

Jonathan R.W. Prag and James Chartrand

19 I.Sicily: Building a Digital Corpus of the Inscriptions of Ancient Sicily

Abstract: This paper presents the I.Sicily project. We focus first upon its original rationale and construction, since this provides explanations for the particular choices and approaches adopted, before exploring some of the challenges faced, as well as current and future developments. We believe that I.Sicily offers an interesting case study of a deliberately open-ended, continuous work-in-progress corpus. The project is constructed on the assumption that collaboration is key to its success, and that collaboration will only increase. We examine the potential for the creation of Linked Open Data, which we consider essential to creating the primary point of reference for the study of Sicilian epigraphy, and to the creation of a resource to support and facilitate research while simultaneously enhancing and supporting the accessibility of Sicilian epigraphy. This last aim is served both directly through the project's web-interface, and indirectly by supporting and facilitating the work of the institutions which curate the majority of the material: we conclude with an illustration of a wide-ranging, museum-based, community collaboration.

Keywords: ancient Sicily, EpiDoc, museums, Linked Data, onomastics

19.1 Background

I.Sicily¹ is a corpus of the inscribed texts from ancient Sicily. This includes the very earliest written texts from the island (late seventh/early sixth century BCE), and extends to late Antiquity and the Byzantine period (seventh century CE). At present, for historical reasons and practical purposes, the primary coverage of the project is texts inscribed on stone (between 4,000 and 5,000 in total; currently 3,246 records). In due course, we will extend coverage to include other inscribed materials (especially metal and ceramic) and portable objects (*instrumentum domesticum*). A pilot project is under development to explore the creation of a sub-corpus of coin-legends in the same format. The epigraphic culture in ancient Sicily includes texts

¹ I.Sicily [<http://sicily.classics.ox.ac.uk>].

Jonathan R.W. Prag, University of Oxford
James Chartrand, Open Sky Solutions



© 2018 Jonathan R.W. Prag and James Chartrand

This work is licensed under the Creative Commons Attribution 4.0 International license (CC BY 4.0)

Brought to you by | University of Oxford - Bodleian Libraries

Authenticated

Download Date | 7/11/19 2:47 PM

written in Phoenician/Punic, Greek, Oscan, Latin, Hebrew, and two of the indigenous languages, Sikel and Elymian (for overviews of Sicilian epigraphy and linguistics, see the contributions in Gulletta, 1999; Tribulato, 2012a).

The original motivation for I.Sicily lies in traditional problems of publication and access. Sicily has a very long tradition of epigraphic study and corpus creation (De Vido, 1999): the first modern history of the island, which included epigraphic texts, is the *de rebus Siculis* of Tommaso Fazello (1558), and the first epigraphic corpus was published by Georg Gualtherus in 1624; Sicily was the subject of some of the earliest volumes of the monumental Berlin projects, *Corpus Inscriptionum Latinarum* (vol. X.2 = Mommsen, 1883) and *Inscriptiones Graecae* (vol. XIV = Kaibel, 1890). However, the rate of both discovery and publication increased rapidly from the late 1880s onwards, and the ability of both the primary publications (such as the gazettes, *Supplementum Epigraphicum Graecum* and *L'Année Épigraphique*) and scholars to keep pace with new material has been limited. The situation is compounded by the very uneven practices in the publication of archaeological excavation, and there is an unknown and not insignificant quantity of unpublished material (often highly fragmentary) languishing in stores across the island. Consequently, the discussion of Sicilian epigraphy has tended to be concentrated very narrowly in the hands of specialists, not simply for disciplinary reasons, but due to the difficulties of comprehensive knowledge (an emblematic example is Manganaro, 1988, an unparalleled discussion of the material of the Roman imperial period, referencing hundreds of texts, and alluding frequently to unpublished or obscure and unreferenced texts). These challenges have become even more visible in recent scholarship with the increased focus upon socio-linguistics, which depends upon the ability to engage with a comprehensive dataset. As Olga Tribulato recently noted, “Arguments [on the linguistic history of ancient Sicily], and the statistics on which they rely, are destined to remain little more than hypotheses, until a comprehensive list of all epigraphic texts from ancient Sicily is assembled” (Tribulato, 2012b, p. 324).

