject them to formal review, perpetuates the underlying problem of perception and acceptance.[21] Arguably even more challenging is the idea that a digital publication,

(14) A rapid summary and introduction with reference to Roman epigraphy in Elliott 2014. *PHI* is online at <https://inscriptions.packhum.org/> (accessed 03.02.2019).

(15) For EpiDoc, see Bodard 2010 and Elliott, Bodard, Cayless *et al.* 2006-2017.

(16) See <https://www.trismegistos.org/index.html> (accessed 31.01.2019), under 'texts'.

(17) The pioneering example of such a digital corpus is *IAph2007* (Reynolds, Roueché, Bodard 2007).

(18) See <https://creativecommons.org/licenses/by/4.0/> (accessed 31.01.2019).

(19) *EDCS* = Clauss, M., Kolb, A., Slaby, W. A. and Woitas, B., *Epigraphik-Datenbank Clauss / Slaby*, available at <http://www.manfredclauss.de/> (accessed 31.01.2019).

(20) See, *e.g.* Ball and Duke 2015.

(21) For high quality review of digital editions, see, e.g., <https://ride.i-d-e.de/> (accessed 31.01.2019); the recent Michigan digital publication of excavation at Gabii (Opitz, Mogetta and Terrenato 2016) is one of the very first actually to receive formal reviews in traditional paper journals (e.g. Poehler 2018).

such as the edition of an individual inscription, has the potential to be repeatedly and frequently updated and emended (and, optimistically, improved), which poses even greater challenges to questions of use, citation, and preservation, if the version cited by a user is no longer identical with the version presented online. It is not within the scope of this paper to address all of these concerns, but they are real, and there is little evidence that many current digital projects have given them sufficient thought (in the sense that there is very little explicit reference to any of these issues in the, frequently very limited, documentation accompanying many digital resources and projects).

## III. NEW HORIZONS IN SICILIAN EPIGRAPHIC STUDIES? OR REPEATING OLD MISTAKES?

By itself, however, does the potentiality of such a digital corpus change anything? In the remainder of this paper, using *I.Sicily* as a case-study, I discuss two distinct examples of, on the one hand (§3.1), what might be considered 'progress' in a move towards open data and open scholarship; and on the other (§3.2) a situation where the full possibilities of a truly *collaborative* open data model remain to be realised by the wider research community.

### *1. Mapping the Epigraphic Habit*

As already noted, an early product of the work leading to the *I.Sicily* project was a study of the epigraphic habit in ancient Sicily, which followed traditional publication practices, referencing a privately held dataset, the basis for which was only briefly described within the article, and the data for which was not made public.[22] As part of the ongoing work associated with *I.Sicily*, elements of the original analysis published in 2002 were recently repeated, employing new analyses of the revised dataset as it was in 2016; discussion and visualisation of these analyses was subsequently published in 2018.[23]

A significant number of analyses are enabled by the uniform collation of this sort of metadata. It becomes possible, for example, to plot the chronological and geographical distribution of lapidary epigraphy on the island of Sicily over time (Figs. 1-2). Sicilian lapidary epigraphy has a distinct archaic phase, consisting primarily of Greek and Punic material; during the Hellenistic period, Greek lapidary epigraphy rapidly increases and dominates; while under the High Empire, Latin epigraphy appears and predominates, as is typical across the Roman empire. One interesting feature of Sicilian epigraphy is the continuity and persistence of the use of Greek for lapidary epigraphy into Late Antiquity. The distribution of Sicilian lapidary epigraphy is, unsurprisingly, a fundamentally urban phenomenon, with a particularly strong concentration in the major coastal centres, and from the first century BC onwards there is a clear correlation with centres that received colonial or municipal status under Rome. It is interesting to note that the quantity of data employed in these assessments is approximately double that available for the original 2002 study (due both to more comprehensive data recording and to improvement of the original data), and yet the overall patterns remain consistent (the principal variation is at the level of individual settlements, such as Halaesa, where the number of published inscriptions has increased, from approximately 15 to nearly 60, but that new data follows the trends already identified).[24]

With improved data, it is possible significantly to refine this sort of analysis. By comparing chronological and geographical distributions between the Archaic/Classical and Hellenistic

(22) PRAG 2002, cf. PRAG 2004.
(23) PRAG 2018.
(24) For the epigraphy of Halaesa see PRESTIANNI 2012 and especially PRAG and TIGANO 2017.

