



From Fragmented Data to Linked History: Developing the FAIR Epigraphic Vocabularies

DISCUSSION PAPER

]u[ubiquity press

PETRA HEŘMÁNKOVÁ

IMRAN ASIF

MARIETTA HORSTER

JONATHAN R. W. PRAG

*Author affiliations can be found in the back matter of this article

ABSTRACT

The digital study of ancient inscriptions has undergone rapid growth. However, much of the data remains scattered across separate databases with varying levels of interoperability, presenting opportunities for improved integration that would facilitate large-scale analysis and enable researchers to draw broader connections. The *FAIR Epigraphy* project aims to overcome these challenges by developing a shared, community-driven framework of standardised vocabularies, terms and concepts for epigraphic data based on documented community needs. Following in the footsteps of the *EAGLE Europeana Project* and aligned with international authorities such as the *Getty Art & Architecture Thesaurus* and major epigraphic databases, the project provides stable, linked open data designed for interoperability across the semantic web.

The *FAIR Epigraphy* framework supports the integration of domain-specific controlled vocabularies with Wikidata to create richer, more accessible knowledge networks. Established vocabularies, such as those for inscription types and bilingualism, are already available as RDF Turtle files with stable URIs, with continuous expansion planned throughout 2025. The controlled vocabularies support consistent terminology across projects and allow researchers to query across databases and connect inscriptions to wider cultural heritage resources. By enabling more coherent and data-rich research, *FAIR Epigraphy* is opening new pathways for collaborative and computational approaches to studying the ancient world.

CORRESPONDING AUTHOR:

Petra Heřmánková

Alte Geschichte, Johannes
Gutenberg University in
Mainz, Mainz, Germany

petra.herbankova@uni-mainz.de

KEYWORDS:

digital epigraphy; CIDOC
CRM; linked open data;
FAIR principles; controlled
vocabularies; Wikidata

TO CITE THIS ARTICLE:

Heřmánková, P., Asif, I.,
Horster, M., & Prag, J. R. W.
(2025). From Fragmented Data
to Linked History: Developing
the FAIR Epigraphic
Vocabularies. *Journal of Open
Humanities Data*, 11: 74,
pp. 1–14. DOI: <https://doi.org/10.5334/johd.428>

(1) CONTEXT AND MOTIVATION

The proliferation of digital epigraphic projects over the past two decades has made hundreds of thousands of ancient inscriptions accessible online, alongside rich metadata about their physical carriers and provenance. However, several competing conceptualisations of epigraphic terminology have emerged across projects, making it challenging for researchers to query across databases or aggregate results from multiple sources (Bodel et al., 2024; Heřmánková et al., 2021; Tupman, 2021). The *FAIR Epigraphy* project,¹ an international collaboration of domain experts in epigraphy and digital humanities, addresses this challenge by developing a comprehensive framework for standardised epigraphic data representation using CIDOC CRM.² Bidirectional integration with Wikidata extends this framework's reach beyond the project's primary focus on Greek and Latin inscriptions, enabling connections with epigraphic materials across multiple writing systems and cultural contexts.³ Wikidata integration also enhances discoverability: researchers and cultural heritage projects can access FAIR Epigraphic vocabularies through Wikidata's infrastructure, whilst epigraphic databases can leverage Wikidata's multilingual labels and cross-references to other knowledge bases.

(1.1) THE FRAGMENTATION CHALLENGE IN DIGITAL EPIGRAPHY

Digital epigraphy has matured considerably since the introduction of EpiDoc, TEI XML tailored to record epigraphy in the early 2000s (Elliott et al., 2017–2025). Major projects such as the *Epigraphic Database Heidelberg*⁴ and the *Epigraphic Database Rome*⁵ have made significant portions of the epigraphic record computationally accessible. However, as documented in our 2022 scoping survey (Heřmánková et al., 2022), most projects employ project-specific data models, terminologies, and classification systems, often without providing machine-readable metadata, formal definitions, or alignment with existing standards. This heterogeneity creates substantial obstacles for researchers seeking to aggregate data sources or conduct large-scale quantitative analyses (Heřmánková et al., 2021; Laurence & Trifilò, 2023; Mullen & Willi, 2024). Different scholarly traditions have developed distinct approaches to categorising inscription types, material supports, and writing techniques. Without shared vocabularies and conceptual frameworks, even basic comparative questions require manual reconciliation of divergent classification schemes, a labour-intensive process that was initially undertaken by the international consortium *EAGLE Europeana Project*, running between 2013 and 2016, to enable cross-project data integration of epigraphy (Orlandi et al., 2014).

(1.2) THE FAIR IMPERATIVE AND COMMUNITY REQUIREMENTS

The FAIR principles (Wilkinson et al., 2016) have become increasingly central to research data management across humanities disciplines. Our 2022 community survey revealed that the majority of surveyed projects in epigraphy were familiar with FAIR principles, but their compliance was varied and depended heavily on the resources available (Heřmánková et al., 2022). Only a minority had implemented concrete measures for interoperability, such as providing data in RDF format or using standardised vocabularies. This gap between awareness and implementation resonates with the explicit, yet still rare, discussion of FAIR principles implementation challenges within the Text Encoding Community and, more specifically, in epigraphy (Creamer et al., 2021). The survey respondents identified controlled vocabularies as

¹ <https://inscriptions.org/> The *FAIR Epigraphy* project (2022–2026), co-directed by Marietta Horster (JGU/BBAW, Germany) and Jonathan Prag (Oxford, UK), provides access to FAIR and linked open data for epigraphy by creating digital tools and resources.

² The FAIR Epigraphic vocabularies are structured within a CIDOC CRM-based ontological framework that builds upon CRMtex 1.0 (Murano & Felicetti, 2020) and EpOnt (Bodard et al., 2021). Our implementation extends CIDOC CRM through CRMtex 2.0 (Murano et al., 2023), incorporating elements from CRMsci (Doerr et al., 2023) and LRMoo (Aalberg et al., 2024) to model the conceptual complexity of inscriptions as both linguistic texts and physical objects with rich study and publication histories. Initial user testing and engagement took place in early 2025, and the full ontology will be published in early 2026 via the *FAIR Epigraphy* website <https://inscriptions.org/> and Zenodo community <https://zenodo.org/communities/fairepigraphy/>.

³ Bidirectional integration establishes explicit links between *FAIR Epigraphic* vocabularies and Wikidata, enabling computationally executable data exchange. The concrete implementation using SKOS (Simple Knowledge Organisation System) properties to maintain conceptual complexity and nuance is described in the Section 3.4.2.

⁴ <https://edh.ub.uni-heidelberg.de/>.

