



Editorial. Wikidata Across the Humanities: Datasets, Methodologies, Reuse

COLLECTION:
WIKIDATA ACROSS
THE HUMANITIES:
DATASETS,
METHODOLOGIES,
REUSE

EDITORIAL

]u[ubiquity press

MAXIME GUÉNETTE

SABINE VON MERING

CAMILLO CARLO PELLIZZARI DI SAN GIROLAMO

ALAN ANG

ANDREA FARINA

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

ABSTRACT

In this editorial, we present *Wikidata Across the Humanities: Datasets, Methodologies, Reuse*, a special collection dedicated to the publication of discussion and data papers that explore the use of Wikidata in the digital humanities. With 45 articles, it is by far the largest special collection in the *Journal of Open Humanities Data*. The collection deliberately spans a wide range of disciplines, from history and linguistics to archaeology and museology, which demonstrate the growing prominence and maturity of Wikidata as a knowledge base that is increasingly used not only in the humanities and social sciences, but also in the cultural heritage sector and GLAM institutions. The success of this collection reflects not only the adoption and dissemination of free and open knowledge, but also the research community's interest in integrating methods and tools that enable the interoperability of existing datasets – an issue that has become increasingly important in recent years. We hope this collection will encourage further researchers to contribute to Wikidata and similar collaborative platforms, not as a peripheral activity, but as an integral part of how humanities data is built and shared.

CORRESPONDING AUTHOR:

Maxime Guénette

Department of History,
Université de Montréal,
Montréal, Canada

maxime.guenette@umontreal.ca

KEYWORDS:

Wikidata; Wikibase; digital humanities; Linked Open Data; research infrastructure

TO CITE THIS ARTICLE:

Guénette, M., von Mering, S., Pellizzari di San Girolamo, C. C., Ang, A., & Farina, A. (2026). Editorial. Wikidata Across the Humanities: Datasets, Methodologies, Reuse. *Journal of Open Humanities Data*, 12: 136, pp. 1–13. DOI: <https://doi.org/10.5334/johd.624>

(1) CONTEXT AND MOTIVATION

Wikidata has emerged as one of the most significant infrastructures for Linked Open Data (LOD) in the contemporary knowledge landscape, serving as a semantic backbone capable of connecting datasets across disciplines, institutions, and languages (Farda-Sarbas & Müller-Birn, 2019; Neubert, 2017). For humanities research in particular, where data are often heterogeneous, historically contingent, and dispersed across archives and collections, Wikidata's capacity to provide stable identifiers, rich metadata, and interoperable links offers a way to make such data findable and reusable across projects. The Journal of Open Humanities Data (JOHD) occupies a distinctive position at the intersection of open scholarship and humanities research. Through the publication of peer-reviewed data papers, JOHD promotes the visibility, accessibility, and long-term usability of humanities datasets. Its editorial principles (Wigdorowitz et al., 2024) hold published data to a standard of transparency and reproducibility while keeping it freely accessible — values that align closely with those underpinning Wikidata as a LOD hub.

Yet, an analysis of JOHD's publication records between 2015–2024, presented at the *Wikidata and Research* conference in 2025 (Farina & McGillivray, 2026), showed limited integration between Wikidata and JOHD: of the 198 articles published at the time, only 13 – fewer than 7% – made any use of Wikidata. Across the publications examined, three distinct patterns emerged. In some cases (e.g., Besnier & Mattingly, 2021; Giovannini & Gallucci, 2024; Röttgermann, 2024), Wikidata was directly embedded in the dataset itself, functioning either as a knowledge base or as a source for textual data (e.g., via Wikisource). In others (e.g., Mambrini & Passarotti, 2023), authors acknowledged the potential for Wikidata integration without having pursued it, typically in the reuse potential sections of their data papers. Finally, some papers (e.g., Farina, 2023) neither used nor referred to Wikidata, even though their datasets offered obvious possibilities for integration and linking. These cases matter for the collection's argument: they suggest that Wikidata's limited footprint in humanities data publishing is not necessarily about potential — it is often a lack of awareness of Wikidata as a tool for linking, enriching, and reusing humanities datasets. That said, the analysis did show a positive trend—the awareness of and adoption of Wikidata among submissions to JOHD have been increasing over time, particularly in the fields of History and Literary Studies. JOHD is therefore particularly relevant as a site for this intervention, not because its publications are necessarily representative of humanities publishing as a whole, but because its focus on datasets, data documentation, reuse, and open scholarship makes it especially well aligned with the affordances of Wikidata. The analysis presented at *Wikidata and Research* also demonstrated that there was already a body of JOHD research whose datasets could potentially benefit from greater integration with Wikidata. The special collection was conceived as an opportunity to build on this existing alignment: to raise awareness of Wikidata among humanities data authors, to demonstrate concrete forms of integration, and to encourage researchers to consider the linkability and reuse potential of their datasets from the outset.

It is against this background that the special collection *Wikidata Across the Humanities: Datasets, Methodologies, Reuse* was conceived. Born directly from the conversations and collaborations that emerged at *Wikidata and Research* in 2025, it brings together contributions that explore, demonstrate, and critically reflect on the relationship between Wikidata and humanities data publishing. The collection has a dual purpose: to show the Wikidata community the value of structured data publication and peer-reviewed documentation, and to show the humanities data community how Wikidata can amplify the impact of their work. In doing so, the collection aims not simply to increase the number of publications that explicitly use Wikidata, but also to raise awareness of its potential as an infrastructure for making humanities datasets more interoperable, discoverable, and reusable. It therefore represents a focused intervention within JOHD, while contributing to the broader effort to strengthen connections between Wikidata and humanities data practices. Its success is assessed here by the number of contributions it has brought together, but also by their disciplinary range, the different ways in which they engage with Wikidata, and the extent to which they address broader questions of interoperability, data modelling, reuse, sustainability, and inclusivity.

(2) DESCRIPTION

The special collection introduces a wide variety of papers on diverse subjects of the humanities such as linguistics, history, museology, or archaeology (Figure 1). It focuses on datasets mainly

set in the Global North (mostly North America and Europe) and discusses recurrent themes and challenges in Wikidata and LOD ecosystems, such as data modelling, ontologies, sustainability in digital humanities, and digital citizen science. The core argumentation of the papers is to showcase how Wikidata can contribute to free and open knowledge in digital humanities, and help its users publish data that can be FAIR (Findable, Accessible, Interoperable, Reusable; [Wilkinson et al., 2016](#)).