periods, and then considering some of the ways in which the material in those periods is distributed typologically and by material, it is possible to reveal several patterns which align very clearly with the broader archaeological trends that have emerged for the later Hellenistic period in Sicily, and which in turn contribute to discussions about civic life in Sicily under the Roman Republic and the nature/extent of cultural phenomena such as 'Romanisation'.[25] In the first place, it is clear that epigraphic culture expands from being a feature of the south and east of the island in the earlier period (the location of the main Greek colonies, together with the Phoenician colony of Mozia), to an island-wide practice over the Hellenistic period (Figs. 3-4). This pattern reflects, firstly the urban foundations of the fourth century BC, and secondly the increasing urban monumentalisation of the later Hellenistic period on the island.[26] In the second place, a shift in practice is observable, from primarily funerary and votive epigraphy in the Archaic/Classical periods, to a much more diverse range of practice in the Hellenistic period, including a wide variety of public epigraphy (Figs. 5-6), which is matched by an increasing monumentalisation of epigraphy in its own right (e.g. the cadastral text from Halaesa and the account inscriptions from Tauromenion, as well as the magistrate lists and other texts from Akrai).[27] This latter set of developments can most visibly be illustrated in the increased use of marble for epigraphy on the island (Figs. 7-8). All marble on Sicily is necessarily imported, which therefore entails both a choice and a cost.[28] The use of marble for inscriptions in the Archaic period is extremely rare.[29] It is much more extensive in the Hellenistic period, and what is particularly striking is that marble is used almost entirely for 'public' epigraphy.[30] All of these patterns very clearly illustrate that the use of public writing played an integral part in the public, civic culture of the island in the later Hellenistic period, and makes a more detailed analysis of the content of such a material textual culture highly desirable.

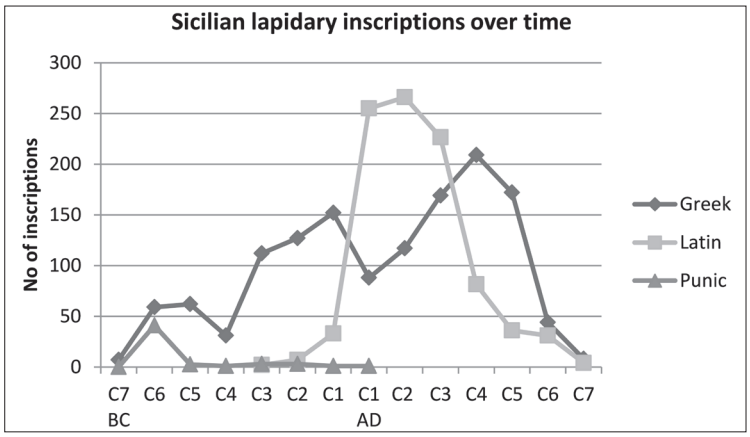



Fig. 1: Frequency of Greek, Latin, and Punic lapidary epigraphy on Sicily over time (other minor languages are omitted). Data from *I.Sicily*, as of 22.6.2016, *n*=2352 dated inscriptions out of a total of 3231 recorded stones.

(25) These analyses are the focus of Prag 2018, where more detailed discussion can be found.

(26) See *e.g.* Campagna 2011 and Prag 2014 for these developments.

(27) Halaesa (*ISic*1174, 3651); Tauromenion (*ISic*1247-1255, 2985-2987, 3081, 3082); Akrai (*e.g. ISic*1028-1035, 1038, 2965): *ISic* references can be resolved in the form <http://sicily.classics.ox.ac.uk/inscription/ISic1174> (accessed 03.02.2019), and full bibliography is available there.

(28) Wilson 1990: 237-42 on sources and use of stone in ancient Sicily.

(29) Five texts are known to me: *ISic*1566 (Selinus); *ISic*3148 (Gela); *ISic*1425 (Naxos); *ISic*1048 (Akrai); *ISic*1435 (Megara Hyblaea, a *kouros* statue).

(30) 37 instances, used almost entirely for decrees, public lists, honorifics and dedications. For references, see the dataset online at <https://doi.org/10.5287/bodleian:YeRab9boO> (accessed 26.05.2019).




Fig. 2: Distribution of all lapidary epigraphy on Sicily over time between the 7th cent. BC and the 7th cent. AD. Data from *I.Sicily*, as of 22.6.2016, $n$ = 3180 located inscriptions out of a total of 3231 recorded stones. (799 inscriptions on Lipara, and 24 on other surrounding islands not displayed)