⁵ <http://www.edr-edr.it/default/index.php>.

one of the most critical infrastructure gaps, emphasising that vocabularies must be developed collaboratively rather than imposed top-down. Alongside identifying this shortcoming, there was a strong desire to avoid limiting discussion about standard vocabularies to a small, exclusive group. Rather, respondents emphasised treating vocabulary development as a central undertaking affecting all projects and individual researchers, one that actively encourages community involvement throughout the process.

The adoption of Linked Open Data (LOD) approaches offers a pathway to FAIR compliance whilst preserving project autonomy. The *EAGLE Europeana Project* pioneered this approach (Liuzzo et al., 2013; Liuzzo, 2015, 2019), developing controlled vocabularies between 2013 and 2016 that established foundational terminology for inscription types, materials, object types, execution technique, dating criteria, decoration and state of preservation.⁶ Following their partial integration into Wikidata in 2015,⁷ these vocabularies provided crucial groundwork for interoperability in the field. After several years of use and implementation by major digital databases, active users identified the need to consolidate, extend, and align the EAGLE vocabularies with contemporary semantic web standards to meet current FAIR requirements and support bidirectional data exchange (Prag et al., 2022). The *FAIR Epigraphy* project emerged from this recognised need to update and expand the existing infrastructure.

(1.3) STRATEGIC ALIGNMENT

The FAIR Epigraphic vocabularies framework,⁸ aligned with existing digital resources, positions digital epigraphy within the broader landscape of cultural heritage linked data. Alignment with the *Getty Art & Architecture Thesaurus* (AAT)⁹ ensures compatibility with museum and archaeological data systems, whilst the CIDOC CRM-based ontology provides a robust conceptual framework for modelling complex relationships between inscriptions, objects, persons, places, and events.¹⁰ Additionally, the FAIR Epigraphic vocabularies maintain explicit alignment with the EAGLE vocabularies, which are partially present in Wikidata,¹¹ and with other Wikipedia items; for alignment examples, see Section 3.4.2. This multi-layered approach reflects the reality that inscriptions exist at the intersection of multiple domains – textual, material, historical, and spatial – and the conceptual differences between existing knowledge recording infrastructure.

(1.3.1) EAGLE in Wikidata

Property P1900 (Eagle ID) functions as an identifier linking Wikidata items to epigraphic concepts within the EAGLE vocabulary system. As of October 2025, Wikidata contains only a subset of the EAGLE vocabularies: P1900 comprises 698 items and seven qualifiers, representing 18.6% of the total 3,747 EAGLE vocabulary entries. Most entries were created via upload by Pietro Liuzzo from the EAGLE project in 2015 and were gradually expanded by Wikidata contributors, but they are far from being complete. As an early-stage integration effort, this work naturally exhibits some structural challenges: entries occasionally overlap or duplicate concepts, and hierarchical relationships and definitions remain underdeveloped, see Sections 3.2 and 3.4.2 for examples. These characteristics reflect the exploratory nature of this preliminary implementation, which nonetheless provides a valuable foundation for future enhancement of the EAGLE vocabulary and its incorporation within Wikidata.

(2) DATASET DESCRIPTION

The FAIR Epigraphic vocabularies are published as openly accessible linked open data resources, designed for computational analysis and semantic web integration. Building upon the EAGLE vocabularies, they address structural issues in existing resources by eliminating duplicate

6 <https://www.eagle-network.eu/resources/vocabularies/>.

7 <https://www.wikidata.org/wiki/Q19371183>.

8 <https://ontology.inscriptiones.org/>.

9 <https://www.getty.edu/research/tools/vocabularies/aat/> The Art & Architecture Thesaurus (AAT) is a structured vocabulary for describing and indexing the visual arts and architecture.

10 For details about the ontology, see footnote 2.

11 <https://www.wikidata.org/wiki/Q19371183>.

entries, establishing clear hierarchical relationships, and providing comprehensive definitions and usage examples for each concept. Rather than attempting to revise EAGLE vocabularies from 2015 retroactively, a thorough community consultation in 2020–2021 established consensus for developing enhanced vocabularies with stable URIs that maintain backward compatibility. This approach ensures that existing projects linking to EAGLE vocabularies can benefit from improvements whilst preserving their current integrations, and enables future projects to adopt more robust semantic foundations from the outset.

REPOSITORY LOCATION

The vocabularies are available at <https://ontology.inscriptiones.org/> with stable URIs for each vocabulary term. The website is hosted via GitHub at <https://github.com/FAIR-epigraphy/ontology> and maintained by the Centre for the Study of Ancient Documents (CSAD), University of Oxford. Individual sets of vocabularies are assigned their own DOIs for citation and persistent identification.

REPOSITORY NAME

Zenodo, FAIR Epigraphy community.¹²

OBJECT NAME

FAIR Epigraphic vocabularies, comprising individual controlled vocabularies for epigraphic data representation, e.g., type of inscription <https://doi.org/10.5281/zenodo.17454135> (Heřmánková et al., 2025), epigraphic bilingualism <https://doi.org/10.5281/zenodo.17569510> (Mullen, 2025).

FORMAT NAMES AND VERSIONS

RDF/Turtle (.ttl) files conforming to W3C standards.

CREATION DATES

Development began in 2022, first published in 2024, with ongoing expansion throughout 2025.

DATASET CREATORS

Individual vocabularies list their specific creators and contributors in their respective VoID (Vocabulary of Interlinked Datasets) metadata files, along with their ORCID identifiers. Petra Heřmánková is responsible for the development of vocabularies, Jonathan Prag for the development of the ontology and Imran Asif for the technical development.

LANGUAGE

English, with multilingual labels planned for future releases (fr, de, es, la). URI patterns and technical documentation in English.

LICENSE

Individual vocabularies within the framework are released under different open licenses depending on the originating project and authors, with details specified in the VoID metadata. Preference is given to the CC BY-SA 4.0 license.

PUBLICATION DATE

Initial vocabularies published 2024-04-01, with continuous updates.

SCOPE AND COVERAGE

The vocabularies currently cover inscription types (building on EAGLE typology) and classifications for bilingualism in epigraphy, see Section 3.3. Planned expansion includes

¹² <https://zenodo.org/communities/fairepigraphy/>.

writing techniques (12/25), dating criteria (12/25), object types (1/26) and materials (2/26). Each vocabulary term includes stable URIs, definitions, examples from major epigraphic databases, and mappings to EAGLE, *EpiVoc* (Brunet et al., 2022), Wikidata, and Getty AAT where applicable.

(3) METHOD

This Section describes the methodological approach employed to develop the FAIR Epigraphic vocabularies, from community requirements gathering through technical implementation and integration with external resources.