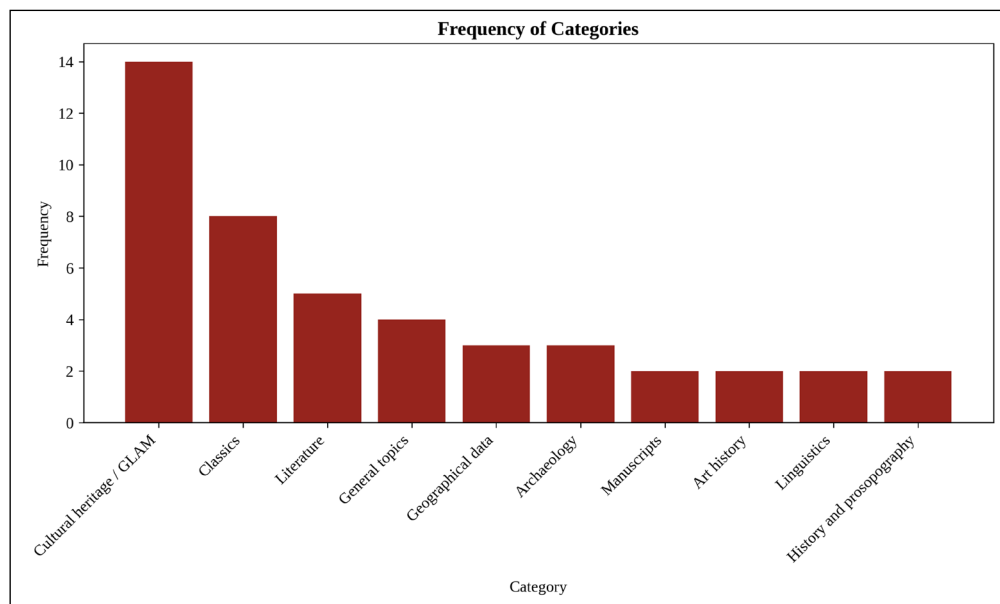


Figure 1 Number of articles across categories in the editorial collection.

Eight papers are dedicated to the field of Classics, ranging from literary studies and philology to prosopography and epigraphy. Berti (2025) shares her work with the Linked Ancient Greek and Latin (LAGL) project, especially the Catalog of Authors and Works: this dataset uses CTS URNs to link authors through a Wikidata property (P12869), enabling users to visualise and query geographic associations and intellectual activities of Classical authors while evaluating Wikidata's flexibility for Named Entity annotations. Also investigating ancient authors, Ripoll-Alberola et al. (2026) explore how Wikidata can be used as a compiling tool to create canonical lists of Ancient Greek and Latin authors. The core of their study lies in analysing how Wikidata is effective for identifying name variations and working as a data reconciliation standard and stable meta-authority across different types of corpora such as modern articles, 20th-century French press, and Early Modern print. Looking more closely at a single author, Blair (2026) examines the utility and challenges of using Wikidata to study the Latin lexicon *De uerborum significatu* by M. Verrius Flaccus and its integration with other linguistics tools and databases like LiLa or Ontolex. Similarly focused on a single text, La Veglia (2026) presents a digital edition of Angelo Poliziano's marginal notes to the Epistle of Sappho to Phaon. He is using the Text Encoding Initiative (TEI) to mark up the text while Wikidata URIs are added to tag Named Entities. Turning from texts to individuals, Fantoli et al. (2026) explore how Wikidata could be used as a knowledge graph for historical Greco-Roman individuals in parallel with other prosopographical resources like Trismegistos, MANTO and Prosopographia Imperii Romani. Through SPARQL queries and cross-search with the *Realencyclopädie der classischen Altertumswissenschaft* (RE), ca. 30,000 individuals were retrieved with variable spatio-temporal extent and accuracy, thus leading to future work that would strengthen prosopography in Wikidata. Gautschy (2025) introduces the Digital Lexicon Iconographicum Mythologiae Classicae (Digital LIMC), a dataset of more than 1,650 mythological figures, archaeological sites, modern places, and cultural institutions linked with Wikidata through a Wikidata property (P13094). In parallel with other external databases such as Pleiades and Ubi Erat Lupa, ongoing efforts to link more items demonstrate how cross-referencing with external datasets strengthens Wikidata's coverage of the ancient world. Moving from mythology to epigraphy, Kubiak-Schneider (2025) introduces a new dataset of 43 Ancient Greek and Middle Aramaic inscriptions from the Temple of Atargatis in Dura-Europos. As part of the International (Digital) Dura-Europos Archive project, Wikidata is used to publish these inscriptions as LOD while exploring the challenges of modelling digital epigraphic sources. Finally, Heřmánková et al. (2025) introduce the FAIR Epigraphy project,

which is developing a shared, community-driven framework of standardised vocabularies, terms and concepts for epigraphic data. Following the EAGLE project, FAIR Epigraphy uses Wikidata as an authority control platform to anchor its controlled vocabularies within a broader cross-disciplinary knowledge graph, enabling SPARQL queries that connect epigraphic classifications to biographical, geographical, and historical data held in external sources.

Three further papers ask how Wikidata can model material culture without flattening its heterogeneity — through interoperable ontologies, Wikibase instances, or as a linking hub for external databases and datasets. Chen et al. (2026) examine how Wikidata enables collaborative, multilingual cultural heritage data modelling while supporting interoperability with existing ontologies such as CIDOC CRM. Thiery et al. (2026) present three case studies where two of them — Irish Ogham stones and Irish holy wells — are modelled directly in Wikidata, while the third, on Campanian Ignimbrite volcanic findspots, runs on a purpose-built “fuzzy-sl Wikibase” linked back to Wikidata through shared identifiers. Together, these case studies argue that Wikidata functions not merely as a data source but as a semantic and social infrastructure, and that heterogeneous datasets across archaeology, epigraphy, and geoscience can be interlinked without being centralised. Finally, Mees and Thiery (2026) explore how Wikidata can serve as a secondary database and linking hub for the Linked Open Samian Ware project, which builds on Samian Research, a database of 250,000 Roman potters’ stamps from over 4,000 sites. It explores issues of data modelling, quality assurance, and FAIRification to strengthen Citizen Science integration and data interoperability at LEIZA, drawing on standards such as CIDOC CRM.