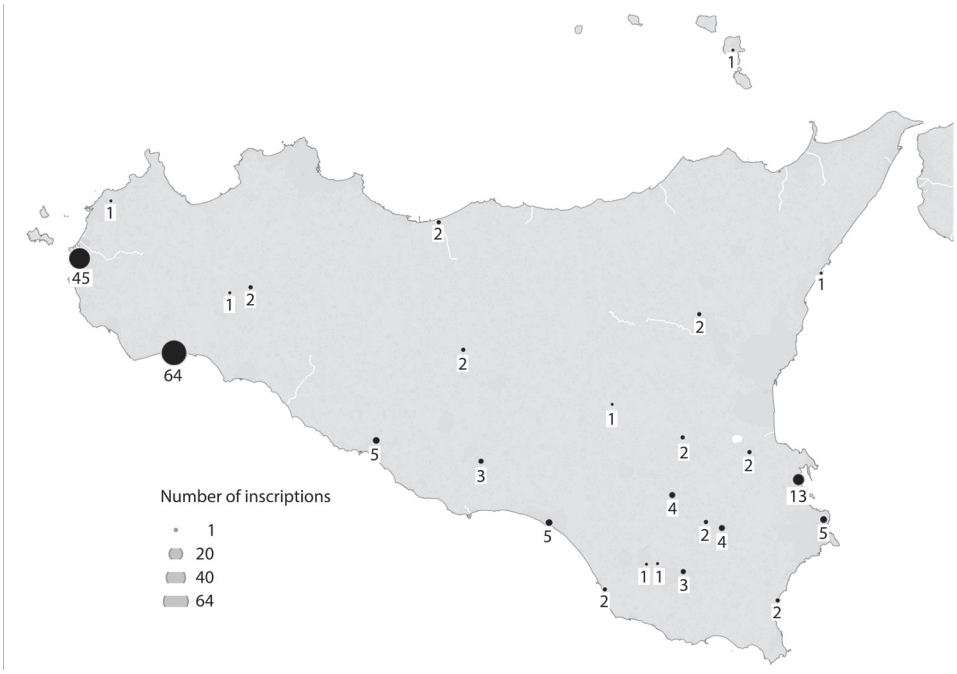



Fig. 3: Distribution of lapidary inscriptions across Sicily in all languages (VII - V cent. BC). Data from *I.Sicily*, as of 22.6.2016; *n* = 179. Map prepared using Tableau Public 10.1, with a map-base from OpenStreetMap.

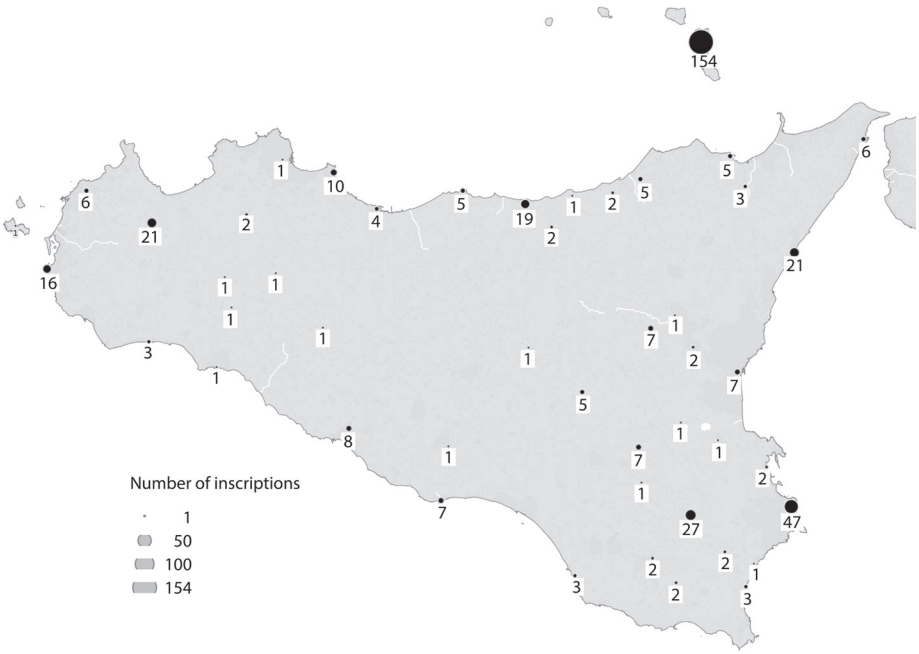



Fig. 4: Distribution of lapidary inscriptions across Sicily in all languages (IV – I cent. BC). Data from *I.Sicily*, as of 22.6.2016; *n* = 433 (a further 2 from Melita are not shown). Map prepared using Tableau Public 10.1, with a map-base from OpenStreetMap.

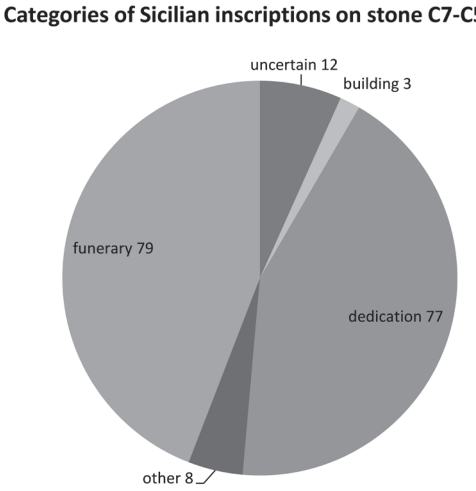



Fig. 5: Types of inscription on stone in Sicily (VII-V cent. BC). Data from *I.Sicily*, as of 22.6.2016; *n* = 179.

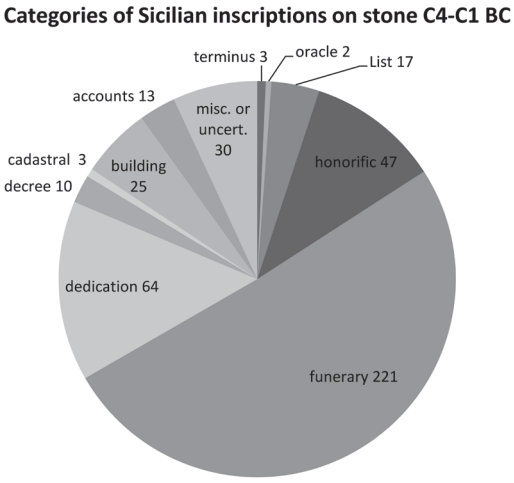



Fig. 6: Types of inscription on stone in Sicily (IV-I cent. BC). Data from *I.Sicily*, as of 22.6.2016; *n* = 435.

