

The Journal of Open Humanities Data. Bridging open data and Wikidata for the Humanities

Andrea Farina and Barbara McGillivray (King's College London)

Wikidata serves as a critical tool for enriching and interconnecting datasets, enabling researchers to explore relationships across diverse domains (Farda-Sarbas and Müller-Birn 2019; Neubert 2017). It offers a centralized platform for integrating identifiers, metadata, and semantic links, and allows for the creation of interoperable and reusable datasets supporting advanced analysis and interdisciplinary research. At the Journal of Open Humanities Data (JOHD), we embrace the potential of Wikidata to amplify the impact of open data for research in the humanities. Through the publication of data papers (peer-reviewed articles describing datasets, their methodologies, and their reuse potential), JOHD ensures that humanities datasets are accessible and reusable. Our mission aligns with the principles of platforms like Wikidata, emphasizing transparency, accessibility, and collaboration to elevate the role of data in advancing scholarly work and public engagement (Wigdorowitz et al. 2024).

This poster highlights the synergies between JOHD and Wikidata, focusing on how the journal's principles of open access, reusability, and reproducibility complement Wikidata's capabilities as a linked open data hub. This collaboration can enhance the value and impact of humanities research in the digital age with JOHD acting as a bridge to encourage humanities scholars to engage with Wikidata by providing guidance on integrating datasets into Wikidata. We present case studies of data papers published in JOHD, showing how they have used Wikidata for dataset creation. For instance, linking place names in historical newspapers to Wikidata (Coll Ardanuy et al. 2022) enhances cultural heritage accessibility. Multilingual cultural heritage information, such as historical Chinese kung fu masters, can be integrated with Wikidata into reusable and human-centered knowledge graphs (Hou and Yuan 2023). Further, Wikidata ensures the reusability and transparency of bibliographical data, supporting JOHD's emphasis on reproducible research (Malínek et al. 2024). We also comment on published datasets that do not mention Wikidata but could potentially benefit from its integration to enhance interdisciplinarity (e.g., Farina 2023).

Finally, we explore how datasets published in JOHD and integrated with Wikidata can enhance the visibility and discoverability of research (cf. McGillivray et al. 2022) by tracking dataset reuse and citation within the Wikidata ecosystem.

References

- Coll Ardanuy, M., Beavan, D., Beelen, K., Hosseini, K., Lawrence, J., McDonough, K., Nanni, F., van Strien, D., & Wilson, D. C. S. (2022). A Dataset for Toponym Resolution in Nineteenth-Century English Newspapers. *Journal of Open Humanities Data*, 8(1), 3, pp. 1–7. DOI: <https://doi.org/10.5334/johd.56>
- Farda-Sarbas, M., & Müller-Birn, C. (2019). Wikidata from a Research Perspective - A Systematic Mapping Study of Wikidata. *ArXiv*, *abs/1908.11153*. <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, pp. 1–7. DOI: <https://doi.org/10.5334/johd.139>
- Hou, Y., & Yuan, L. (2023). Building a Knowledge Graph of Chinese Kung Fu Masters From Heterogeneous Bilingual Data. *Journal of Open Humanities Data*, 9: 27, pp. 1–12. DOI: <https://doi.org/10.5334/johd.136>
- Malínek, V., Umerle, T., Gray, E., Heibi, I., Király, P., Klaes, C., Korytkowski, P., Lindemann, D., Moretti, A., Panušková, Ch., Péter, R., Tolonen, M., Tomczyńska, A., & Vimr, O. (2024). Open Bibliographical Data Workflows and the Multilinguality Challenge. *Journal of Open Humanities Data*, 10: 27, pp. 1–14. DOI: <https://doi.org/10.5334/johd.190>
- McGillivray, B., Marongiu, P., Pedrazzini, N., Ribary, M., Wigdorowitz, M., & Zordan, E. (2022). Deep Impact: A Study on the Impact of Data Papers and Datasets in the Humanities and Social Sciences. *Publications*, 10(4), 39. <https://doi.org/10.3390/publications10040039>
- 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.
- 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>