Against this backdrop, Jonathan Prag originally attempted to create just such a list of the lapidary inscriptions of Sicily. This was carried out within the framework of a PhD on Roman Sicily (London, 1999–2004), of which the initial results were published as a quantitative analysis (Prag, 2002), in order to assess epigraphic culture on the island. That project did not concentrate on the texts themselves, but on creating a reference list based upon bibliographic citations, together with a limited amount of metadata. The original list was created in a flat table in MS Access 97 (upgraded several times subsequently). This dataset was intermittently maintained and updated on a series of private computers over the following decade, during which time its value as a research tool became increasingly apparent.² The same period witnessed

² Facilitating e.g. Prag, 2003; 2007; 2008; 2010.

the development of the EpiDoc TEI-XML standard,³ and in 2011 several bids were submitted to funding bodies to transform the existing dataset into an EpiDoc corpus. The primary funding for the creation of I.Sicily was provided by a grant of £80,000 from the John Fell Fund of the University of Oxford, which was used over the period 2013–2015.⁴

The principal development work undertaken over that period consisted of (a) the transformation of the legacy dataset from an Access table to a set of EpiDoc files; (b) the construction of the necessary back-end and front-end tools to make a usable corpus with a flexible web interface.⁵ In its final form, the original table held data across some 40 different fields, for c. 3,200 records; 18 of these fields detailed publication history (corpora references and other bibliography); the other fields recorded information on the language, date, provenance, current location, epigraphic type, form and material of the inscriptions, together with a free-text field recording further information about the inscription and fields to record any autopsy undertaken. Almost all of this data was derived from existing publications. After extensive cleaning of the data, the conversion from the original MS Access dataset was developed through a pipeline of known conversions passing from MS Access to CSV to TEI P5 XML. The subsequent XSLT transformation of the table of data from TEI P5 XML to EpiDoc XML provided an ideal opportunity to enrich the existing dataset, both to normalise the data and to lay the foundations for Linked Open Data. This was done, principally, by the embedding of reference to multiple external authority lists (local correspondence lists were created in CSV files during the pre-conversion cleaning of the original data to facilitate this alignment). This enabled the incorporation of Pleiades and Geonames URIs on the “ref” attribute for `<placeName type="ancient">` and `<placeName type="modern">`, as well as the inclusion of representative decimal-degree location data in a `<geo>` element, to simplify local mapping. EAGLE vocabularies were incorporated for `@ref` on `<material>`, `<objectType>`, `<rs type="execution">` (in `<layout>`) and for epigraphic type on `<term>` within the `<textClass>` element.⁶ Two new resources were created as part of the process of transforming the data: an open bibliography in Zotero⁷ and a new museums database.⁸ URIs are maintained for both sets of data (for bibliographic items

³ [<https://sourceforge.net/p/epidoc/wiki/Home/>]; Bodard, 2010.

⁴ Additional small supporting grants have been provided by the John Fell Fund, by the Warden and Scholars of Merton College, Oxford, and by the Craven Committee of the Faculty of Classics, University of Oxford.

⁵ (a) was undertaken by Dr James Cummings (then Senior Academic Research Technology Specialist, IT Services, University of Oxford); (b) has been undertaken by James Chartrand, Open Sky Solutions.

⁶ [<https://pleiades.stoa.org/>]; [<http://www.geonames.org/>]; [<https://www.eagle-network.eu/resources/vocabularies/>].

⁷ [<https://www.zotero.org/groups/382445/isicily/items>].

⁸ The museum database is an ongoing project initiated by Dr Michael Metcalfe, with an online interface at: [<http://sicily.classics.ox.ac.uk/museums>].

these are already published as RDF by Zotero), and during the process of conversion reference to both was incorporated on the <repository> and <bibl> elements in the TEI in anticipation of Linked Open Data. The one significant element of metadata which was normalised but not externally referenced was the dating information, and reference to, for example, [<http://perio.do/>] remains a future possibility.