(3.1) COMMUNITY REQUIREMENTS AND COLLABORATIVE DEVELOPMENT

Development commenced with a comprehensive scoping survey conducted between February and April 2022, gathering responses from digital epigraphic projects across Europe and North America (Heřmánková et al., 2022). The survey identified controlled vocabularies as the most critical infrastructure gap, with projects emphasising that vocabularies must emerge from collaborative, bottom-up processes reflecting diverse scholarly traditions. The *Epigraphy.info* Vocabularies Working Group¹³ determined through community consultation between 2020 and 2021, that, rather than attempting to revise the completed EAGLE vocabularies, developing enhanced vocabularies with improved structure and backward compatibility would better serve both existing and future projects. The *FAIR Epigraphy* project implemented this community feedback by consolidating existing resources into an enhanced framework that addresses interoperability challenges within the discipline and enables integration with the broader semantic web ecosystem.

(3.2) VOCABULARY CONSOLIDATION AND EXTENSION

The EAGLE vocabularies,¹⁴ developed between 2013 and 2016 (Liuzzo et al., 2013; Liuzzo, 2015), provided crucial groundwork but required substantial consolidation and extension. The original EAGLE vocabularies suffered from structural limitations: a flat hierarchy that obscured conceptual relationships, numerous duplicate entries, and conceptual conflation with other vocabulary categories, e.g., graffiti or *instrumentum domesticum* as a type of inscription,¹⁵ instead of technique or type of inscribed object. Moreover, the lack of detailed definitions elucidating their scope created ambiguity for data creators (e.g., what constitutes a ‘*proskynema*’ inscription).¹⁶ Where applicable, we systematically reviewed each EAGLE vocabulary term, identifying and resolving duplicates whilst establishing hierarchical relationships using SKOS (Simple Knowledge Organisation System) properties, including broader, narrower, and related terms, to express hierarchical and associative connections. Terms were verified against usage in major epigraphic databases, reflecting the status in 2025, and where projects employed divergent terminology for equivalent concepts, we conducted alignment exercises to establish mappings based on close reading and evaluation of inscriptions and consultations with resource owners and maintainers. We expanded coverage to accommodate functional classifications underrepresented in the original scheme, particularly for multilingual contexts. Each term includes formal definitions in English, scope notes clarifying boundaries with related concepts, and additional verified usage examples drawn from published corpora.

(3.3) VOCABULARY CONTENT AND COVERAGE

The *Type of Inscription* vocabulary (Heřmánková et al., 2025) categorises inscriptions according to their primary societal function and content, encompassing 122 vocabulary terms organised under 15 top-level categories, ranging from official legal documents and religious dedications

¹³ https://epigraphy.info/vocabularies_wg/.

¹⁴ <https://www.wikidata.org/wiki/Q19371183>.

¹⁵ graffiti: <https://www.eagle-network.eu/voc/typeins/lod/370> and <https://www.eagle-network.eu/voc/typeins/lod/127>; instrumentum: <https://www.eagle-network.eu/voc/typeins/lod/136>.

¹⁶ EAGLE provides no definition: <https://www.eagle-network.eu/voc/typeins/lod/423> vs FAIR Epigraphic Vocabulary: https://ontology.inscriptiones.org/type_of_inscription/Proskynema “Inscriptions associated with acts of reverence, adoration, or worship, such as bowing or kneeling before a deity or sacred object.”

to commercial records and personal commemorations. The vocabulary prioritises functional categories over formal characteristics, enabling analysis of epigraphic practice across contexts and periods. Building upon *EAGLE Project* foundations (Liuzzo et al., 2013), it consolidates the original classification through systematic review of duplicates and establishment of hierarchical relationships, as well as comprehensive scholarship-based definitions.

The *Bilingualism in epigraphy* vocabulary (Mullen, 2025) classifies inscriptions indicating concurrent use of two or more languages. The 33 vocabulary items, structured into language-focused and script-focused subcategories, address a critical gap: whilst digital projects increasingly recognise bilingual features, systematic tagging using established schema has been lacking. Developed by Mullen, building upon Adams (Adams, 2003) and engagement with modern sociolinguistics scholarship, terms include bibliography and usage examples linked to the *LatinNow* webGIS database.¹⁷

Planned vocabulary releases throughout 2025 and early 2026 address further core epigraphic categories: Execution Technique (methods of physical text production), Inscribed Object (physical support typologies), and Dating Criteria (evidence types for chronological establishment).¹⁸ These vocabularies build upon the foundational *EAGLE* categories, applying the same consolidation methodology of removing duplicates, establishing hierarchical relationships, providing comprehensive definitions with usage examples, and maintaining SKOS-based alignments.

(3.4) MULTI-PLATFORM INTEGRATION STRATEGY

Linking several knowledge bases positions epigraphic data within multiple overlapping semantic networks. The FAIR Epigraphic vocabularies maintain explicit alignments with the *EAGLE* vocabularies, their Wikidata references, and Wikidata items not originated by *EAGLE*. This alignment using SKOS mapping properties (exactMatch, closeMatch, broadMatch, narrowMatch, relatedMatch) allows for the expression of precise semantic relationships, whilst providing enhanced structure and extended coverage beyond the original *EAGLE* terms. Alignment with reference thesauri, such as Getty AAT, or *EpiVoc* (Brunet et al., 2022), ensures compatibility with museum and archaeological data systems, particularly for materials, object types, and techniques. The result of predominantly manual evaluation of the concepts, these alignments use SKOS mapping properties to document semantic relationships between FAIR and related terms in other knowledge bases. This multi-vocabulary alignment strategy maximises discoverability and enables cross-domain queries bridging epigraphy, archaeology, art history, and museum collections, as well as improving transparency within the disciplines.

(3.4.1) Wikidata as a Central Knowledge Graph

A central component of our integration strategy is Wikidata, a free, collaborative, and multilingual knowledge graph that functions as a central repository for structured data (Vrandečić & Krötzsch, 2014). Its structure is fundamentally entity-centric, meaning every concept, person, place, or object is represented as a unique ‘item’ (identified by a ‘Q-number,’ such as Q121614747 for ‘funerary inscription’). These items are then linked by ‘properties’ (identified by ‘P-numbers’) to create simple, factual statements (Ripoll et al., 2025). This structure makes Wikidata a powerful tool for authority control and for linking together different datasets (Vrandečić & Krötzsch, 2014). The planned implementation of bidirectional workflow integrates the FAIR Epigraphic vocabularies’s URIs to matching Wikidata items via a new dedicated P property for authority control, see Section 3.4.2 for exact steps. This connection helps position our domain-specific terms within Wikidata’s vast, cross-disciplinary network, making them easier to find and use. That said, as we explore further in Section 4, Wikidata’s entity-centric approach poses some real challenges when we try to align it with our own event-based ontology, using the same SKOS-based alignment.¹⁹

¹⁷ <https://gis.latinnow.eu/>.