The following five papers explore the extent of literary corpora that can be linked with Wikidata’s data model — from importing corpora directly into it, to building parallel Wikibase representations, to extracting structured data back out, to avoiding the use of Wikidata altogether because its notability criteria exclude uncatalogued material. They engage with corpora ranging in scale from a German-language fiction collection of over 4,000 works to a single set of theatrical productions. Rohrbacher and Schritteser (2026) discuss a case study regarding the integration into Wikidata of a large literary corpus, specifically a corpus about German-language fiction and non-fiction from 1780 to 1930, with 4,215 literary works by 1,064 authors. The study describes the process of uploading the data into Wikidata, focusing on the tools which were used and on the challenges encountered. Wolter (2026) is using three productions of Ferdinand von Schirach’s *Terror* to build a Wikibase data model capable of representing scenography, costumes, and staging with sufficient granularity to support structured SPARQL queries over otherwise ephemeral performance data. De Monaco (2025) describes a dataset built from a corpus of 493 pseudonymous Italian authors, migrated into Wikidata from a biographical dictionary; the dataset includes a dump of the items edited in the import and of the results of the queries on the imported items. Where the previous studies focus on importing data into Wikidata, Untner (2025) works in the other direction, presenting a modular data model and Python-based workflow to extract bibliographic data from Wikidata and convert them into an ontology based on CIDOC CRM for authors, on IFLA LMRoo for works, and on INTRO (the Intertextual, Interpictorial, and Intermedial Relations Ontology) for (inter)textual phenomena. The ALMEDA project, presented by Aangenendt et al. (2025), makes the case for choosing Wikibase over Wikidata altogether, because the project “requires a robust ontology to repeal the effects of colonial library standards on literary description.” The authors argue that Wikidata’s open editability cannot itself guarantee data veracity, and that the notability criteria Wikidata uses to compensate — requiring verifiable external references — are difficult to satisfy for ALMEDA’s less formal, previously uncatalogued materials. A private Wikibase instance lets ALMEDA build its own multilingual ontology for African literary materials that “often go uncatalogued in libraries and archives,” while still aligning its structure with Wikidata where possible to enable federated search in the future.

Two papers examine the same project, Digital Scriptorium (DS), at different levels of detail. Jones (2026) presents the project, a digital catalogue of pre-modern and early modern manuscripts with over 50 member institutions; the LOD part of the catalogue, interconnected with Wikidata, can be explored through Wikidata SPARQL queries. McCandless and Coladangelo (2025) take a closer look at the database and focus specifically on its efforts to model pre-modern manuscripts in Wikidata. After the necessary ontology refinements in Wikidata, a repeatable and scalable workflow has been developed for transforming metadata from DS into Wikidata statements.

Two papers in linguistics ask how language-related data can be validated before entering the LOD framework. Samo and Ursini (2025) present a dataset of geonyms extracted from Wikidata, useful for cross-linguistic research. The data extracted from Wikidata are geographically validated using OpenStreetMap (OSM), a community-edited map database. In a similarly structured pipeline, Lindemann et al. (2025) describe the steps to encode in a Wikibase instance, interlinked with Wikidata, the inflected forms of over 8,000 Latin verbs. This pipeline allows for integrating Latin inflected forms and their attestations in the LOD framework.

Three papers are interested in art history, testing Wikidata's reliability for visual and biographical data as a benchmark against other databases, as structured input for vision-language models, and as a standard for cataloguing fellowship recipients. De Beyssat (2025) investigates the completeness and the accuracy of Wikidata items regarding the artists who exhibited at the Paris Salon between 1848 and 1880. Wikidata shows some structural tendencies that are more broadly present in historiography; at the same time, it is sometimes more accurate than professional art history databases, thanks to its volunteer-based model. Shifting from evaluation to method, Schneider et al. (2026) present a pipeline that integrates Vision-Language Models (VLMs) with structured metadata from Wikidata, using visual artworks having Iconclass notation (P1257) that represent identifiable concepts in this standardised art-historical vocabulary. Turning to a more applied case, Westerby et al. (2025) present the work done in the Scholars Data Project to put together the data about four residential programs in art history and visual arts in the United States: Wikidata is used to standardise the affiliations of the fellowship recipients.

Two papers deal with spatial data, addressing geographical data directly. Lamar (2025) presents a workflow that applies Wikidata-based geocoding directly to plain-text files. He proceeds in two phases; firstly Named Entity Recognition (NER), and then finding the relevant Wikidata item and extracting its coordinates. Where Lamar works from unstructured plain text, Munke (2026) starts from an already structured source: a bibliography of travel reports in Saxony from the late Middle Ages and the early modern period, which has been firstly converted in a closed database and secondly exported semi-automatically to Wikidata.

The next three papers discuss historical and prosopographical data. Baránek (2026) explains the workflow to import into Wikidata the data contained in a biographical dictionary, the *Biographisches Lexikon zur Geschichte der böhmischen Länder* (BLGL); and for items that already exist in Wikidata, comparing Wikidata's statements against the data drawn from the BLGL. Taking a different approach, Dilley (2026) studies 310 Wikidata items regarding writing systems, concluding that Wikidata is presently "the best open-source linked dataset on the topic," despite significant gaps and inconsistencies. Returning to the question of data import, Kim-Baumann (2026) presents a workflow for integrating historical data into Wikidata, specifically a dataset regarding Berlin's public transport network during the Cold War; it also proposes strategies compatible with Wikidata's existing standards to model discontinuities in operation periods and geographical uncertainty.

Thirteen papers focus on cultural heritage and GLAM institutions (Galleries, Libraries, Archives, and Museums), making this one of the largest thematic areas represented in the special collection. Collectively, these contributions demonstrate how Wikidata has evolved into a collaborative infrastructure for documenting collections, modelling cultural knowledge, improving interoperability, and making heritage resources more accessible and reusable across institutional and disciplinary boundaries. Together, they span conceptual discussions, practical workflows, large-scale data integration, and community infrastructures that together illustrate the breadth of Wikidata's role in the GLAM sector.