The final element that was incorporated during the conversion process was the epigraphic text itself, since this was not included in the original dataset. This was done through an automated process, using available digitally published texts, exploiting the inclusion of existing digital identifiers in the original dataset (I.Sicily URIs are also aligned with Trismegistos text numbers, which facilitates further alignment with other digital epigraphic databases and corpora).⁹ The vast majority of these texts (generously made available, e.g., by the EDR project) were themselves not originally created in EpiDoc, and so automated conversions were applied, either by providers at source (as in the case of EDR) or at the point of capture and incorporation.¹⁰ Such automated transforms are not perfect, and commonly the underlying published source of the text is not captured through this process. Consequently, while more or less functional texts have been incorporated into approximately two-thirds of the EpiDoc files, all of these require human checking, further editing and appropriate attribution (all I.Sicily records carry a visible “status” indicator of “edited”, “draft” or “unchecked”). This is a pressing need, not least to ensure user-acceptance of the corpus, and is independent of the long-term aim to conduct autopsy and revision for all the inscriptions in the corpus (although the two steps can obviously be combined). At the same time, some hundreds of files remain without any data in the text division, and almost all require the inclusion of a translation. This creates both a challenge and an opportunity, which we discuss below.

The conversion was a one-time process, and subsequent editing has been managed through the use of XML editors and the interface provided by the I.Sicily website and eXist. The correspondence lists created for the upgrading of the data during conversion continue to be maintained, serving as local authority lists, in order to facilitate standardisation and external referencing in the continued editing of existing XML records and in the creation of new records. Where necessary, additional local authority lists will be created (e.g. for names and persons), when the current state of external authorities is insufficient. At present, for the purposes of data management and version control, the XML files and correspondence/authority lists are managed in an open-access GitHub repository.¹¹ For the purposes of actual digital publication and searching, the latest version of the XML records are held on a server hosted by

⁹ [<http://www.trismegistos.org/>]; 2563 records currently aligned.

¹⁰ [<http://www.edr-edr.it/default/index.php>]: import of converted EDR texts was kindly facilitated by Pietro Liuzzo.

¹¹ [<https://github.com/JonPrag/ISicily>].

the Faculty of Classics (University of Oxford), in an eXist database for xQuery access. URIs are maintained for the inscriptions and the museums with an eye to Linked Open Data, and both are manipulated through a RESTful API; the bibliography is published as Linked Open Data and edited directly in Zotero. The records are queried and viewed through a web interface built with AngularJS and jQuery JavaScript components. Mapping is provided in the browser by the Google Maps API. The search interface as a whole has been built very much with the difficulties of researchers in mind, exploiting new JavaScript libraries to create a spreadsheet-like interface that is flexible and reasonably intuitive, and facilitates easy export of search results.¹²

Images were not part of the original dataset (for the same reasons that texts were not). In the conversion, a standard template for the <facsimile> element was created in the EpiDoc, but individual image data needs to be edited into the XML files as the images become available. Currently this is a slow, manual task. We aim to make high-resolution imagery available wherever possible. In the web-interface, ZPR (Zoom, Pan, Rotate) image-viewing is provided by the IIP image server (which also enables the generation of IIIF metadata) and the OpenSeadragon JavaScript library.

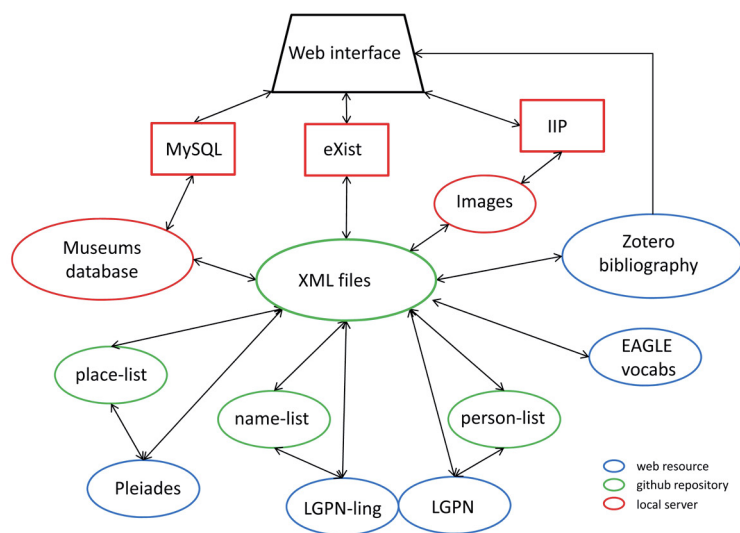


Figure 19.1: Graphic representation of the data organisation of I.Sicily

¹² We gratefully acknowledge the generosity of [<https://www.ag-grid.com/>] who made the Enterprise version of their JavaScript grid available to us for free. For an overview of the search functionality of I.Sicily, see [<https://isicily.wordpress.com/how-to/>].