¹⁸ The *FAIR Epigraphy* platform (<https://ontology.inscriptiones.org/>) is open to host additional developed structured vocabularies related to digital epigraphy. For expression of interest, please get in touch with the authors.

¹⁹ Although mapping between CIDOC and Wikidata models is not always straightforward, several recent projects have demonstrated that transforming Wikidata data into CIDOC CRM structure is eminently feasible (Koch et al., 2024; Untner, 2025). Mapping is required only for integration across both models, not for independent data extraction.

(3.4.2) Bidirectional integration: ‘Type of inscription’ vocabulary

Bidirectional integration with Wikidata is currently in development. This bidirectional model will enable epigraphic projects to query FAIR Epigraphic vocabularies for standardised terms whilst accessing broader contextual information from Wikidata, or from external sources linked in Wikidata, for example, retrieving biographical data about honorands or geographical information about findspots. As a first step, we established inverse links in the FAIR RDF files pointing to corresponding Wikidata Q-numbers.

In order to illustrate the workflow and current state of development, we briefly review the coverage in Wikidata across the FAIR and EAGLE vocabularies, with specific reference to the ‘type of inscription’ vocabulary, one subset of the FAIR Epigraphic vocabularies. The vocabulary contains 334 items in EAGLE, of which 82 are present in Wikidata (24.5% coverage). The FAIR ‘type of inscription’ vocabulary contains 122 items; the lower number of items compared to EAGLE is a result of deduplication and reconceptualisation of some of the terms. Out of the 82 EAGLE items existing in Wikidata, we identified 40 as having an equivalent in the FAIR vocabulary. We linked them in the FAIR RDF file, using the SKOS properties `exactMatch` (0 terms), `closeMatch` (33), `relatedMatch` (29), `broadMatch` (1), and `narrowMatch` (29). Preference was given to wider mapping (`close`, instead of `exactMatch`), due to the insufficient definitional clarity in some of the existing EAGLE and Wikidata items. However, the brevity of textual definitions seems to be a common problem in many sources, plausibly due to their general scope.

For example, the type of inscription ‘Defixio’, a curse (tablet), has the following definition in EAGLE: “Also listed by Anne Glock from volumes II/7 (Corduba); VI 8, 2 and 3 (Rome); XIII (Gaul and Germania); XVII/4,1 (Milestones) in CIL.”²⁰ In Wikidata, the `relatedMatch` ‘curse tablet’²¹ has a definition “*type of votive tablet*”, which is technically not incorrect, but it neglects the magic aspect of the text. The FAIR Epigraphic vocabularies provide the following definition “*Curse tablets or binding spells, aimed at causing harm to others or seeking revenge. These thin lead tablets, often pierced with a nail and usually rolled up, contain curses of people or animals, written in cursive script and labelled with magical signs. They were found on riverbanks, in wells and especially in graves. The anonymous texts are not addressed to the public, but to chthonic deities such as Hecate or the dii inferi. Despite their formulaic wording, these tablets offer scope for personalisation [Schmidt 2004, 48–50].*”²² The need to develop a detailed description resulted from comprehensive community and literature scoping, which revealed no agreed-upon definitions or a clearly defined source of truth, despite the discipline’s centuries-long existence. Additionally, in this particular instance, we provided 15 examples of curse tablets from major online databases, including texts both in Greek and Latin, so users can make their own informed decision (see [Figure 1](#) for the full record).

Linking to Wikidata records beyond EAGLE, we have identified that 104 FAIR items have at least one or more distant equivalent, extending the coverage of FAIR in Wikidata to 85%: SKOS `exactMatch` (0 items), `closeMatch` (49), `relatedMatch` (64), `broadMatch` (3), `narrowMatch` (13). 18 FAIR ‘type of inscription’ items do not currently exist in Wikidata and will need to be created. These mostly relate to highly domain-specific categories, such as healing texts,²³ mortgage stones,²⁴ or *proskynema*²⁵ inscriptions. After additional mapping, we were unable to associate six out of 122 items with any existing vocabulary outside of Wikidata, including Getty AAT or EAGLE. In those cases, we have provided multiple examples from epigraphic databases and bibliographic citations, so that users can make informed decisions themselves, e.g., the case of boundary disputes.²⁶

Once the active development of new FAIR vocabularies is completed in early 2026, we plan to implement the second step of the bidirectional linking process in three stages:

20 <https://www.eagle-network.eu/voc/typeins/lod/76.html>.

21 <https://www.wikidata.org/wiki/Q1430557>.

22 https://ontology.inscriptiones.org/type_of_inscription/Defixio.

23 https://ontology.inscriptiones.org/type_of_inscription/Healing_text.

24 https://ontology.inscriptiones.org/type_of_inscription/Mortgage_stone.

25 https://ontology.inscriptiones.org/type_of_inscription/Proskynema.

26 https://ontology.inscriptiones.org/type_of_inscription/Boundary_dispute.