Four papers in the special collection address the conceptual and organisational foundations of Wikidata adoption in the cultural heritage sector. First, Bemme and Munke (2026) examine the rapid improvement of open metadata for German archives in Wikidata following a community call for contributions. They show how collaboratively curated archival metadata can strengthen the visibility, interoperability, and networking of archives in Saxony and neighbouring regions. Building on this question of interoperable infrastructure, Meiners and Bulle (2025) explore the strategic role of external identifiers in connecting heritage databases and argue for treating identifier properties as an essential component of sustainable, interoperable metadata rather than merely technical attributes. Similarly concerned with structural consistency, Mathios et al. (2026) examine inconsistencies in Wikidata's bibliographic domain and argue that aligning its ontology more closely with established library models such as IFLA LRM would improve

consistency and increase Wikidata's reuse across GLAM applications. Their discussion highlights ontology refinement as a prerequisite for reliable knowledge graph integration. Turning from ontology to practice, Dişli et al. (2025) assess how art museums adopt open data best practices in Wikidata, benchmarking compliance with criteria such as machine-readable data, licensing clarity, and API access. Their study provides practical guidance for institutions seeking to improve the openness and reusability of their cultural heritage data.

The next group of papers explores these questions of metadata openness and infrastructure in collection-based case studies. Akhlaq et al. (2026) use LOD to recover the often-overlooked contributions of African workers involved in the Tendaguru excavations, demonstrating how Wikidata can support more inclusive representations of museum collections and their colonial histories. Campos Castelo Branco and von Mering (2025) reconstruct colonial film expeditions to the Brazilian Amazon by linking expeditions and associated individuals to films, specimen-holding collections, archival records, and publications, illustrating the potential of Wikidata to bridge film history, natural history, and provenance research. In a related effort to recover marginalised histories, Sorensen et al. (2026) demonstrate how Wikidata can serve as an authority identifier to reconcile records across distributed Smithsonian collections, improving the representation, visibility, and discoverability of women in cultural heritage data. Complementing these examples, Foster et al. (2026) examine the reconciliation of museum collection management records with Wikidata using OpenRefine, drawing on the Index of American Design at the National Gallery of Art. Their paper highlights the methodological challenges of aligning local and Wikidata identifiers while demonstrating how linked data workflows and visualisation tools can support ongoing data quality and collection stewardship. Shaw et al. (2026) examine the migration of volunteer-curated heritage websites into Wikidata, showing how fragile web-based cultural datasets can be preserved, structured, and reused through collaborative digitisation efforts involving students and Wikimedia UK.

Other contributions focus on the integration of institutional data at national scale. Pensa et al. (2026) demonstrate how coordinated workflows can support the large-scale integration of museum data into Wikidata. Their paper presents a scalable approach for engaging GLAM institutions in contributing structured, openly accessible information, thereby strengthening both the visibility of museums and the broader LOD landscape. Complementing this work, Roller et al. (2026) describe the integration of data on more than 11,000 Italian libraries into Wikidata, highlighting the challenges of data reconciliation, entity alignment, and iterative human curation in building a comprehensive national knowledge graph. Finally, two papers broaden the perspective to the infrastructure underpinning these activities. Sichani et al. (2026) present the Congruence Engine Data Register as a community-driven infrastructure for connecting distributed cultural heritage collections, demonstrating how Wikibase and Wikidata can support sustainable, interoperable, and inclusive digital heritage infrastructures. Sohmen et al. (2026) make the case for a division of labour between Wikidata and Wikibase by reporting on TIB's Open Science Lab work reconciling and federating data between three project-specific Wikibase instances and Wikidata: Manor Houses in the Baltic Sea Region, the Corpus of Baroque Ceiling Paintings (hosted on the Semantic Kompakkt Wikibase), and the Federation of German-Speaking Architectural Collections. Their workflows rely on OpenRefine for reconciliation and federated SPARQL queries for retrieval, developed through the Wikibase4Research service for NFDI4Culture.

Beyond the domain-specific contributions, four papers examine the infrastructures, workflows, and methodological foundations needed to keep humanities data usable well after a project ends. Together, they highlight Wikidata not simply as a database, but as part of a broader ecosystem for integrating, curating, and disseminating knowledge across domains.

Gotham (2026) examines the opportunities and trade-offs of coordinating open humanities data across multiple platforms through the OpenScore initiative. Using openly licensed digital musical scores as a case study, he argues that Wikidata is most effective as part of a broader ecosystem of complementary infrastructures, highlighting the importance of platform choice, community coordination, and sustainable data stewardship. Trognitz et al. (2026) demonstrate how Wikidata can serve as a sustainable reconciliation anchor for metadata curation and long-term preservation. Using examples from the ARCHE repository and its aggregators, they illustrate the mutual benefits of integrating Wikidata into repository workflows, improving metadata quality while strengthening the shared LOD ecosystem. Centelles Velilla and

Ferran-Ferrer (2025) present an ontology-first, community-driven workflow that integrates humanities datasets into Wikidata using reusable modelling patterns, knowledge graphs, and neuro-symbolic AI. Their approach improves data quality, interoperability, and explainability while enabling SPARQL-based analysis and retrieval-augmented generation (RAG). It also demonstrates how community-governed modelling can promote more inclusive knowledge graphs through gender-aware data curation. Loll (2026) presents a harmonised open dataset of Germany's public vocational schools, systematically modelled as Wikidata items and linked to occupational classifications, geographic information, and external identifiers. The dataset demonstrates how Wikidata can facilitate the integration, harmonisation, and reuse of public educational data, enabling spatial analyses, studies of vocational education provision, and research on school naming.

(3) DISCUSSION

Having outlined the collection's scope and contents, we now turn to what this scope suggests about Wikidata's place in humanities research. The discussion that follows is fourfold: first, an assessment of the collection's community engagement and thematic scope, then a closer examination of whose institutions and geographies shaped that scope, and finally a look at the solutions currently being used to address Wikidata's inclusivity issues.

(3.1) COMMUNITY ENGAGEMENT

With 34 discussion papers and 11 data papers, *Wikidata Across the Humanities: Datasets, Methodologies, Reuse* is the largest special collection of the journal at the time of writing this article. The scale of the collection is significant, but its value is better understood through the breadth and diversity of the contributions, which suggest that Wikidata is becoming a cornerstone across multiple disciplines of the humanities: no longer a general-purpose reference tool on the margins of scholarship, but something researchers actively build on — for domain-specific data modelling, scholarly collaboration, or Citizen Science initiatives.