All of the above creates a rather complex and atomised data management structure, with XML files, authority lists, Zotero bibliography, images, and museums database held in diverse locations and curated in different ways (see Figure 19.1 for a graph). At the same time, it can be argued that this creates a very flexible system, exploiting open-source tools where possible and using standardised formats to ensure maximum interoperability, with the result that preservation and maintenance overheads are kept to a minimum. This approach is particularly well adapted to the very fluid data flows involved in curating and publishing a complex set of data that is subject to continual revision and improvement, and a continuous drip of minor updates, rather than the one-off presentation of a static dataset.

19.2 Challenges and Ambitions

19.2.1 Text-Editing and Annotation

As already noted, one of the immediate challenges faced by I.Sicily is the need to edit the text division for a large number of epigraphic texts. This task has several aspects and phases to it, each of which offers different challenges and potential solutions:

- (a) the editing of existing or missing texts, based upon published editions;
- (b) the inclusion or revision of texts based upon autopsy;
- (c) the development of a full critical apparatus for a complete edition combining (a) and (b);
- (d) the extension of mark-up, such as to record onomastic, prosopographic, or linguistic information.

With over 3,000 records, notwithstanding the fact that many are short funerary texts, this is a substantial task requiring a considerable investment of time.

Basic revision and editing (i.e. task (a)) provides a ready opportunity for developing EpiDoc training, since the I.Sicily records offer a rich set of material for students to practise editing using common tools such as the oXygen XML editor, as well as to become familiar with the basics of GitHub, which provides a convenient data management tool. At the same time, students can gain credit for their work since I.Sicily makes full use of the <resp> and <change> elements, and publishes that information in the HTML and PDF editions generated from the EpiDoc. A teaching support grant from the University of Oxford in 2015 facilitated the embedding of EpiDoc teaching within existing epigraphic teaching at the masters level, creating the necessary supply and demand relationship; and volunteer encoders have been

forthcoming.¹³ Needless to say, such a collaborative approach requires that the documentation of the precise structure of the EpiDoc mark-up employed needs to be rigorous and available in advance in order to minimise irregularities in the edited files. The greatest challenge, however, is simply one of human resource: the resulting rapid increase in the generation of revised files, which require management and curation prior to release, creates a potential bottleneck, unless additional resources of time (or money to buy additional support) become available.

The contribution of more comprehensive revision (i.e. task (b)) based upon new information (especially through autopsy), or else of a new record for a text not previously included, in both cases including new or revised metadata, is a more challenging scenario. In principle, this can be managed through the same set of mechanisms as task (a). However, on the one hand, the free marking-up of metadata creates greater risks of irregularities; and on the other, many of those submitting such information will come from outside the academy and/or will neither have access to nor familiarity with, e.g., XML-editing (we return below to the collaborative approach responsible for this situation). Such a situation creates a need for alternative solutions to data entry, since it is both more empowering for the contributor, and more efficient for the editor, if this process can be as direct as possible (notwithstanding that a more basic approach is always possible, with an editor taking on the task of transforming data submitted in any form into a compliant XML file). At present we are experimenting with the use of an online web form,¹⁴ which allows submission of a flexible range of data, while also constraining data formats for some fields and offering pre-set choices for metadata fields where authority lists exist. The form is used to generate a pre-populated XML file from the project's EpiDoc template, which is then submitted for editing. The form is still in development and, in line with the overall initial focus of the project, is again focused more on rich metadata than text-editing. A robust, web-based GUI for direct editing and revision of the actual epigraphic text remains a desideratum, but is not an immediate priority (contributors are currently left free to submit the text itself in whatever format they feel most comfortable). Pilot contributions of several sorts are underway using the form, with one set of collaborators repurposing the HTML form for use by students at a local school in Sicily (see below).

The creation of a comprehensive critical apparatus to support a final edited text (i.e. task (c) above) is a long-term desideratum, enabling the effective capture and comparison of the full information from past editions as well as fresh autopsy, but it is also a more complex challenge. In the first place, this remains an area slated for future development within the wider EpiDoc community and a relatively underdeveloped

¹³ "Digital Techniques in the study of Ancient Epigraphy: transforming MSt/MPhil teaching", £ 2884 from the Humanities Division of the University of Oxford, academic year 2014/15.