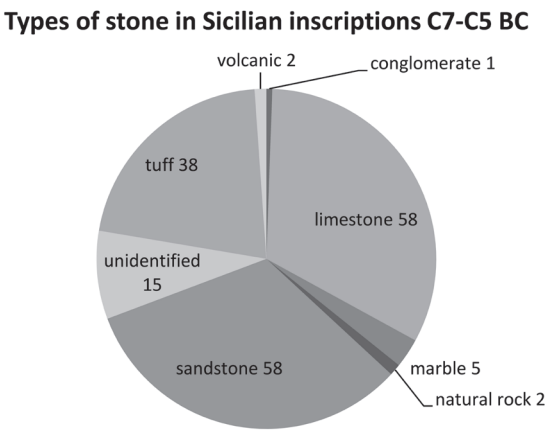



Fig. 7: Types of stone used for inscriptions in Sicily (VII-V cent. BC). Data from *I.Sicily*, as of 22.6.2016; *n* = 179.

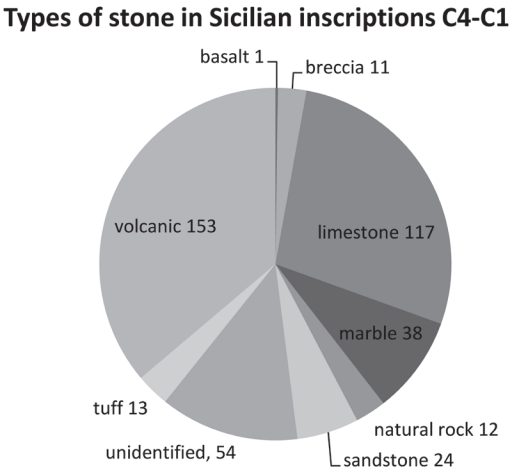



Fig. 8: Types of stone used for inscriptions in Sicily (IV-I cent. BC). Data from *I.Sicily*, as of 22.6.2016; *n* = 423.

As noted above, and illustrated by the foregoing analysis, one significant motivation for the digital collection of inscriptions and metadata is that it provides the basis for data analysis. It is, nonetheless, striking that although an increasing number of projects are creating such data, surprisingly few of them make the data available (or at least available in a format that can relatively easily be exploited). Data of this sort could not readily (if at all) be extracted, for example, from the PHI database of Greek inscriptions, nor from the Clauss-Slaby database of Latin inscriptions, nor from the Epigraphic Database Roma, all of which simply present results as a continuous page of HTML, and sometimes do not even explicitly address the question of data use.[31] At the same time, despite the fact that a growing number of such resources do make such data available (most obviously the Epigraphic Database Heidelberg; it is also an ambition of the EAGLE project), the number of analyses attempted by those from outside the individual projects remains very few indeed.[32]

The final point for discussion within this section, however, resumes the opening contrast with the study of 2002, which, typically for such a study, merely described the collection of the data, but did not preserve or publish the underlying dataset. Consequently, the 2002 study is essentially unreproducible and untestable. The 2018 study differs in two fundamental respects. In the first place, the study was conducted using data extracted from an open access online resource.[33] In principle, this means that any individual is free to conduct the same analysis at any time, and the availability of data in CSV format and the existence of a growing number of reasonably intuitive and open source data analysis and visualisation tools means that only a very limited technical ability and resource is required to undertake such analysis.[34] However, there is one significant flaw in this assumption, which reflects a major challenge in the new publication models emerging in the digital environment. If one were to filter and download the equivalent data sets from *I.Sicily* today, one would observe that both the total number of files has increased since 2016 and the content of individual files has been refined. Theoretically, because *I.Sicily* employs a GitHub repository for the management of the editorial workflow of the EpiDoc files, before they are uploaded to the server which supports the website, it is possible to go back to the state of the files at a particular date.[35] However, this only provides access to the c.3,000 EpiDoc files, without the web interface, and consequently, a rather higher level of technical ability would be needed to extract the equivalent data directly from the EpiDoc XML files. This situation reflects a much more general problem for the publication and citation of digital resources that are undergoing continuous editing and revision, unlike traditional hard-copy publication. In the case of *I.Sicily*, the use of GitHub provides a partial solution, but the reality is that without a systematic and rigorous process of publicly archiving each file in an accessible and referenceable format on every occasion that it is revised, the problem persists, both at the level of the dataset as a whole and at the level of individual inscriptions

(31) For *PHI* and *EDCS* see nn.12 and 17 above; *EDR* is at <http://www.edr-edr.it/default/index.php> (accessed 03.02.2019).

(32) A rare example noted by this author is McCleery 2016, with thoughtful observations on the challenges at 43 with nn.5-7 and *passim*. For *EDH* see < https://edh-www.adw.uni-heidelberg.de/data> (accessed 02.02.2019); *EAGLE* aims to make searches available in CSV and PDF to registered users < https://www.eagle-network.eu> (accessed 02.02.2019); note the review by Charles Hedrick, 14.10.2018, at <https://classicalstudies.org/node/29654> (accessed 26.05.2019).

(33) A summary of the possible ways of extracting data from *I.Sicily* is outlined at <https://isicily.org/how-to/#opendata> (accessed 02.02.2019). *I.Sicily* went live on 1 January 2017.

