

Open Research Data: Drawbacks to think about (Editorial)

José L. Arquero (Editor)

<https://orcid.org/0000-0002-7086-8812>

Universidad de Sevilla (Spain)

<https://dx.doi.org/10.12795/EDUCADE.2025.i16.01>

The relevance of Open research data (ORD) policies

Open research data (ORD) is recognized as a fundamental pillar of open science, rooted in the core principle of increasing accessibility, usability, and reusability of scientific findings (Hahnel et al., 2024). These policies are crucial for enhancing transparency, integrity, and reproducibility across diverse research fields (Stieglitz et al., 2020). ORD is viewed as essential for tackling massive, interdisciplinary global challenges (Borycz et al., 2023). Making data openly available facilitates reuse, speeds up innovation, and promotes interdisciplinary cooperation (Borgerud & Borglund, 2020). Beyond societal benefits, data sharing could result in personal advantages for researchers, including increased citation rates, improved reputation, and career benefits (Borgerud & Borglund, 2020; Hahnel et al., 2023; Stieglitz et al., 2020).

The political support for ORD is substantial, driven primarily by public institutions such as the European Commission, which has promoted open science (González-Teruel et al., 2022). International research funding institutions—e.g. the European Union, the National Institutes of Health (NIH), the National Science Foundation, and the German Research Foundation—increasingly make open access to research data a top priority and a prerequisite for funding (Borycz et al., 2023; Stieglitz et al., 2020). Governments, such as the Swedish government, have also mandated that publicly funded research results and data must be openly accessible as far as possible (Borgerud & Borglund, 2020), and recently in Spain the evaluation criteria used by the ANECA to assess research also value and promote open science for both results and data. Furthermore, major academic publishers, including PLOS and Springer Nature, enforce mandates requiring authors to provide access to their underlying data, often through Data Availability Statements (DAS) (Hahnel et al., 2024). Compliance is guided by the **FAIR** principles (Findable, Accessible, Interoperable, and Reusable), which are central to ensuring data quality and usability (Borycz et al., 2023; Hahnel et al., 2024; Stieglitz et al., 2020). These concerted efforts aim to create a lasting cultural shift in scientific practices, maximizing the return on public investment and accelerating global knowledge advancement (Hahnel et al., 2024).

ORD problems and barriers

However, Open research data (ORD) policies encounter numerous drawbacks, problems, and limitations. These constraints stem from individual researcher factors, legal

and ethical hurdles, deficiencies in infrastructure and resources, and inadequate recognition within the scientific evaluation system (González-Teruel et al., 2022; Stieglitz et al., 2020; Hahnel et al., 2023).

Examining a set of relevant publications (Borgerud & Borglund, 2020; Borycz et al., 2023; González-Teruel et al., 2022; Hahnel et al., 2023; Hahnel et al., 2024; Stieglitz et al., 2020) on the topic: the main barriers identified could be classified as follows:

1. Challenges Related to Researchers and Research Culture

Researchers' reluctance and attitudes constitute major barriers to ORD. Many lack knowledge and experience in data management—nearly three-quarters report never having received support for planning or sharing data (Hahnel et al., 2023). Research materials are often perceived as private property, and confusion persists over what constitutes "data," particularly in disciplines where outputs are qualitative, artistic, or theoretical.

A significant concern is fear of competition and loss of value. The academic reward system prioritises novel data production, making researchers hesitant to share datasets that could give competitors an advantage (Fecher et al., 2015). Furthermore, creating open data is time-consuming: formatting, curating, and documenting data require substantial time and resources, often without institutional compensation.

2. Legal, Ethical, and Confidentiality Barriers

Research involving human subjects poses inherent conflicts with openness. Legal restrictions concerning transparency, personal data protection, and national security often prevent full publication of datasets. The protection of personal integrity remains a key concern, as ensuring anonymity is complex, particularly in qualitative studies where participants or locations may be identifiable (González-Teruel et al., 2022).

Anonymisation difficulties frequently discourage sharing. Qualitative datasets may include sensitive content that is hard to remove without distorting meaning, and misinterpretation risks persist even after anonymisation. The absence of clear legal and ethical guidance further complicates compliance, while economic interests and confidentiality agreements, especially in industry-funded projects, limit openness.

3. Infrastructure and Resource Limitations

ORD implementation is hindered by insufficient coordination, funding, and infrastructure. Data preservation and accessibility are often underfunded, as data management is viewed as an indirect rather than core research cost (Borgerud & Borglund, 2020). Many researchers report unsatisfactory or nonexistent data infrastructures and a lack of common metadata standards—most repositories rely solely on the Dublin Core schema (Hahnel et al., 2023).

Technical and preservation challenges include ensuring long-term file compatibility and the persistence of storage systems. Researchers frequently lack practical support, such as training, user-friendly repositories, and clear guidelines, while bureaucratic complexity and poor coordination among institutions further impede data sharing.

4. Disciplinary and Data-Type Limitations

The heterogeneity of research disciplines makes universal ORD policies difficult to implement. Attitudes toward data sharing vary significantly: social scientists, for instance, are generally less willing to share data than researchers in the natural or information sciences (Borycz et al., 2023). Qualitative data present unique difficulties, such as anonymisation challenges and heavy contextual dependence, which make external reuse problematic. Consequently, no single solution fits all disciplines or data types.