¹⁴ [<http://sicily.classics.ox.ac.uk/login/inscription-submit/>]. Use of the form requires the user to register, but is not restricted.

area among the majority of existing projects.¹⁵ In the second place, even with such structural choices resolved, a tool that would enable editing of this part of the text mark-up, without the user having to engage directly with the increasingly complex XML involved, would be non-trivial to construct. However, examples do already exist within the wider TEI community of manuscript studies (e.g. Burghart, 2016). Part of the problem is that the demand for such an interface is more limited, since the level of already specialist knowledge entailed makes the user-group for such a tool too small to warrant the investment, at least at the scale of a project like I.Sicily. All of this implies that, in the short-term at least, this area is likely to be a significant roadblock in the final editorial development of the dataset.

A final area of text annotation, which we are currently attempting to address, is the indexing of terms within the ancient text (task (d) above). Here too, our interest lies in trying to facilitate multiple contributors, often without the ability to work directly in the XML, and not simply in resolving the problems of choosing between internal and external authority lists (where the latter even exist; see below). The two issues are, however, inter-related, since incorporating the direct referencing of external authority lists requires a different set of tools from simply building an internal list. Emblematic is the particular challenge presented by the indexing of names and individuals.¹⁶ For the present, we treat the annotation of names and individuals as a discrete task, separate from general text-editing, and we are therefore content to employ a separate editing tool in order to enable the rapid annotation of names and persons across the full set of texts, by multiple contributors. The “micro-editor” for this purpose is being developed through the participation by I.Sicily in the CANARIE-funded Canadian Writers Research Collaboratory project, as one of a number of open-source tools for TEI-based projects.¹⁷ We are attempting to leverage this development work with a grant from the John Fell Fund of the University of Oxford, which will permit the necessary development work within the *Lexicon of Greek Personal Names* (individuals) and the new *LGNP-Ling* database (names).¹⁸ The latter will enable the publication of URIs for both named individuals in ancient Greek (i.e. persons) and names as linguistic entities, addressable via an API.

¹⁵ Cf. [<http://www.stoa.org/epidoc/gl/latest/supp-apparatus.html>] and recent discussion on the Mark-Up list at [<http://lsv.uky.edu/scripts/wa.exe?A1=ind1710&L=markup#3>].

¹⁶ See especially [<https://snapdrgn.net/>].

¹⁷ Canadian Writing Research Collaboratory Extension, grant of \$CDN 456,139 for 2017-2019 from CANARIE: see [<http://beta.cwrc.ca/>].

¹⁸ See [<http://www.lgpn.ox.ac.uk/>] and [<http://clas-lgpn2.classics.ox.ac.uk/>] for the XML database and [<http://admin.exist-db.org:41233/exist/apps/lgpn-ling/about.html>] for the new linguistic database of names.

19.2.2 Linked Open Data?

Referencing external authority lists provides an opportunity to enable greater interoperability and the creation of Linked Open Data. As has previously been observed, while EpiDoc is a huge step forward in our ability to record and represent ancient inscriptions in a rich, machine-readable, digital format, nonetheless it risks perpetuating some of the traditional challenges posed by rich but ultimately non-standardised datasets, since almost every EpiDoc project develops its own customisations and an EpiDoc file, “consists in a monolithic, self-descriptive and self-standing information unit” (Casarosa et al., 2014, p. 24, p. 28). One (partial) solution to this challenge is the use of externally referenceable controlled vocabularies – as noted above, extensive use of such reference has been incorporated into the I.Sicily EpiDoc files.