(34) For the 2018 publication, the mapping was done using Tableau Public <https://public.tableau.com/> (accessed 03.02.2019), other data analysis using MS Excel.

(35) In practice, *I.Sicily* only adopted GitHub in September 2017 (https://github.com/JonPrag/ISicily), so retracing to the original commit, at <https://github.com/JonPrag/ISicily/commit/6aa985e71b08e5f4f766a45b5ef02807288d49e9#diff-b8d4b5ef527bf3ba015eea52570c9325> (accessed 03.02.2019) would still not exactly match the dataset of the 2018 study based on 2016 data.

and their editions: the version of the edition that a scholar cites in 2016, will not necessarily be the same as the version that is currently visible on the web in 2019; and yet, for any subsequent reader, it is necessary to be able to control that scholarly citation against the version (or dataset) cited, and not the subsequently revised version that is now visible. In fact, this problem is solvable (*I.Sicily* currently plans to make use of the Zenodo repository for a referenceable archive of previous editions[36]), but few projects have yet systematically addressed the challenge, nor have academic practices necessarily woken up to the challenge.

The 2018 study differs from the 2002 study in a second and more obvious way, however, and in a way which directly bypasses the challenges outlined in the previous paragraph. The data extracted from *I.Sicily* for the purposes of that study can itself be published. In some ways it is curious that this point should even need to be made explicit; an equivalent study in the sciences would now require the publication of the underlying data (subject to any ethical/privacy considerations). In fact, it is a failure of the 2018 study that it does not make this point explicit within the publication (this quite simply reflects the gradual nature of the culture change underway - it only occurred to the author that this needed to be done after final submission of the manuscript). The data employed in that study, and which was employed for the maps and charts that are reproduced in this paper, is all freely available and referenceable via a DOI, as downloadable CSV files in the Oxford University Research Archive (ORA).[37] Consequently, it is possible for anyone quickly to reproduce the data analysis of that discussion and to compare the dataset to any other; or indeed to redeploy the data for further analysis or research.

This final point opens up a whole further set of possibilities and challenges relating to questions of data re-use. In principle the open publication of data (and very often not simply 'raw' data in a standardised form, but heavily curated data, enriched with additional metadata) should lay the foundations for increased collaboration and further study building upon such work. Failure to publish the data in open format closes off such possibilities, risks data loss (and so wasted effort), and forces redundant repetition of labour (compare the observations about historical linguistics in §1 above). The last part of this paper offers a second case-study, which engages with some of these issues.

### *2. Civic and Municipal Institutions in Sicily*

One frequent and important area of research into Hellenistic and Roman Sicily, which is heavily dependent upon epigraphic evidence, is the study of the civic institutions of the urban centres of the island. Studies of the institutional elements of both periods are legion.[38] Such studies are a mixture between those which aim at a more or less systematic tabulation of the evidence, and those which (more or less systematically) deploy an underlying survey of the evidence as the basis for a discursive analysis. However, what all of these studies have in common is that, at best, the evidence is listed in a series of references, generally buried within the text and (consequent upon past publication practices), lacking any uniformity of reference. Consequently, each study begins more or less from scratch, exploiting some or all of its predecessors to identify references and supplementing these with additional material in uneven fashion – and the ability of the reader meaningfully to compare

(36) Zenodo, named after the Alexandrian librarian Zenodotos, is funded by CERN and committed to Open Science: <http://about.zenodo.org/> (accessed 03.02.2019). A beta deposit of the dataset as a whole is already in place, at <https://doi.org/10.5281/zenodo.2556744> (accessed 26.05.2019).

(37) https://doi.org/10.5287/bodleian:YeRab9boO (accessed 26.05.2019) = Prag 2019.

(38) Merely *exempli gratia*, Cordano 1999, Cordano 2012, Ghinatti 1996, Ghinatti 2000, Manganaro 1988, Prag 2008, Sherk 1993, Vera 1996.

the coverage of these studies is extremely limited. Quite apart from the question of whether there is any meaningful advance in the sum of knowledge from one study to the next, the degree of repetition of effort in each subsequent study is (or should be) extremely depressing. (A further set of questions then arise as to the degree of openness in the existing literature about dependence upon earlier studies, which in fact adds weight to the arguments considered above in favour of publishing not only conclusions but also the underlying data; and in this context, it deserves emphasising that some of the issues that are now surfacing in the world of digital humanities are far from being merely 'digital' problems, but reflect more fundamental issues of best practice in scholarship.)

A notable example of this can be illustrated through the major research project EMIRE, co-directed by the late Mireille Cébeillac-Gervasoni and Clara Berrendonner, within the framework of UMR 8210 ANHIMA of the French CNRS (and which ran in parallel with a second project, 'Les pouvoirs locaux depuis l'Antiquité', co-directed by Mireille Cébeillac-Gervasoni and Laurent Lamoine at the Centre d'Histoire « Espaces et Cultures » of the Université Blaise-Pascal, Clermont-Ferrand). This project united a large international team of researchers (including the current author), who together compiled a very substantial (in principle, comprehensive) dataset of the evidence for municipal institutions across the western Roman empire. The output of the project lay in a series of major conference volumes, which united the works of analysis undertaken by the participants of the data collected on a city or regional or thematic basis.[39] The data itself, however, was never published, although copies of the underlying datasets (in very variable formats) circulated among the participants on CD-ROM. In retrospect, an enormous work of data collection, with very considerable potential for further research, was effectively discarded (although doubtless individual participants, like the current author, retained some version of their own datasets, and possibly even the original CD-ROM).