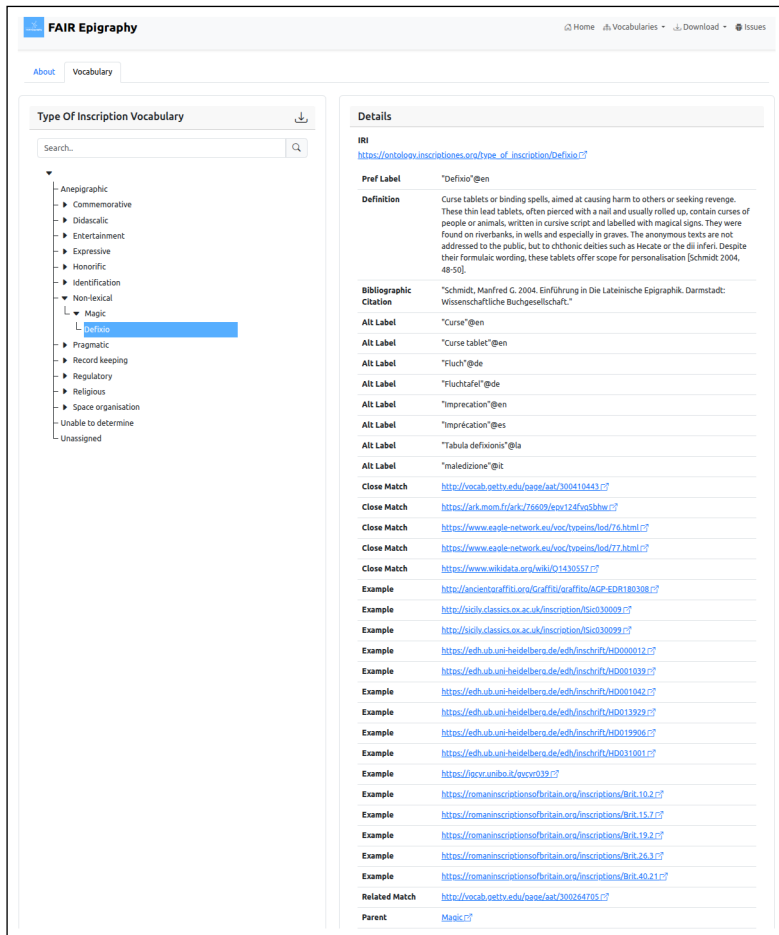


Figure 1 FAIR Epigraphic vocabularies entry for the 'Defixio', curse (tablet) type of inscription, showing SKOS-mapped aligned resources, 15 examples, bibliographic references, and translations with common variations of the term in multiple languages (English, German, Latin, Italian).

1. Propose a new authority control property (P) called FAIR Epigraphic vocabularies ID;
2. Link existing Wikidata records with FAIR Epigraphic vocabularies URIs, using the SKOS generic mapping relation property (P4390) with the following values: exact match (Q39893449), close match (Q39893184), broad match (Q39894595), narrow match (Q39893967), and related match (Q39894604);
3. Create new items (Q) for the FAIR Epigraphic vocabularies' terms not already represented in Wikidata.

(3.5) TECHNICAL INFRASTRUCTURE AND QUALITY CONTROL

The vocabularies are published as RDF/Turtle files conforming to W3C standards, with each term assigned a stable, dereferenceable URI following linked data best practices. The technical infrastructure comprises a GitHub repository for version control and collaborative development, a web interface providing human-readable documentation,²⁷ and VoID (Vocabulary of Interlinked Datasets) metadata documenting provenance, licensing, and creator information. Individual vocabularies receive their own DOIs to enable precise citation via Zenodo. Quality control encompasses both technical validation (automated checks for RDF syntax correctness, URI stability, and conformance to SKOS and CIDOC CRM specifications) and scholarly review by the domain experts (*Epigraphy.info* Vocabularies Working Group, *FAIR Epigraphy* partners). We maintain a transparent issue-tracking system via GitHub, enabling community members to report errors or propose modifications, with changes undergoing review by vocabulary creators before incorporation.

The practical use of FAIR Epigraphic vocabularies will be discussed in detail in a separate paper, but given that the EpiDoc guidelines²⁸ provide only suggestions, our scoping confirms that practice differs widely. For example, the EpiDoc guidelines suggest the following structure for inscription type, see [Figure 2](#).²⁹

27 <https://inscriptiones.org/>.

28 <https://epidoc.stoa.org/gl/latest/app-allidx.html>.

29 <https://epidoc.stoa.org/gl/latest/supp-desctexttype.html>.

```

<textClass>
  <keywords scheme="hgv" >
    <term> Brief (privat) </term>
  </keywords>
</textClass>

```

Figure 2 Structure suggested by the EpiDoc guidelines to record controlled vocabularies (i.e., type of inscription).

Not all projects employ controlled vocabularies; where they do, implementation varies considerably. The ideal scenario in [Figure 3](#) uses the suggested EpiDoc structure and maintains both the data structure and terminology of the original resource (*I.Sicily* in this case) and alignment with the controlled vocabulary of their choice (*FAIR Epigraphic vocabularies* in this case):

```

<textClass>
  <keywords scheme="https://ontology.inscriptiones.org/type_of_inscription" >
    <term ana="#function.funeraary" ref="https://ontology.inscriptiones.org/type_of_inscription/Funeraary">funeraary </term>
  </keywords>
</textClass>

```

Figure 3 An ideal implementation of controlled vocabularies in EpiDoc.

We strongly advocate that users provide links to their chosen controlled vocabulary, which represents their conceptual understanding, that may or may not necessarily align with the FAIR Epigraphic vocabularies, either directly in EpiDoc files or as attribute metadata in their databases, to ensure transparency and facilitate future research.

(4) RESULTS AND DISCUSSION

The vocabularies demonstrate robust technical implementation and interoperability with existing linked data infrastructures. Each term receives a stable URI enabling persistent citation and computational integration, with comprehensive metadata including English definitions, examples, scope notes clarifying usage boundaries, and bidirectional alignments to Wikidata. The implementation employs CIDOC CRM as its foundational framework, providing hierarchically organised categories with formal SKOS relationships expressing semantic connections between terms. Technical implementation demonstrates robust interoperability. SPARQL queries successfully retrieve vocabulary terms and traverse alignments to Wikidata entities, enabling researchers to combine local epigraphic classifications with external knowledge graphs. Getty AAT alignments enable cross-domain queries linking epigraphic data with museum collections and archaeological datasets.

However, implementation revealed substantial challenges, particularly regarding integration with Wikidata. The planned bidirectional integration proved complex due to fundamental structural asymmetries between the FAIR ontology's event-based CIDOC CRM framework and Wikidata's entity-centric model. As explained in Section 3.4.1, Wikidata is excellent at representing and linking clear, individual entities, such as an inscription, a person, or a findspot. Our framework, on the other hand, is designed to focus on the events and relationships that tie those entities together. For instance, we do not just document that an inscription exists; we model the entire writing event, i.e., who carved it, who commissioned it, who it was for, when and where it happened, and so on ([Bodard et al., 2021](#)) and the entire history of subsequent study and evaluation.

Wikidata does not currently offer the tools to fully capture these kinds of rich, layered relationships. It is not just a technical limitation; it is a difference in how the two systems understand and structure knowledge. As a result, syncing data between the two is not simply a matter of writing code. It is a process of conceptual translation and figuring out how to map the complexities of the two models. For this reason, the bidirectional mapping has to be done manually. That is not a failure of technology; it is an essential act of scholarly interpretation. But it does come at a cost. The process is time-consuming, and because both our FAIR Epigraphic vocabularies and Wikidata are always changing, there is a real risk of things falling out of sync over time. Maintaining alignment requires ongoing attention and effort.