The Wikidata community itself offers a striking example of this self-reflexive use: Wikidata editors created a Wikidata item for the collection (Q136325279).¹ Subsequently, the community created items for every article published as part of the collection and linked each one to the collection item. As a result, it systematically represented this extensive body of scholarly work on Wikidata within Wikidata itself, making the collection more visible, more discoverable, and more readily integrated into the broader scholarly knowledge graph. This shows that, for the Wikidata community, curating a collection like this one is not just about publishing articles — it is a form of scholarly responsibility the community takes seriously in its own right.²

(3.2) THEMATIC VARIETY

The composition of the collection itself points in the same direction. The majority of contributions do not simply publish a dataset — they reflect on methods, workflows, and the epistemic consequences of using Wikidata as a research infrastructure. This roughly three-to-one ratio of discussion to data papers suggests that Wikidata has reached a point of critical mass in the humanities where the community is no longer only building with it, but thinking through it: interrogating what it means to model historical knowledge as structured data, what is gained and lost in the translation, and what responsibilities come with contributing to a platform that operates at the scale described above.

Taken together, this pattern confirms what Zhao (2023) had anticipated: Wikidata's greatest potential lay not just in hosting data, but in serving as a platform and technology stack for publishing LOD and enabling exchange between external resources and Wikidata itself.

(3.3) GEOGRAPHIC DISTRIBUTION OF THE AUTHORS

Yet if datasets are arguments, it matters who is making them and from what perspective. The geographic distribution of authors by their institutional affiliation (Figure 2) clearly shows that

1 <http://www.wikidata.org/entity/Q136325279>.

2 See also DARIAH-EU WG DHwiki (Lindemann & Candela, 2026) for similar interests.

there is a high level of bias toward contributions from the Global North. In total, 107 authors are affiliated with institutions in 20 countries. Europe and North America alone account for 101 of them ($\approx 94\%$), while the remaining 6 ($\approx 6\%$) come from Turkey, Brazil, and China. Germany and the United States by themselves represent 51 authors, appearing in nearly half of the contributions ($\approx 48\%$). This concentration may partly reflect the networks of the collection's guest editors and its international scientific committee, as well as the central role of Wikimedia Deutschland within the Wikidata ecosystem. It also raises the question of whether the call for contributions reached primarily our own network, leaving much of the global Wikidata community unaware of it altogether. Similar patterns of over-representation of the Global North have also been identified in recent community surveys conducted by Wikimedia Deutschland (Dittrich, 2021; WMDE UX team, 2025).

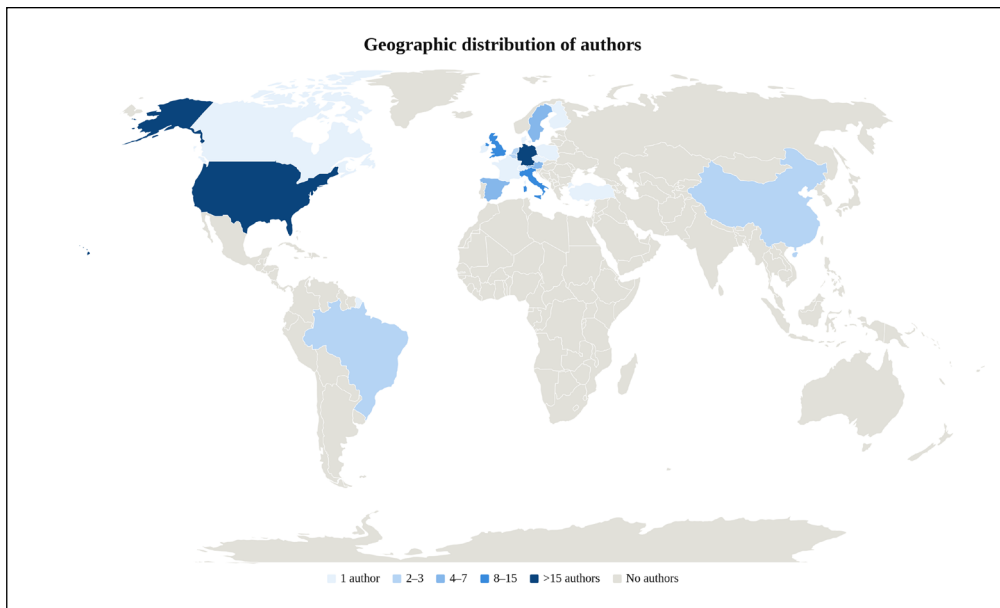


Figure 2 Geographic distribution of authors based on their institutional affiliation.

This figure makes evident how the asymmetries, biases, and inequalities of the web are replicated by Wikidata and its editing community (Das et al., 2025). With large language models and artificial intelligence (AI) now increasingly relying on Wikidata for grounded, machine-readable facts, the geographic bias of its global community is not just about who gets represented — it is about whose knowledge these systems treat as fact.³ It is not a problem exclusive to Wikidata either. Risam (2019) has shown how the rise of digital humanities has been celebrated as an endeavour committed to openness, access, and knowledge democracy — yet one that poses fundamental problems of omission with respect to race, gender, and nation. Digital humanities projects, even when conceived with open mandates, invariably rely on the institutional infrastructure of the Global North, including the universities that sponsor them, the libraries that digitise their collections, and the academics who have the time to carry them out. In fact, making this data available online does not address these disparities but rather institutionalises them within datasets, ontologies, and knowledge graphs that can involuntarily exclude communities from the Global South.

Furthermore, publication practices and access to research funding may also contribute to such geographic asymmetries. Open-access publication costs can constitute a potential barrier for scholars whose institutions or projects do not provide dedicated funding for publication, although the JOHD offers fee waivers for authors without access to such funding. The present data do not allow us to determine to what extent publication costs contributed to the geographic distribution observed here. Rather, the pattern may reflect a combination of factors, including the networks through which the call circulated, existing inequalities in participation in the Wikidata and digital humanities communities, and broader differences in access to research resources and infrastructures.

³ For AI's integration in Wikidata, see for example the Embedding Project (https://www.wikidata.org/wiki/Wikidata:Embedding_Project) and the Model Context Protocol (MCP) (<https://www.wikidata.org/wiki/Wikidata:MCP>).