This example is not presented in order to criticise the specific project, since such 'criticism' could be levelled at almost every such project, including many that explicitly create digital resources such as databases, but then do not make the database publicly accessible to the wider research community.[40] Indeed, the EMIRE project was enormously stimulating of further research, a remarkable collaborative effort, and the volumes it produced were themselves of very great value. However, returning to the specific framework of the study of ancient Sicily and its civic institutions, it is instructive to examine the progress of subsequent study.

Back in 2008, the present author contributed the data and study of Sicily within the framework of the EMIRE project, using the 'private' database already described, which preceded the *I.Sicily* project. The result was a discursive work of synthesis, containing references to much of the data within the footnotes of the discussion.[41] The challenge in this case is to consider how this sort of work might be constructively built upon, rather than simply repeated. Although the *I.Sicily* project unites all the epigraphic texts which provide the evidence upon which the 2008 EMIRE study was built, the specific metadata on institutions attested within those texts and gathered for that study was not carried forward into *I.Sicily*. However, as noted, *I.Sicily* now contains all the texts of relevance, both metadata from the original database and the texts themselves; crucially, the format adopted, that of EpiDoc XML, permits the annotation of such additional information on the texts themselves

(39) Principally Berrendonner, Cébeillac-Gervasoni and Lamoine 2008; Lamoine, Berrendonner and Cébeillac-Gervasoni 2010.

(40) Although one reviewer of the 2008 volume did note that 'Une annexe précisant la méthodologie mise en œuvre lors des dépouillements aurait été utile au lecteur en lui permettant de mieux comprendre la démarche du projet et de mieux apprécier l'apport des différentes contributions, dont chacune mériterait une analyse détaillée' (Saliou 2009, online at: <http://histara.sorbonne.fr/cr.php?cr=690> (accessed 03.02.2019)).

(41) Prag 2008: 73-81. Retrospectively, the Sicilian dataset constructed for the project is now available at <https://doi.org/10.5287/bodleian:pzMe0PZAk> (accessed 07.02.2019).

(e.g. every instance of *duumvir* could be identified and marked up as such within the texts), in a way that enables automated indexing and data extraction, so that datasets for this sort of research can be readily and repeatedly generated in future. Furthermore, it is possible to record the individual responsible for such annotation of the data (i.e. the creation of the enriched dataset), so that each subsequent step in the research on this material both enriches the total available data, and is credited appropriately. Theoretically, this means that collaborative work reiterated on a single open dataset results in a continuous improvement of the data, rather than the loss of data or the repetition of effort. This is obviously an ideal situation, and it should be immediately acknowledged that a combination of existing academic culture (in particular the pressures induced by funding bodies and national research evaluation structures, which insist upon innovation, individual leadership, etc.), lack of resource (both financial and human), and a lack of agreed technical standards in a number of areas mean that this ideal is some way off becoming a reality.

The ambition deserves highlighting, however, given that in this particular case a major example of (currently) non-open-access reduplication of effort can now be observed in the digital project *ARCAIT* (=ARchivio delle Costituzioni Antiche in ITalia).[42] This project is gathering evidence (epigraphic and literary, with supporting imagery) relative to the institutional history of the urban communities across ancient Italy and Sicily.[43] However, the project, as currently presented, offers only the most basic free-text search facility and presents individual plain text editions of individual documents, without any visible overarching structure or external linking. The site is presented under the Principal Investigator's personal copyright, with no indication of whether or how the data may be accessed or employed, beyond simple on-line viewing, or citation of the individual records in their existence as plain-text web-pages. There is a very considerable amount of labour going in to this project, but the objective of the digital aspect remains very unclear. The first point being made here is that the labour risks being wasted, because it is seemingly closed off to other use, tacit about its methods and aims, and consequently a dead-end beyond the immediate aims of the specific project. The second point will obviously appear to be more self-interested, given the author's role in *I.Sicily*, but is not intended as such, since the fundamental point regards the advantage of the academic (and broader) community as a whole, and is equally applicable to any such project. Many of these documents are already available in *I.Sicily* (and indeed in other more or less open resources, such as *EDR*), and it would be entirely possible to carry out annotation of the relevant information on the *I.Sicily* files (whether within *I.Sicily*, or on an independent copy of the *I.Sicily* data). Adopting such an avowedly open and collaborative approach would, among other things: (a) require significantly less time and effort than the complete reassembling of the basic dataset; (b) potentially ensure that the data was in a standardised open source format, in turn open to integration with other datasets (ultimately through linked data structures); (c) enable others to continue to build on that research in the future through cumulative enrichment of the core data set; (d) potentially facilitate the long-term preservation of the data (since even if, e.g., *I.Sicily* were not to survive, the re-use of the same data collaboratively by multiple projects itself preserves the data).[44]

The crucial point is that this sort of approach does not require that anyone wanting to work on Sicilian epigraphy must work within, or subordinate to *I.Sicily*. Rather the failure of projects

(42) ARCAIT (Archivio Costituzioni Antiche in Italia), <www.arcait.it> (accessed 03.02.2019), directed by Loredana Cappelletti of the University of Vienna. It must be emphasised that, in singling out this project for discussion in this context, no criticism of the intellectual content or quality of the project itself is intended.