When manually implementing the bidirectional mapping, we encountered three main problems. First, the EAGLE records in Wikidata lack connections between other existing Wikidata items, either creating duplicate concepts within the Wikidata knowledge base or missing a concept linkage. For example, the record 'funerary inscription' (Q121614747) created by EAGLE lists only *EAGLE Project* as its external resource. The record 'epitaph' (Q1772) is more widely used

and linked to at least 40 other external resources, but not to EAGLE. Second, whilst the EAGLE and FAIR vocabularies were created by domain specialists who receive attribution for their work, other Wikidata entries may have been created by specialists, enthusiasts, or contributors without detailed knowledge backgrounds, leading to inconsistent quality and reliability across the dataset. The varying depth of detail makes it difficult to determine whether linked terms constitute close or related matches, even for domain specialists. Therefore, when linking concepts across data sources, we preferred `relatedMatch` as a precautionary approach when in doubt, whereas others might have chosen `closeMatch`. Third, implementing the SKOS relational mapping of concepts in Wikidata is necessary yet non-trivial, and as it is not a widely adopted practice, the documentation is sparse. We have identified the necessary steps and describe them in Section 3.4.2.

The bottom-up approach of building vocabularies proved both essential and challenging. In the discussions between 2020 and 2022, the *Epigraphy.info* Vocabularies Working Group identified disagreements regarding term granularity, particularly for inscription types and between projects focused on specific time periods, geographical regions, or epigraphic traditions. We resolved this through a relatively generic design of the vocabulary based on domain expertise, detailed textual descriptions, multiple examples across the field, as well as iterative user testing and continuous refinement based on feedback. Moreover, the hierarchical design of FAIR Epigraphic vocabularies enables classification at appropriate levels of granularity, allowing multiple tags from the same vocabulary per record. To enhance the usefulness of the FAIR Epigraphic vocabularies, we consulted with numerous projects, including projects using the Wikidata schema, such as *IDEA* (Thornton et al., 2024b) or *Greek Metrical Inscriptions* (Ortimini, 2025). Limitations of the FAIR Epigraphic vocabularies remain in vocabulary coverage and its primary focus on Latin and Greek epigraphic cultures, restricting current utility for comprehensive epigraphic data modelling.

The Wikidata sources for epigraphy, however, are represented by more than just EAGLE. The recent efforts of WikiProjects *IDEA* (Thornton et al., 2024a, 2024b) and *Epigraphy*³⁰ testify to the need to improve the foundations made in 2015 by EAGLE and coordinate the proliferation of Wikidata resources related to epigraphy, as well as their relevance to non-Wikidata-based conceptual models, like the one of *FAIR Epigraphy*. The latest projects and activities at academic conferences suggest a dynamically expanding field (Azzolini, 2025; Ortimini, 2025),³¹ with potential to benefit all participants, regardless of choice of underlying conceptual model. For a comprehensive list of resources, see the WikiProject *Epigraphy*.³²

(5) IMPLICATIONS/APPLICATIONS

The FAIR Epigraphic Ontology & Vocabularies establish foundational infrastructure for interoperable epigraphic data, with implications extending beyond technical standardisation to reshape collaborative research practices. By providing shared semantic frameworks grounded in community consensus, the vocabularies enable genuinely comparative analysis across previously fragmented datasets. The *FAIR Epigraphy Browser*³³ demonstrates this potential as a proof of concept for linking distinct datasets via RDF. Researchers can now browse or formulate SPARQL queries spanning multiple epigraphic databases, retrieving all honorific inscriptions on marble, for instance, without first having to manually reconcile divergent classification schemes. This transforms the range of feasible research questions, particularly for computational approaches requiring large, structured datasets.

Despite ongoing development, FAIR Epigraphic vocabularies are already reflected upon in digital epigraphy and cultural heritage studies (Tamrazyan & Hovhannisyan, 2025), listed amongst resources and tools for digital epigraphy at the Wikibase instance *Greek Metrical Inscriptions*³⁴ (Ortimini, 2025), and referenced within specialised thesauri, e.g., *Thesaurus poésie*

30 https://www.wikidata.org/wiki/Wikidata:WikiProject_Epigraphy.

31 https://www.wikidata.org/wiki/Wikidata:WikiProject_Epigraphy/Linked_Pasts_11 in December 2025.

32 https://www.wikidata.org/wiki/Wikidata:WikiProject_Epigraphy and https://greek-metrical-inscriptions.wikibase.cloud/wiki/Main_Page.

33 <https://browser.inscriptions.org/>.

34 https://greek-metrical-inscriptions.wikibase.cloud/wiki/Main_Page.

épigraphique.³⁵ Individual vocabularies are adopted by partner epigraphic projects such as *I.Sicily* (Prag, 2017–2025) or *LatinNow* (Mullen et al., 2025; Mullen & Willi, 2024). Several major epigraphic projects plan future implementation of the vocabularies, such as *Roman Inscriptions of Britain* (RIB),³⁶ or *Carmina Epigraphica Latina Online* (CLEO).³⁷

The vocabularies' integration with Wikidata positions epigraphic data within broader knowledge graphs encompassing biographical, geographical, and historical information. Once properly implemented, an inscription mentioning a magistrate (Q126733465)³⁸ links not only to controlled vocabularies defining the term 'magistrate' (Q4594605) but to Wikidata entities documenting that individual's career, family connections, and historical context. As a practical example, serves linking of emperors in Wikidata with the *Inscriptions of Roman Tripolitania* dataset, done through the project's authority lists (Rouéche et al., 2022). This facilitates research bridging epigraphy with prosopography, historical geography, and social network analysis, enabling scholars to trace how individuals, families, and institutions employed inscriptions within broader communicative strategies.

For digital humanities infrastructure, the project demonstrates practical approaches to developing domain-specific ontologies within established frameworks. The methodology, community requirements gathering, consolidation of existing vocabularies, and alignment with external authorities, provides a replicable model for other subdisciplines confronting similar standardisation challenges. The experience highlights both the potential and limitations of integrating specialised scholarly vocabularies with general-purpose platforms like Wikidata, offering lessons for future linked open data initiatives. The distributed nature of linked open data presents both opportunities and challenges: whilst improvements to one dataset can theoretically propagate across interconnected resources, ensuring that changes cascade appropriately across multiple platforms requires careful coordination. Without robust version control and clear communication channels, updates risk creating temporary inconsistencies that may confuse users encountering different versions of aligned vocabularies across different platforms.

The adoption of linked open data principles offers substantial benefits to the epigraphic research community that justify the initial learning curve and implementation effort. By establishing machine-readable, interoperable vocabularies, researchers gain the ability to aggregate data across projects, discover previously hidden patterns in large corpora, and connect epigraphic evidence with broader historical and archaeological contexts. The investment in LOD infrastructure pays dividends through enhanced discoverability, reduced duplication of effort in data entry and classification, and the creation of sustainable digital resources that outlast individual projects. Training researchers in linked data principles through concrete epigraphic examples, using the vocabularies' transparent structure and comprehensive documentation, enables the community to understand how controlled vocabularies facilitate data integration whilst grasping the intellectual decisions underlying classification schemes. This supports critical data literacy development alongside traditional philological skills, equipping scholars to participate effectively in increasingly collaborative, digitally mediated research environments.