(3.4) MAKING WIKIDATA MORE INCLUSIVE

To address this form of “digital colonialism”, Wikidata and its community of editors need to develop digital methods that reflect a broader range of communities and knowledge traditions. This also means building multilingual infrastructures that do not take Western conventions of human-computer interaction as their unmarked default. Building global knowledge needs to be done by a global community. Low-barrier platforms like Wikidata can help narrow the gap between the Global North and South — but only if contributors are encouraged to participate more evenly across both regions.

Multiple strategies and initiatives are already in place and growing to make Wikidata a more inclusive place. GLAM institutions in particular are at the forefront of this. In response to calls from communities of origin, GLAM institutions are working to open up their collections, increase transparency, and make information accessible as a first step — with co-curation and (digital) repatriation or restitution following as further stages of this process. Projects such as TheMuseumsLab,⁴ International (Digital) Dura-Europos⁵ and Restitution of Knowledge⁶ illustrate efforts to involve communities of origin directly, through fellowships and knowledge exchange, edit-a-thons or other collaborative formats. These initiatives depend on relationships built on trust, which in turn raises questions of data ethics and data sovereignty, the CARE principles (Carroll et al., 2020), and the treatment of traditional and Indigenous knowledge within Wikidata’s data model. Geographical bias is not the only content gap Wikidata’s editing community must address. Similar content gaps affect marginalised communities, with initiatives such as Art + Feminism edit-a-thons working to address underrepresentation of women and queer identities in Wikidata (Melis et al., 2024).

(4) CONCLUSION

Taken against the objectives set out at the outset, the collection provides several indications of qualitative success beyond its size. First, its disciplinary breadth demonstrates that the relationship between Wikidata and humanities data is not confined to a single research community: the contributions span classics, archaeology, literary studies, linguistics, art history, cultural heritage, and other fields, while addressing datasets ranging from individual texts and inscriptions to large-scale institutional collections. Second, the collection demonstrates a substantial diversity of uses for Wikidata, including authority control, entity reconciliation, data publication, ontology development, knowledge-graph construction, geocoding, and connections with Wikibase and other LOD infrastructures. Third, many contributions move beyond demonstrating technical integration to critically examine questions of data modelling, quality, sustainability, provenance, interoperability, and inclusivity. In this respect, the collection has achieved an important part of its intended purpose: it does not present Wikidata simply as a technology to be adopted, but documents a range of ways in which humanities researchers are incorporating it into their existing data practices while critically assessing its limitations and implications.

The collection also provides evidence of increased dialogue between the Wikidata and humanities data communities. Its contributions bring together researchers working with Wikidata directly and researchers approaching it through established humanities datasets, repositories, ontologies, and institutional infrastructures. The subsequent representation of the collection and its individual articles within Wikidata itself further illustrates this reciprocal relationship. These developments should not, however, be taken as evidence that the collection has already achieved its broader objective of transforming humanities data publishing or making Wikidata more representative. Such impacts require longer-term assessment, including evidence of dataset reuse, new Wikidata contributions, subsequent collaborations, and changes in publishing and data-management practices. The present collection can therefore best be understood as an important step towards these goals rather than as their endpoint.

4 <https://themuseumslab.org/en>.

5 <https://duraeuroposarchive.org/>.

6 https://www.wikidata.org/wiki/Wikidata:WikiProject_TheRoK.

The predominance of discussion papers over data papers in the collection is also significant. Most contributions reflect on datasets, but also on workflows, challenges, modelling choices, and the implications of representing humanities knowledge as structured data. This suggests that Wikidata is now understood both as a research infrastructure and as an object of scholarly inquiry. At the same time, the collection highlights a complementary relationship between Wikidata and peer-reviewed humanities data publishing. While Wikidata provides a global, collaborative, and continuously evolving knowledge graph, data and discussion papers within JOHD provide the scholarly context, documentation, and methodological transparency necessary to understand and reuse individual datasets. High-quality, openly licensed datasets can in turn provide valuable sources and references for Wikidata. Together, these practices contribute to a more robust ecosystem for producing, documenting, connecting, and reusing humanities data.

However, the collection also makes clear that the future development of Wikidata as a humanities infrastructure depends on addressing persistent challenges of representation, participation, and inclusivity. The geographical concentration of contributions, as well as broader content gaps within Wikidata, reflects wider structural inequalities in global knowledge production. Expanding Wikidata's role in the humanities therefore requires not only technical improvements, but also sustained efforts to support participation from diverse communities, incorporate different knowledge traditions, and develop practices attentive to data ethics and sovereignty.

Ultimately, what these contributions show is that Wikidata has become more than a technological tool: it is a social and scholarly infrastructure through which humanities researchers negotiate how knowledge is represented, connected, and shared. Its future potential lies precisely in this dual role, as both a platform enabling interoperable open knowledge and a space where the humanities can critically examine the processes through which knowledge itself is produced.

COMPETING INTERESTS

Andrea Farina is Senior Editor and Senior Social Media Editor for the Journal of Open Humanities Data (JOHD). Alan Ang is Senior Partner Manager at Wikimedia Deutschland, where he works on data partnerships related to Wikidata and Wikibase. Camillo Carlo Pellizzari di San Girolamo is Regional Coordinator for Tuscany at Wikimedia Italia.

AUTHOR CONTRIBUTIONS

Maxime Guénette: Conceptualisation, Writing – Original Draft (Section 2, 3 and 4), Writing – Review & Editing, Visualisation.

Sabine von Mering: Conceptualisation, Writing – Original Draft (Sections 2, 3 and 4), Writing – Review & Editing, Visualisation.

Camillo Carlo Pellizzari di San Girolamo: Conceptualisation, Writing – Original Draft (Sections 2, 3 and 4), Writing – Review & Editing.

Alan Ang: Conceptualisation, Writing – Original Draft (Section 4).

Andrea Farina: Conceptualisation, Data Curation, Supervision, Writing – Original Draft (Sections 1 and 4), Writing – Review & Editing.