(43) The project self-describes on its homepage thus: 'uno strumento interdisciplinare di informazione e ricerca bibliografica, di studio e approfondimento, sulla storia politica e costituzionale dei popoli e delle comunità dell'Italia antica e della Sicilia in età preromana e romana (ca. VIII sec. a.C. - ca. IV sec. d.C.)' <www.arcait.it> (accessed 03.02.2019).

(44) Loosely, this is one of the principles of LOCKSS ('Lots Of Copies Keeps Stuff Safe'), < https://www.lockss.org/about/preservation-principles> (accessed 03.02.2019).

working on the same core data (a) to adopt open data principles and (b) to collaborate in the production, curation, and publication of that data, means that we will for the foreseeable future continue to advance as slowly as before, and large amounts of effort and resource will be repeatedly duplicated and therefore wasted. The current author can hardly claim the moral high ground, since almost everything I published previously was itself, until this point, unsupported by open access data. But with the steady rise of digital technologies and digital techniques, and with an increasing number of digital resources now available, the potential, as illustrated here with specific reference to the study of Hellenistic and Roman Sicily, is very considerable indeed. The opportunity for real advances, engendered by an ever increasing integration of different categories of evidence across the world of linked open data, which enable the asking (and answering) of new questions, is - or should be - genuinely exciting.

Jonathan R. W. Prag
Merton College, University of Oxford

*Bibliography*

Ball, A., Duke, M., 2015, *How to Cite Datasets and Link to Publications* (DCC How-to Guides), Edinburgh., <http://www.dcc.ac.uk/resources/how-guides> (accessed 31.01.2019).

Berrendonner, C., Cébeillac-Gervasoni, M., Lamoine, L. (eds.), 2008, *Le quotidien municipal dans l'Occident romain*, Clermont-Ferrand.

Campagna, L., 2011, 'Exploring Social and Cultural Changes in *Provincia Sicilia*: Reflections on the Study of Urban Landscapes,' in F. Colivicchi (ed.), *Local Cultures of South Italy and Sicily in the Late Republican Period: Between Hellenism and Rome*, Portsmouth, RI, p. 161-183.

Cordano, F., 1999, 'Le istituzioni delle città greche di Sicilia nelle fonti epigrafiche,' in M. I. Gulletta (ed.), *Sicilia Epigraphica. Atti del convegno internazionale, Erice, 15-18 ottobre 1998*, Pisa, vol. 1, p. 149-158.

Cordano, F., 2012, 'Eponimi ufficiali nella Sicilia di età ellenistica,' in C. Ampolo (ed.), *Agora greca e agorai di Sicilia*, Pisa, p. 77-80.

De Vido, S., 1999, '*Corpora* epigrafici siciliani da Gualtherus a Kaibel,' in M. I. Gulletta (ed.), *Sicilia Epigraphica. Atti del convegno internazionale, Erice, 15-18 ottobre 1998*, Pisa, vol. 1, p. 221-250.

Elliott, T., 2014, 'Epigraphy and Digital Resources,' in C. Bruun and J. Edmondson (eds.), *The Oxford Handbook of Roman Epigraphy*, Oxford, p. 78-86, <https://doi.org/10.1093/oxfordhb/9780195336467.013.005> (accessed 31.01.2019).

Elliott, T., Bodard, G., Cayless, H. *et al.*, 2006-2017, *EpiDoc: Epigraphic Documents in TEI XML*, <http://epidoc.sf.net> (accessed 31.01.2019).

Ghinatti, F., 1996, *Assemblee greche d'Occidente*, Turin.

Ghinatti, F., 2000 [2004], 'Le organizzazioni civiche siceliote,' *Kokalos* 46(1), p. 31-73.

Hardwicke, T. E., Mathur, M. B., MacDonald, K. *et al.*, 2018, 'Data Availability, Reusability, and Analytic Reproducibility: Evaluating the Impact of a Mandatory Open Data Policy at the Journal *Cognition*,' *Royal Society Open Science* 5(8), <https://doi.org/10.1098/rsos.180448> (accessed 31.01.2019).

Lamoine, L., Berrendonner, C., Cébeillac-Gervasoni, M. (eds.), 2010, *La praxis municipale dans l'Occident romain*, Clermont-Ferrand.

Manganaro, G., 1988, 'La Sicilia da Sesto Pompeo a Diocleziano,' *ANRW* 2.11.1, p. 3-89.

McCleery, R., 2016, 'Being Roman, Writing Latin? Consumers of Latin inscriptions in Achaia,' *Chronika* 6, p. 41-57, <http://www.chronikajournal.com/resources/McCleery%202016.pdf> (accessed 02.02.2019).

Mimbrera Olarte, S., 2012, *Fonética y morfología del Dorio de Sicilia (siglos VII-I a.C.)*, Madrid.