Ongoing development priorities through 2025 focus on expanding vocabulary coverage whilst strengthening technical infrastructure. Planned additions include comprehensive inscribed object vocabularies, execution technique classifications, materials, and dating criteria categories. The project's open, collaborative governance model via GitHub³⁹ ensures continued responsiveness to evolving scholarly needs through transparent issue-tracking and version control. Although the *FAIR Epigraphy* project concludes in February 2026, the infrastructure will continue to be maintained and developed through ongoing support from CSAD in Oxford and the efforts of the *Epigraphy.info* community, acknowledging that effective scholarly infrastructure emerges from sustained dialogue between diverse research communities rather than top-down standardisation.

35 <https://ark.mom.fr/ark:/76609/poe-j01pqj275rgkdbmwqhvw>.

36 <https://romaninscriptionsofbritain.org/>.

37 <https://institucional.us.es/cleo/>.

38 https://ontology.inscriptiones.org/type_of_inscription/Magistrates.

39 <https://github.com/FAIR-epigraphy/>.

DATA ACCESSIBILITY STATEMENT

Individual datasets are accessible under CC licenses at <https://ontology.inscriptiones.org/> and in the FAIR Epigraphy Zenodo Community <https://zenodo.org/communities/fairepigraphy/>.

GenAI DECLARATION

GenAI tools, such as Grammarly, were used for proofreading and copyediting of the manuscript.

ACKNOWLEDGEMENTS

We gratefully acknowledge the foundational work of the *EAGLE Europeana Project* in establishing controlled vocabularies for digital epigraphy. We are particularly grateful to Pietro Liuzzo, Silvia Evangelisti, and Alex Mullen for their ongoing support and sharing of resources. We extend our thanks to the *FAIR Epigraphy* partners and [Epigraphy.info](https://epigraphy.info) community, especially the members of the Vocabularies Working Group, for their invaluable contributions to the development of community-wide standards and their ongoing engagement in advancing FAIR principles in digital epigraphy. We thank the anonymous reviewers whose detailed comments substantially improved the manuscript.

FUNDING STATEMENT

This research was supported by the Arts and Humanities Research Council (AHRC) [grant number AH/W010682/1] and the Deutsche Forschungsgemeinschaft (DFG) [grant number 468455971].

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS


Petra Heřmáňková: Conceptualisation, Data curation, Methodology, Project administration, Resources, Validation, Writing – original draft, Writing – review & editing


Imran Asif: Data curation, Methodology, Software, Visualisation, Writing – original draft, Writing – review & editing


Marietta Horster: Conceptualisation, Funding acquisition, Methodology, Supervision, Writing – original draft, Writing – review & editing

Jonathan R. W. Prag: Conceptualisation, Funding acquisition, Methodology, Supervision, Writing – original draft, Writing – review & editing

AUTHOR AFFILIATIONS

Petra Heřmáňková  orcid.org/0000-0002-6349-0540
Alte Geschichte, Johannes Gutenberg University in Mainz, Mainz, Germany

Imran Asif  orcid.org/0000-0002-1144-6265
Faculty of Classics, University of Oxford, Oxford, United Kingdom

Marietta Horster  orcid.org/0000-0003-1434-224X
Alte Geschichte, Johannes Gutenberg University in Mainz, Mainz, Germany

Jonathan R. W. Prag  orcid.org/0000-0003-3819-8537
Merton College, University of Oxford, Oxford, United Kingdom

REFERENCES

Aalberg, T., Riva, P., & Zumer, M. (2024). *LRMoo: Object-oriented definition and mapping from the IFLA Library Reference Model (Version 1.0)* [Data set]. <https://cidoc-crm.org/lrmoo/ModelVersion/version-1.0>