The epigraphic community has been among the leaders in the move towards the Linked Open Data approach in ancient world studies (Geser, 2016, p. 10). The stand-out example is the work of the EAGLE project, creating a set of SKOS vocabularies to enable cross-lingual referencing of core epigraphic metadata concepts.¹⁹ However, as yet, the overall ontological framework has not been established to enable the full publication of EpiDoc files as RDF, and only very partial examples of the possibilities exist.²⁰ A number of reasons can be suggested (Geser, 2016 offers a thoughtful analysis in the context of archaeological data), and two might be singled out. The first, is the fact that both controlled vocabularies and referenceable authorities for many epigraphic elements are still lacking. The EAGLE vocabularies themselves are still a work-in-progress, currently lacking a clear framework for community development (this is said to be in hand), and they are not consistently adopted since they are themselves mostly aligned to larger vocabularies (e.g. DAI and Getty). As the EAGLE project itself disarmingly observes on the vocabularies landing page, “perhaps one day we will be able to do nice things as those Pelagios, Pleiades and SNAP-DRGN do [*sic*], also based on these vocabularies.” However, even the reference to SNAP-DRGN is optimistic, since currently online prosopographies themselves are a work-in-progress (the projected work on the LGPN database, referenced above, will hopefully help move this forward). The principal area where such referencing is currently possible is in the realm of geographical data. Having referenced place-name information in I.Sicily with Pleiades URIs, we have been able to generate the necessary RDF export for Pelagios, in a working demonstration of the possibilities of Linked Open Data.²¹ However, it remains the case that for most such projects, this is currently the one effective area

¹⁹ [<https://www.eagle-network.eu/resources/vocabularies/>].

²⁰ Contrast the work by the numismatic community at [<http://nomisma.org/>].

²¹ See [<http://peripleo.pelagios.org/about>] and [<http://peripleo.pelagios.org/ui#selected=http://sicily.classics.ox.ac.uk/pelagios-data/isicily-pelagios-dataset>].

where Linked Open Data is a practical reality, and this is due to the success of the Pleiades gazetteer.²² The second reason is the outstanding need to create a map from EpiDoc to a set of RDF ontologies (which entails choosing the ontologies themselves, the appropriate terms within the ontologies and, where no appropriate ontologies exist, creating a new ontology with new terms). Initial work has been undertaken on mapping EpiDoc to CIDOC-CRM (Casarosa et al., 2014) and a further discussion of epigraphic ontologies took place at the recent Open Epigraphic Data Unconference (London, 15 May 2017).²³ It is clear that trying to coordinate this work with others would be best in the long term, but it remains difficult to coordinate in the short term. Consequently, it remains tempting to move ahead independently and seek to publish a smaller subset of some basic RDF (as with the geographical data), mapping independently without consultation, on the assumption that such mappings could later be changed, and with the aim of encouraging further development.

In any event, I.Sicily has chosen to privilege external authority lists wherever possible, in anticipation of Linked Open Data. However, in many cases the incomplete nature of such lists means that an internal authority list is also necessary, and unless those internal lists are also maintained, published, and potentially externally aligned in the future, Linked Open Data remains a hope rather than a reality. Currently, we appear to be in something of a vicious circle, since the resource required to get Linked-Open-Data-ready is not negligible, while the demonstrable short-term (and even medium-term) gains from such activity are few and far between, meaning that there is little incentive.

19.2.3 Collaboration and Outreach

Although the core data of the initial instantiation of I.Sicily is derived from existing publications, moving forward we aim fully to revise each inscription record on the basis of identification of the original object and full autopsy. Such an approach is impossible without the collaboration of the museums that hold the majority of the material.²⁴ I.Sicily has therefore been constructed in a deliberately museum-centric fashion, publishing a gazetteer of Sicilian museums.²⁵ This enables the direct linking of epigraphic records to museum collections, and in turn the effective online publication of individual catalogues of museums' epigraphic collections. On the one

²² [<https://pleiades.stoa.org/>]; [<http://commons.pelagios.org/>].

²³ [<https://github.com/EpiDoc/OEDUC/wiki>].

²⁴ More broadly, the overall objective of I.Sicily as a comprehensive corpus for Sicily is impossible without extensive collaboration from a wide group of experts: something which the model of digital publication with explicit attribution of responsibility clearly facilitates.