AUTHOR AFFILIATIONS

Maxime Guénette  orcid.org/0009-0006-2076-1220

Department of History, Université de Montréal, Montréal, Canada

Sabine von Mering  orcid.org/0000-0003-2982-7792

Collection Future, Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany

Camillo Carlo Pellizzari di San Girolamo  orcid.org/0000-0003-2699-1693

Classe di Lettere e Filosofia, Scuola Normale Superiore, Pisa, Italy

Alan Ang  orcid.org/0009-0003-7195-3124

Wikimedia Deutschland, Berlin, Germany

Andrea Farina  orcid.org/0000-0002-1948-9008

Department of Digital Humanities, King's College London, London, United Kingdom

- Aangenendt, G., Harris, A., Wu, D., Maen, A., & Oberst, U. (2025). Wikibase Solutions for African literary metadata. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.433>
- Akhlaq, S., Bates, E. I., & Depraetere, M. (2026). Making the labour visible behind the Tendaguru dinosaur collection through Wikidata. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.476>
- Baránek, D. (2026). Machine learning assisted conversion of biographical records into Wikidata triples. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.466>
- Bemme, J., & Munke, M. (2026). Call for edits nearby: Open archives metadata from Saxony. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.430>
- Berti, M. (2025). Linked Ancient Greek and Latin (LAGL) and Wikidata: Structuring and reusing data of classical literature. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.423>
- Besnier, C., & Mattingly, W. (2021). Named-entity dataset for Medieval Latin, Middle High German and Old Norse. *Journal of Open Humanities Data*, 7, 23. <https://doi.org/10.5334/johd.36>
- Blair, S. (2026). Linked data, fragmented knowledge: Towards a digital Quellenkritik of Verrius Flaccus' lexicon *De uerborum significatu*. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.455>
- Campos Castello Branco, R., & von Mering, S. (2025). Behind the picture: German filmmakers as agents of natural history collections. The case study of Adolf Freiherr von Dungen in the Amazon. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.443>
- Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE Principles for Indigenous data governance. *Data Science Journal*, 19(1), 43. <https://doi.org/10.5334/dsj-2020-043>
- Centelles Velilla, M., & Ferran-Ferrer, N. (2025). Using Wikidata's ontology in practice: A neuro-symbolic, community-centred workflow for integrating and reusing humanities datasets. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.439>
- Chen, A. H., Guénette, M., Thornton, K., Seals-Nutt, K., & Martin, E. (2026). Interoperability as equity: collaborative cultural heritage knowledge graphs as a tool to shape inclusive ontologies. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.463>
- Das, P., Karnam, S. K., Soni, A. B., & Mukherjee, A. (2025). Social biases in knowledge representations of Wikidata separates Global North from Global South. *Proceedings of the 17th ACM Web Science Conference 2025*, 12–21. <https://doi.org/10.1145/3717867.3717882>
- de Beyssat, C. D. (2025). Victims of posterity. Identifying gaps on 19th-Century french art history with Wikidata. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.399>
- De Monaco, S. (2025). The “Dizionario degli scrittori Italiani contemporanei pseudonimi” in Wikidata: Pseudonymous authors in Wikidata: Dataset and queries. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.441>
- Dilley, P. C. (2026). An assessment of Wikidata on premodern writing systems. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.482>
- Dişli, M., Candela, G., Gutiérrez, S., & Fontenelle, G. (2025). Open data practices of art museums in Wikidata: A compliance assessment. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.438>
- Dittrich, J. (2021). *Wikidata community survey 2021*. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Wikidata_Community_Survey_2021.pdf
- Fantoli, M., Boano, V. I., de Graaf, E., & Pellizzari di San Girolamo, C. C. (2026). Wikidata as a knowledge base for people of the Greco-Roman World. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.457>
- Farda-Sarbas, M., & Müller-Birn, C. (2019). Wikidata from a research perspective – A systematic mapping study of Wikidata. *arXiv*. <https://doi.org/10.48550/arXiv.1908.11153>
- Farina, A. (2023). Lost at sea: A dataset of 25+ SEA words morpho-semantically annotated in Ancient Greek and Latin. *Journal of Open Humanities Data*, 9(24), 1–7. <https://doi.org/10.5334/johd.139>
- Farina, A., & McGillivray, B. (2026). The Journal of Open Humanities Data: Bridging open data and Wikidata for the humanities. In E. Marangoni & C. C. Pellizzari di San Girolamo (Eds), *Wikidata e la ricerca. Condividere esperienze. Atti del Convegno, Firenze, 5–6 giugno 2025* (pp. 337–341). Firenze University Press. <https://doi.org/10.36253/979-12-215-1010-2.41>
- Foster, A., Norling, S., & Westerby, M. J. (2026). Towards a linked open index: reconciling museum records to Wikidata for Index of American Design constituents. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.426>
- Gautschy, R. (2025). Digital LIMC – A knowledge graph for classical mythology. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.403>
- Giovannini, L., & Gallucci, G. (2024). Allacci Digitale: A historical dataset for early modern Italian drama. *Journal of Open Humanities Data*, 10, 55. <https://doi.org/10.5334/johd.250>