- Adams, J. N. (2003). *Bilingualism and the Latin Language* (1st edn). Cambridge University Press. <https://doi.org/10.1017/CBO9780511482960>
- Azzolini, A. C. M. (2025). Altinum. A Wikidata project for Latin epigraphy and prosopography. *Epigraphy.Info IX*, Aarhus, Denmark. https://epigraphy.info/documents/workshop_9/Poster_8.pdf
- Bodard, G., Cayless, H., Cenati, C., Cooley, A. E., Elliott, T., Evangelisti, S., Felicetti, A., Granados, P., Grieshaber, F., Gruber, E., Hershkowitz, A., Hill, T., Kiiskinen, H., Kollatz, T., Levivier, A., Liuzzo, P. M., Luciani, F., Mannocci, A., Mataix, E., ... Weise, F. (2021). *Modeling Epigraphy with an Ontology*. <https://doi.org/10.5281/ZENODO.4639507>
- Bodel, J., Prag, J. R. W., & Roueché, C. (2024). Open scholarship: Epigraphic corpora in the digital age. In P. Fröhlich & M. N. Caballero (Eds.), *L'épigraphie au XXIe siècle. Actes du XVIe Congrès International d'Épigraphie Grecque et Latine, Bordeaux, 29 août-02 septembre 2022* (pp. 91–117). Ausonius.
- Brunet, M., Rousset, M., & Morlock, E. (2022). *EpiVoc: Thésaurus épigraphique* [Data set]. OpenTheso. <https://thesaurus.mom.fr/?idt=th61>
- Creamer, A., Lembi, G., Mylonas, E., & Satlow, M. (2021). Archiving a TEI Project FAIRly. *Journal of the Text Encoding Initiative, Issue 14*. <https://doi.org/10.4000/jtei.4324>
- Doerr, M., Kritsotaki, A., Rousakis, Y., Hiebel, G., Theodoridou, M., & Velios, Athanasios. (2023). *CRMsci: An Extension of CIDOC-CRM to support scientific observation* (Version 2.0) [Data set]. <https://cidoc-crm.org/crmsci/ModelVersion/version-2.0>
- Elliott, T., Bodard, G., Mylonas, E., Stoyanova, S., Tupman, C., & Vanderbilt, S. (2017–2025). *EpiDoc Guidelines: Ancient documents in TEI XML* (Version 9.7). <https://epidoc.stoa.org/gl/latest/>
- Heřmánkuvá, P., Asif, I., Prag, J., & Horster, M. (2025). *FAIR Epigraphic Vocabularies: Type of Inscription* (Version 1.0.0) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.17454135>
- Heřmánkuvá, P., Horster, M., & Prag, J. R. W. (2022). *Digital Epigraphy in 2022: A Report from the Scoping Survey of the FAIR Epigraphy Project* (Version v1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.6610696>
- Heřmánkuvá, P., Kaše, V., & Sobotkova, A. (2021). Inscriptions as data: Digital epigraphy in macro-historical perspective. *Journal of Digital History, 1*(1), 99–141. <https://doi.org/10.1515/jdh-2021-1004>
- Koch, I., Ribeiro, C., Poveda-Villalón, M., Rico, M., & Teixeira Lopes, C. (2024). Enriching Archival Linked Data Descriptions with Information from Wikidata and DBpedia. In A. Antonacopoulos, A. Hinze, B. Piwowarski, M. Coustaty, G. M. Di Nunzio, F. Gelati, & N. Vanderschantz (Eds.), *Linking Theory and Practice of Digital Libraries* (Vol. 15177, pp. 396–412). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-72437-4_23
- Laurence, R., & Trifilò, F. (2023). *Mediterranean Timescapes: Chronological Age and Cultural Practice in the Roman Empire* (1st edn). Routledge. <https://doi.org/10.4324/9781315267708>
- Liuzzo, P. M. (2015). EAGLE and EUROPEANA: Architecture Problems for Aggregation and Harmonization. Symposium on Cultural Heritage Markup, Washington, DC. <https://doi.org/10.4242/BalisageVol16.Liuzzo01>
- Liuzzo, P. M. (2019). *Digital Approaches to Ethiopian and Eritrean Studies* (1st edn). Harrassowitz, O. <https://doi.org/10.2307/j.ctvrnfr3q>
- Liuzzo, P. M., Fasolini, D., & Rocco, A. (2013). *Content harmonisation guidelines, including GIS and terminologies—Second Release (D 2.2.2, version n 4.0)* [Deliverable: Europeana network of Ancient Greek and Latin Epigraphy]. https://www.eagle-network.eu/wp-content/uploads/2013/06/EAGLE_D2.2.2_Content-harmonisation-guidelines-including-GIS-and-terminologies-Second-Release.pdf
- Mullen, A. (2025). *FAIR Epigraphic Vocabularies: Bilingualism in epigraphy* (Version 1.0.0) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.17569510>
- Mullen, A., & Willi, A. (Eds.) (2024). *Latinization, Local Languages, and Literacies in the Roman West*. Oxford University Press. <https://doi.org/10.1093/9780191994760.001.0001>
- Mullen, A., Vanderbilt, S., Willi, A., Houten, P., & Stoyanova, S. (2025) 'LatinNow Project Data'. Version 0.1. November 14; GitHub, <https://github.com/LatinNow/LatinNowProjectData>
- Murano, F., & Felicetti, A. (2020). *Definition of the CRMttx: An Extension of CIDOC CRM to Model Textual Entities* (Version 1.0) [Data set]. <https://cidoc-crm.org/crm-ttx/ModelVersion/version-1.0>
- Murano, F., Felicetti, A., & Fafalios, P. (2023). *Definition of the CRMttx: An Extension of CIDOC CRM to Model Textual Entities* (Version 2.0) [Data set]. https://cidoc-crm.org/sites/default/files/CRMttx_v2.0_June_2023.pdf
- Orlandi, S., Santucci, R., Casarosa, V., & Liuzzo, P. M. (Eds.). (2014). *Information Technologies for Epigraphy and Cultural Heritage: Proceedings of the First EAGLE International Conference*. Sapienza Università Editrice. <https://www.eagle-network.eu/wp-content/uploads/2015/01/Paris-Conference-Proceedings.pdf>
- Ortimini, P. (2025). Wikibase for epigraphy: The instance Greek Metrical Inscriptions. *Epigraphy.Info IX*, Aarhus, Denmark. Zenodo. <https://doi.org/10.5281/zenodo.15102921>
- Prag, J. R. W. (2017–2025). *I.Sicily: Inscriptions of Ancient Sicily*. <http://sicily.classics.ox.ac.uk>

- Prag, J. R. W., Horster, M., Hermankova, P., & Asif, I. (2022). *FAIR Epigraphy: Too many databases? Or just not enough linked data?* Zenodo. <https://zenodo.org/record/7732691>
- Ripoll, M., Reeves, N., Kurteva, A., Simperl, E., Peñuela, A. M., & Diepold, K. (2025). *Filling in the Blanks? A Systematic Review and Theoretical Conceptualisation for Measuring WikiData Content Gaps* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2505.16383>
- Roueche, C., Vagionakis, I., & BODARD, G. (2022). *Inscriptions of Roman Tripolitania 2021: The authority lists* (p. 892946 Bytes) [Data set]. King's College London. <https://doi.org/10.18742/19182014.V1>
- Tamrazyan, H., & Hovhannisyan, G. (2025). Cultural Categorization in Epigraphic Heritage Digitization. *Heritage*, 8(5), 148. <https://doi.org/10.3390/heritage8050148>
- Thornton, K., Seals-Nutt, K., & Chen, A. (2024a). Dura-Europos Stories: Developing interactive storytelling applications using knowledge graphs for cultural heritage exploration. *Semantic Web*, 1–15. <https://doi.org/10.3233/SW-243552>
- Thornton, K., Seals-Nutt, K., & Chen, A. (2024b). Encoding Archaeological Data Models as Wikidata Schemas: Utilizing Shape Expressions to Structure Collaborative Linked Open Data for Digital Storytelling Within the International Dura-Europos Archive. *The International Journal of Technology, Knowledge, and Society*, 21(1), 69–83. <https://doi.org/10.18848/1832-3669/CGP/v21i01/69-83>
- Tupman, C. (2021). Where Can Our Inscriptions Take Us? Harnessing the Potential of Linked Open Data for Epigraphy. In I. Velasquez Soriano & D. Espinosa Espinosa (Eds.), *Epigraphy in the Digital Age: Opportunities and Challenges in the Recording, Analysis and Dissemination of Inscriptions* (pp. 115–128). Archaeopress. <https://doi.org/10.2307/j.ctv1xsm8s5.15>
- Untner, L. (2025). *Wikidata to CIDOC CRM (wiki2crm)* (Version 1.0) [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.17396141>
- Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledgebase. *Communications of the ACM*, 57(10), 78–85. <https://doi.org/10.1145/2629489>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>

TO CITE THIS ARTICLE:

Heřmánková, P., Asif, I., Horster, M., & Prag, J. R. W. (2025). From Fragmented Data to Linked History: Developing the FAIR Epigraphic Vocabularies. *Journal of Open Humanities Data*, 11: 74, pp. 1–14. DOI: <https://doi.org/10.5334/johd.428>

Submitted: 27 October 2025

Accepted: 25 November 2025

Published: 17 December 2025

COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Journal of Open Humanities Data is a peer-reviewed open access journal published by Ubiquity Press.