Opitz, R., Mogetta, M., Terrenato, N. (eds.), 2016, *A Mid-Republican House from Gabii*, Ann Arbor, online only: <https://doi.org/10.3998/mpub.9231782> (accessed 31.01.2019).

Poehler, E. E., 2018, 'Review of Opitz, R., Mogetta, M., and Terrenato, N. (eds.), *A Mid-Republican House from Gabii* (2016)', *Journal of Roman Studies* 108, p. 192-194.

Prag, J. R. W., 2002, 'Epigraphy by Numbers: Latin and the Epigraphic Culture in Sicily,' in A. E. Cooley (ed.), *Becoming Roman, Writing Latin?*, Portsmouth, RI, p. 15-31.

Prag, J. R. W., 2003, 'Nouveaux regards sur les élites locales de la Sicile républicaine,' *Histoires et sociétés rurales* 19, p. 121-132.

Prag, J. R. W., 2004, *Sicily and the Roman Republic 241-44 BC: Provincialization and Provincial Identities*, PhD thesis, Dept. of History, University College London, <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.420608> (accessed 31.01.2019).

Prag, J. R. W., 2007, 'Ciceronian Sicily: The Epigraphic Dimension,' in J. Dubouloz, S. Pittia (eds.), *La Sicile de Cicéron, Lectures des Verrines*, Besançon, p. 245-271.

Prag, J. R. W., 2008, 'Sicilia and Britannia: Epigraphic Evidence for Civic Administration,' in C. Berrendonner, M. Cébeillac-Gervasoni, L. Lamoine (eds.), *Le Quotidien municipal dans l'Occident romain*, Clermont-Ferrand, p. 67-81.

Prag, J. R. W., 2010, 'Sicilia Romana tributim discripta,' in M. Silvestrini (ed.), *Le tribù romane. Atti della XVI^e Rencontre sur l'épigraphie (Bari 8-10 ottobre 2009)*, Bari, p. 305-311.

Prag, J. R. W., 2014, 'Cities and Civic Life in Late Hellenistic Roman Sicily,' *Cahiers du Centre Gustave Glotz* 25, p. 165-208.

Prag, J. R. W., 2018, 'The Birth of Epigraphic Culture in the Western Mediterranean: Sicilian Epigraphic Culture in the Later Hellenistic Period,' in F. Beltrán Lloris and B. Díaz Ariño (eds.), *El nacimiento de las culturas epigráficas en el occidente Mediterráneo. Modelos romanos y desarrollos locales (ss. III-I a.E.)*, Madrid, Anejos de *AEspA* LXXXV, p. 131-144 (author manuscript online: <https://ora.ox.ac.uk/objects/pubs:836613>).

Prag, J. R. W., 2019, 'Quantification of Lapidary Epigraphy for Ancient Sicily', University of Oxford, <https://doi.org/10.5287/bodleian:YeRab9boO> (accessed 04.06.2019).

Prag, J. R. W. and Chartrand, J., 2018, '*I.Sicily*: Building a Digital Corpus of the Inscriptions of Ancient Sicily,' in A. De Santis and I. Rossi (eds.), *Crossing Experiences in Digital Epigraphy. From Practice to Discipline*, Warsaw, p. 240-252, <https://doi.org/10.1515/9783110607208-020> (accessed 01.05.2019).

Prag, J. R. W. and Tigano, G., 2017, *Alesa Arconidea*: *il lapidarium*, Palermo.

Prag, J. R. W., Chartrand, J., and Cummings, J., 2017, 'I.Sicily: An Epidoc Corpus for Ancient Sicily,' in S. Orlandi, R. Santucci, P. M. Liuzzo, F. Mambrini (eds.), *Digital and Traditional Epigraphy in Context. Proceedings of the Second EAGLE International Conference. Rome 27-29 January 2016*, Rome, p. 83-96, <http://digilab-epub.uniroma1.it/index.php/Antichistica/issue/view/38> (accessed 31.01.2019).

Prestianni Giallombardo, A. M., 2012, 'Spazio pubblico e memoria civica. Le epigrafi dall'agora di Alesa,' in C. Ampolo (ed.), *Agora greca e agorai di Sicilia*, Pisa, p. 171-200.

Reynolds, J., Roueché, C., Bodard, G., 2007, *Inscriptions of Aphrodisias*, <http://insaph.kcl.ac.uk/iaph2007> (accessed 31.01.2019).

Sherk, R., 1993, 'The Eponymous Officials of Greek Cities V: The Register,' *ZPE* 96, p. 267-295.

Tribulato, O., 2012, '*Siculi bilingues*? Latin in the Inscriptions of Early Roman Sicily,' in O. Tribulato (ed.), *Language and Linguistic Contact in Ancient Sicily*, Cambridge, p. 291-325.

Vera, D., 1996, 'Augusto, Plinio il Vecchio e la Sicilia in età imperiale. A proposito di recenti scoperte epigrafiche e archeologiche ad Agrigento,' *Kokalos* 42, p. 31-58.

Wilson, R. J. A., 1990, *Sicily under the Roman Empire: the archaeology of a Roman province, 36 B.C.-A.D. 535*, Warminster.