²⁵ [<http://sicily.classics.ox.ac.uk/museums>].

hand, this serves the needs of researchers who want to be able to locate individual inscriptions for study. On the other, this makes the corpus of direct value to the museums themselves, both as a service for the curatorial staff and as a potential tool for virtual display of material and other forms of increased accessibility.²⁶

As noted above, this creates challenges in the work of collaborative recording, and we are experimenting with several models. The most productive and exciting of these to date has been a joint project with the Museo Civico Castello Ursino of Catania, the city of Catania, the Liceo artistico statale M.M. Lazzaro, and the CNR Istituto di Scienze e Tecnologie della Cognizione (ISTC) at Catania (Agodi et al., 2018). Exploiting the possibilities of the Italian Ministry of Education, Universities and Research (MIUR) “alternanza scuola-lavoro” programme (i.e. work experience for school students), we have worked with students and teachers from a large secondary school in Catania on the work of cataloguing the epigraphic collection of the Catania civic museum. A group from the CNR-ISTC (the “EpiCUM project” directed by Dr Daria Spampinato) has in turn worked with the students, developing a version of our own HTML record form to enable the students to input data into an automatically generated XML file. The CNR-ISTC project is in turn using the I.Sicily template for a digital catalogue of the non-Sicilian inscriptions in the collection (EpiCUM). All parties worked together to curate a permanent exhibition (“Voci di pietra”) in the museum of a selection of 35 inscriptions from ancient Catania, which opened on 14 July 2017.²⁷ The EpiCUM project is also developing a parallel virtual exhibition, in part based upon the I.Sicily EpiDoc files. The students undertook cleaning, recording and conservation work in the museum prior to the exhibition, and played a leading role in the design and production of the exhibition itself. Subsequently, they have continued cataloguing and recording the c. 500 inscriptions in the museum’s collection. With additional funding from the University of Oxford, a follow-up collaboration is now being planned with a second Liceo at the Museo Archeologico Regionale “Paolo Orsi”, in Siracusa. An approach of this sort creates many problems of its own, but two very clear advantages can be observed: firstly, a very much more rapid aggregation of (genuinely high quality) data; secondly, a real sense of community engagement and empowerment, bringing local epigraphic material into the public consciousness, rendering it comprehensible as ‘voices of stone’ from a community’s past.

²⁶ Note e.g. the *izi.travel* project, which is very active in Sicily and which can link to I.Sicily for epigraphic objects (see [<https://izi.travel/it/4d91-museo-archeologico-regionale-paolo-orisi/it>] for an example of a museum tour).

²⁷ The exhibition was supported by a Knowledge Exchange Fellowship from The Oxford Research Centre in the Humanities, 2016/17. Press coverage includes: [http://www.corriere.it/foto-gallery/cultura/17_agosto_03/catania-romana-raccontata-voci-pietra-48d4802e-77ec-11e7-84f5-f24a994b0580.shtml] and [<http://www.globusmagazine.it/110708-2/#.WiR9MDdpGwV>].

19.3 Conclusions

There are a number of further challenges presented by the I.Sicily corpus which we have not considered here, such as the complications presented by a very non-uniform corpus covering not only a very extended period in time (and so, e.g., Archaic texts compared to Christian texts), but also an increasingly wide variety of materials, and in particular a rich mixture of languages, not all of which have a Unicode character set. From a practical perspective, the current state of the relevant technologies and limited availability of resources makes an undertaking of this sort extremely challenging, above all if one seeks to build an open, collaborative project, rather than a closed, local dataset resulting in a static publication. From a purely scientific perspective, the greatest challenge remains the acceptance not only of a born-digital publication, but also of a publication that is not stable in the traditional sense and has no clear single publication date. Transparency and rigorous, detailed attribution of responsibility appear, to us, to be the most effective responses to this, hopefully temporary, problem. Nonetheless, we have been hugely encouraged by the enthusiasm with which colleagues, museums, local authorities, and local communities have embraced the project so far, and we remain fundamentally optimistic about the potential for the future – not least because of the strength of the EpiDoc community itself.