- Gotham, M. (2026). Wherever I lay my hat is home?: A complex case study of crowd-sourcing, coordination, and cross-platform integration for hosting open humanities data. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.456>
- 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(1). <https://doi.org/10.5334/johd.428>
- Jones, M. (2026). Visualizing the linked manuscript archive: A linked open data approach to analyzing special collections data. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.444>
- Kim-Baumann, N. J. (2026). Modelling temporal public transport systems in Wikidata: A workflow for integrating historical infrastructure data. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.458>
- Kubiak-Schneider, A. (2025). Epigraphical datasets of the Temple of Atargatis from Dura-Europos for the WikiProject International (Digital) Dura-Europos Archive. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.445>
- La Veglia, A. (2026). The digital scholarly edition of Angelo Poliziano's marginal notes to the Epistle of Sappho to Phaon. Encoding of multiple edition levels and implementation of Wikidata Items and CTS URNs. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.493>
- Lamar, A. K. (2025). An HPC-ready, Wikidata-based workflow for exploratory geocoding of unstructured textual corpora. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.401>
- Lindemann, D., & Candela, G. (2026, May 27). DARIAH-EU WG DHwiki—A first year of activity. DARIAH Annual Event 2026, Rome, Italy. Zenodo. <https://doi.org/10.5281/zenodo.20409899>
- Lindemann, D., Pellegrini, M., Mambrini, F., & Passarotti, M. (2025). Wikidata and LiLa for Latin: Enabling interoperability and access to inflected forms and corpus attestations. *Journal of Open Humanities Data*, 11(83). <https://doi.org/10.5334/johd.464>
- Loll, I. M. (2026). Open public vocational school data for Germany. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.422>
- Mambrini, F., & Passarotti, M. C. (2023). The LiLa Lemma Bank: A knowledge base of Latin canonical forms. *Journal of Open Humanities Data*, 9, 28. <https://doi.org/10.5334/johd.145>
- Mathios, K., Doğan, E. A., & Nelson, J. (2026). Refining Wikidata's bibliographic domain to increase reuse in GLAMs. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.454>
- McCandless, R. A., & Coladangelo, L. P. (2025). Integrating premodern manuscript metadata into Wikidata: A case study in ontology design and linked data reuse. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.431>
- Mees, A. W., & Thiery, F. (2026). Bridging archaeological Samian Ware data and the knowledge graph: Potentials and challenges of using Wikidata as a linking hub. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.427>
- Meiners, H.-L., & Bulle, K. (2025). Disrupting the dust: Identifier properties and the future of cultural heritage metadata in Wikidata. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.412>
- Melis, B., Paolini, C., Fioravanti, M., & Metilli, D. (2024). What does it mean to be queer in Wikidata? Practices of gender representation within a transnational online community. *Communication, Culture and Critique*, 17(3), 200–207. <https://doi.org/10.1093/ccc/tcae029>
- Munke, M. (2026). Modelling historical travelling in Wikidata: The database “Digital Travelogues. Saxony in Travelogues of the Late Middle Ages and Early Modern Times”. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.437>
- Neubert, J. (2017). Wikidata as a linking hub for knowledge organization systems? Integrating an authority mapping into Wikidata and learning lessons for KOS mappings. *NKOS@TPDL*, 1–12.
- Pensa, I., Benvenuti, F., & Bertacchini, E. (2026). Empowering Italian GLAMs: A workflow for building a comprehensive open dataset of museums on Wikidata. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.477>
- Ripoll-Alberola, L., Le Bris, M.-M., & Fischer, J. P. (2026). Canons across time: compiling lists of ancient authors with Wikidata. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.442>
- Risam, R. (2019). *New digital worlds: Postcolonial digital humanities in theory, praxis, and pedagogy*. Northwestern University Press. <https://doi.org/10.2307/j.ctv7tq4hg>
- Rohrbacher, K., & Schrittmesser, D. (2026). Bridging the gaps: integrating bibliographic metadata into Wikidata for literary corpora. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.483>
- Rolleri, D., Marchetti, A., & di San Girolamo, C. C. P. (2026). Documenting Italian libraries on Wikidata: From local projects to a multilayered national knowledge graph. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.478>
- Röttgermann, J. (2024). The collection of eighteenth-century French novels 1751–1800. *Journal of Open Humanities Data*, 10, 31, 1–5. <https://doi.org/10.5334/johd.201>

- Samo, G., & Ursini, F.-A. (2025). Generic geonyms: Exploring Wikidata for crosslinguistic prototypical semantics. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.432>
- Schneider, S., Springstein, M., Stalter, J., Ritter, D., Ewerth, R., & Müller-Budack, E. (2026). Uncovering biases through faceted image retrieval: Wikidata as a data provider for art-historical research. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.436>
- Shaw, J., Young, G., Cranston, M., Thomas, S., & Ross, K. S. (2026). From websites to Wikidata: Digitising Scotland's stories. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.471>
- Sichani, A.-M., Kono, K., & Winters, J. (2026). Building cultural heritage data infrastructures with Wikidata: The case of the Congruence Engine Data Register. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.448>
- Sohmen, L., Rossenova, L., & Blümel, I. (2026). Wikidata 4 Open Culture: Lessons learned from hands-on work with cultural heritage data in the expanded Wikibase ecosystem. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.440>
- Sorensen, A. H., Menyuk, R., Shapiro, D., & Naples, R. (2026). Piloting Wikidata as an authority identifier: The (In)visible Women Project at the Smithsonian Institution. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.467>
- Thiery, F., Schenk, F., Stefan, D., Distel, A.-K., & Thiery, P. A. (2026). From stones and wells to tephra: Leveraging Wikidata and Wikibase for interdisciplinary humanities linked data. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.425>
- Trognitz, M., Mandell, R. A., Štuhec, S., & Palacz, J. (2026). Wikidata as a reconciliation anchor: Curating data for long-term preservation. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.450>
- Untner, L. (2025). From Wikidata to CIDOC CRM: A use case scenario for digital comparative literary studies. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.421>
- Westerby, M. J., Ferrara, L., & Um, N. (2025). A practical join: Using Wikidata to aggregate awards and affiliations data from residential fellowship programs in art history. *Journal of Open Humanities Data*, 11(1). <https://doi.org/10.5334/johd.429>
- Wigdorowitz, M., Ribary, M., Farina, A., Lima, E., Borkowski, D., Marongiu, P., Sorensen, A. H., Timis, C., & McGillivray, B. (2024). It takes a village! Editorship, advocacy, and research in running an open access data journal. *Publications*, 12(3), 24. <https://doi.org/10.3390/publications12030024>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. 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>
- WMDE UX Team. (2025). *Wikidata community survey 2024 report*. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Wikidata_Community_Survey_2024_Report.pdf
- Wolter, V. (2026). Developing a data model for theatre productions in a Wikibase instance: A case study approach. *Journal of Open Humanities Data*, 12(1). <https://doi.org/10.5334/johd.434>
- Zhao, F. (2023). A systematic review of Wikidata in Digital Humanities projects. *Digital Scholarship in the Humanities*, 38(2), 852–874. <https://doi.org/10.1093/lc/fqac083>

TO CITE THIS ARTICLE:

Guénette, M., von Mering, S., Pellizzari di San Girolamo, C. C., Ang, A., & Farina, A. (2026). Editorial. Wikidata Across the Humanities: Datasets, Methodologies, Reuse. *Journal of Open Humanities Data*, 12: 136, pp. 1–13. DOI: <https://doi.org/10.5334/johd.624>

Submitted: 24 July 2026

Accepted: 10 September 2026

Published: 15 October 2026

COPYRIGHT:

© 2026 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 <https://creativecommons.org/licenses/by/4.0/>.

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