

## FAIRness of research data in humanities

Ljiljana Poljak Bilić<sup>1</sup>, Kristina Posavec<sup>2</sup>

<sup>1</sup>University of Split Library, Croatia

<sup>2</sup>University of Zagreb University computing centre SRCE, Croatia

### ABSTRACT

Research data are considered the primary result and output of scientific research, and sharing and reusing data are key aspects of the transition to open science on a European level. However, there are many unanswered questions regarding research data management, such as understanding which data might be shared and under what conditions, who might share them, or even what research data is (Borgman, 2012). In the humanities, research data are among the most diverse of all scientific disciplines because nearly any data on human activity can be considered research data, such as newspapers or photographs. Thus, the boundary between data and publication in this context is often quite vague (Borgman, 2008; Thoegersen, 2018).

This paper aims to examine and answer several research questions related to datasets in the field of humanities. The first research question will analyse the types of research data represented in humanities (1). In literature, a consensus has yet to be reached on the definition, and often, research data in the humanities are used as an umbrella term that includes different types of sources for research. DARIAH-DE defines research data in humanities as all sources/materials and results collected, written, described, and/or evaluated in the context of a research and research questions in the field of human and cultural sciences, presented in machine-readable form for archiving, citation and further use. This definition aims to account for the particular characteristics of human-scientific research and the resulting heterogeneity of the underlying data (DARIAH- DE, 2021). For the purpose of this research, the authors have chosen to adopt the DARIAH- DE definition of research data which will be the base for determining types of data represented in humanities. The research was conducted for humanities fields, including philosophy, theology, philology, history, art history, art science, archaeology, ethnology and anthropology, religious studies, and interdisciplinary humanities science. An analysis of datasets published in institutional and thematic repositories in the field of humanities in Croatia and Europe was conducted using platforms such as Zenodo, Digital Academic Archives and Repositories (DABAR), CLARIN, CROSSDA, and DARIAH. The research aimed to identify the most prevalent data types in the humanities found in repositories.

In a study on the storage of research data from the humanities in repositories, Buddenbohm et al. (2016) note that a culture of sharing and reusing research data has not yet been established; although research data are, to some extent, stored in repositories, they are difficult to find. In light of this, the second research question will examine the extent to which datasets in the field of humanities are represented in repositories in open access and under which licences (2). The third research question will examine to what extent research data align with FAIR principles (3). The evaluation of dataset FAIRness will

be conducted using the principles outlined in the Wilkinson et al. (2016) FAIR guidelines and Routledge Open Research data guidelines, as mentioned in Grant (2022) for FAIR humanities data. In addition, based on FAIR guidelines, authors will provide a practical framework for managing research data in humanities, which can be used to assess the quality of research datasets. The paper will present the types of research data in humanities that are represented in repositories, assess their level of openness, licensing, and alignment with FAIR data principles.

## KEYWORDS

data management; FAIR data; humanities; open science; research data

## REFERENCES

1. Borgman, C. L. (2008). Data, disciplines, and scholarly publishing. *Learned Publishing*, 21(1), 29–38. <https://doi.org/10.1087/095315108X254476>
2. Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059–1078. <https://doi.org/10.1002/asi.22634>
3. Buddenbohm, S. et al. (2016). State of the art report on open access publishing of research data in the humanities. DARIAH. <https://halshs.archives-ouvertes.fr/halshs-01357208v2>
4. DARIAH-DE. n.d. Accessed 21 April 2023. <https://de.dariah.eu/en/weiterfuehrende-informationen>
5. Harrower, N. M., et al. (2020) Sustainable and FAIR Data Sharing in the Humanities: Recommendations of the ALLEA Working Group. E-Humanities Digital Repository of Ireland [Distributor], Digital Repository of Ireland. <https://doi.org/10.7486/DRI.tq582c863>
6. Pawlicka, U. (2017). Data, Collaboration, Laboratory: Bringing Concepts from Science into Humanities Practice. *English Studies*, 98(5), 526–541. <https://doi.org/10.1080/0013838x.2017.1332022>
7. Grant, R. (2022). Reusable, FAIR Humanities data. *International Journal of Digital Curation*, 17(1), 15. <https://doi.org/10.2218/ijdc.v17i1.820>
8. Thoegersen, J. L. (2018). "Yeah, I Guess That's Data": Data Practices and Conceptions among Humanities Faculty. *Portal - Libraries and the Academy*, 18(3), 491–504. <https://doi.org/10.1353/pla.2018.0030>
9. Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>