Bibliography

- Agodi, S, Cristofaro, S., Noto, V., Prag, J., & Spampinato, D. (2018). Una collaborazione tra museo, enti di ricerca e scuola: l'epigrafia digitale e l'alternanza scuola lavoro. *Umanistica Digitale*, 2. doi: 10.6092/issn.2532-8816/7298. Retrieved from [https://umanisticadigitale.unibo.it/article/view/7298], 2018/06/18.
- Bodard, G. (2010). EpiDoc: Epigraphic Documents in XML for Publication and Interchange. In F. Feraudi-Gruénais (Ed.), *Latin on Stone: Epigraphic Research and Electronic Archives* (pp. 1–17). Lanham, MD: Lexington Books.
- Burghart, M. (2016). The TEI Critical Apparatus Toolbox: Empowering Textual Scholars through Display, Control, and Comparison Features. *Journal of the Text Encoding Initiative*, 10. doi: 10.4000/jtei.1520. Retrieved from [http://jtei.revues.org/1520], 2017/11/17.
- Casarosa, V., Manghi, P., Mannocci, A., Rivero Ruiz, E., & Zoppi, F. (2014). A Conceptual Model for Inscriptions. In S. Orlandi, R. Santucci, V. Casarosa, & P.M. Liuzzo (Eds.), *Information Technologies for Epigraphy and Cultural Heritage. Proceedings of the First EAGLE International Conference, Paris* (pp. 23–40). Rome: Sapienza Università Editrice. Retrieved from [https://www.eagle-network.eu/wp-content/uploads/2015/01/Paris-Conference-Proceedings.pdf], 2017/12/03.
- 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* (vol. I, pp. 221–250). Pisa: Edizioni della Normale.
- Fazello, T. (1558). *F. Thomæ Fazelli... de rebus Siculis decades duæ*. Panormi: Panormi.
- Geser, G. (2016). *Towards a Web of Archaeological Linked Open Data* (ARIADNE WP15 Study. V1.0, 6 October 2016). Salzburg Research, Austria. Retrieved from [http://www.

- ariadne-infrastructure.eu/content/download/8392/49194/version/2/file/ARIADNE_archaeological_LOD_study_10-2016.pdf], 2017/12/03.
- Gualtherus, G. (1624). *Siciliæ obiacentium insular. et Bruttiorum antiquæ tabulæ, cum animadversionib.* Messanae: apud Petrus Bream.
- Gulletta, M.I. (Ed.). (1999). *Sicilia Epigraphica. Atti del convegno internazionale, Erice, 15-18 ottobre 1998* (2 vols.). Pisa: Edizioni della Normale.
- Kaibel, G. (1890). *Inscriptiones Italiae et Siciliae* (= IG XIV). Berlin: Georgius Reimerus.
- Manganaro, G. (1988). La Sicilia da Sesto Pompeo a Diocleziano. *Aufstieg und Niedergang der römischen Welt*, 2.11.1, 3–89.
- Mommsen, T. (1883). *Inscriptiones Bruttiorum Lucaniae Campaniae Siciliae Sardiniae Latinae. Pars posterior. Inscriptiones Siciliae et Sardiniae* (= CIL X.2). Berlin: G. Reimer.
- Prag, J.R.W. (2002). Epigraphy by numbers: Latin and the epigraphic culture in Sicily. In A.E. Cooley (Ed.), *Becoming Roman, Writing Latin?* (Journal of Roman Archaeology Supplementary Series 48) (pp. 15–31). Portsmouth, RI.
- Prag, J.R.W. (2003). Nouveaux regards sur les élites locales de la Sicile républicaine. *Histoires et sociétés rurales*, 19, 121–132.
- Prag, J.R.W. (2007). Ciceronian Sicily: the epigraphic dimension. In J. Dubouloz & S. Pittia (Eds.), *La Sicile de Cicéron, Lectures des Verrines* (pp. 245–271). Besançon: Presses universitaires de Franche-Comté.
- Prag, J.R.W. (2008). Sicilia and Britiannia: Epigraphic Evidence for Civic Administration. In C. Berrendonner, M. Cébeillac Gervasoni, & L. Lamoine (Eds.), *Le Quotidien municipal dans l'Occident romain* (pp. 67–81). Clermont-Ferrand: Presses Universitaires Blaise-Pascal.
- Prag, J.R.W. (2010). Sicilia Romana tributim discripta. In M. Silvestrini (Ed.), *Le tribù romane. Atti della XVIe Rencontre sur l'épigraphie (Bari 8-10 ottobre 2009)* (pp. 305–311). Bari: Edipuglia.
- Tribulato, O. (Ed.). (2012a). *Language and linguistic contact in ancient Sicily*. Cambridge: Cambridge University Press.
- Tribulato, O. (2012b). *Siculi bilingues?* Latin in the inscriptions of early Roman Sicily. In Tribulato, O (Ed.), *Language and linguistic contact in ancient Sicily* (pp. 291–325). Cambridge: Cambridge University Press.