5. Recognition and Evaluation System Problems

The current academic system often conflicts with ORD principles. *Lack of credit and recognition* remains a primary barrier: researchers note that data citations rarely contribute to career advancement or funding success (Hahnel et al., 2023; Hahnel et al.,

2024). The persistence of traditional evaluation metrics—citations, journal impact, and individual reputation—discourages open data practices (González-Teruel et al., 2022). Competition for resources intensifies this reluctance. As funding and career advancement depend on producing new data, researchers may perceive sharing as detrimental to their competitive standing (Borgerud & Borglund, 2020). Moreover, defining when a research project is “complete” and ready for archiving remains challenging, especially in large or longitudinal studies.

Special reference to research in education

In my opinion, the more constraining factors in the case of educational research are those related to *Recognition and Evaluation System*

Data collection in social sciences often involves substantial intellectual, temporal, and financial investment. Designing valid instruments, obtaining ethical approvals, recruiting participants, administering surveys, and transcribing responses represent a labour-intensive process that can span months or even years (Corti et al., 2019). Consequently, immediate open dissemination of such data may allow other researchers to reuse the dataset without investing comparable resources, potentially undermining the original team's ability to capitalise on their effort through publications or further analyses (Fecher et al., 2015).

Moreover, the value of primary data in the social sciences is closely tied to its novelty and contextual richness. Premature data sharing can erode the competitive edge of researchers, particularly in environments where academic success is strongly publication-driven (Tenopir et al., 2015). Additionally, the lack of robust mechanisms for crediting data producers exacerbates the problem, as citations to datasets do not yet carry the same academic weight as journal articles (Borgman, 2012).

Finally, in smaller or specialised fields, openly shared datasets may enable competitors to anticipate or replicate forthcoming findings, potentially pre-empting original contributions. Therefore, while open research data policies promote transparency and reproducibility, they must be balanced with recognition and protection of researchers' intellectual labour.

In conclusion, although the willingness to share data has increased over time—particularly within governmental and academic sectors influenced by policy mandates—levels of satisfaction with available resources remain low, even in regions characterised by high collaboration and data-sharing practices (Borycz et al., 2023). The persistent lack of alignment between scientific practices and institutional policies creates uncertainty that can only be mitigated through appropriate regulatory, evaluative, and funding frameworks (González-Teruel et al., 2022). In particular, misaligned reward systems may significantly limit the adoption of open data practices. Addressing these challenges requires sustained institutional support, harmonised standards, and academic incentives that duly recognise the effort invested in data sharing. It is worth noting that a typical research paper may include more than fifty references, yet often relies on a single dataset.

REFERENCES

- Borgerud, C., & Borglund, E. (2020). Open research data, an archival challenge? *Archival Science*, 20, 279–302. <https://doi.org/10.1007/s10502-020-09330-3>
- 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>

- Borycz, J., Olendorf, R., Specht, A., Grant, B., Crowston, K., Tenopir, C., Allard, S., Rice, N. M., Hu, R., & Sandusky, R. J. (2023). Perceived benefits of open data are improving but scientists still lack resources, skills, and rewards. *Humanities and Social Sciences Communications*. <https://doi.org/10.1057/s41599-023-01831-7>
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2019). *Managing and sharing research data: A guide to good practice* (2nd ed.). SAGE.
- Fecher, B., & Friesike, S. (2014). Open Science: One term, five schools of thought. In S. Bartling & S. Friesike (Eds.), *Opening Science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing* (pp. 17–47). Springer. https://doi.org/10.1007/978-3-319-00026-8_2
- González-Teruel, A., López-Borrull, A., Santos-Hermosa, G., Abad-García, F., Ollé, C., & Serrano-Vicente, R. (2022). Drivers and barriers in the transition to open science: The perspective of stakeholders in the Spanish scientific community. *Profesional de la información*, 31(3), e310305. <https://doi.org/10.3145/epi.2022.may.05>
- Hahnel, M., Smith, G., & Campbell, A. (2024). *Bridging policy and practice in data sharing* [The State of Open Data Report - Special Edition]. Figshare. <https://doi.org/10.6084/m9.figshare.27337476>
- Hahnel, M., Smith, G., Scaplehorn, N., Schoenenberger, H., & Day, L. (2023). *The State of Open Data 2023*. Figshare. <https://doi.org/10.6084/m9.figshare.24428194>
- Monteagudo-Haro, P., & Prieto-Gutierrez, J. J. (2024). Datos abiertos de investigación en repositorios universitarios españoles. *Revista Española de Documentación Científica*, 47(3), e397. <https://doi.org/10.3989/redc.2024.3.1581>
- Stieglitz, S., Wilms, K., Mirbabaie, M., Hofeditz, L., Brenger, B., López, A., & Rehwald, S. (2020). When are researchers willing to share their data? Impacts of values and uncertainty on open data in academia. *PLoS one*, 15(7), e0234172. <https://doi.org/10.1371/journal.pone.0234172>
- Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D., & Dorsett, K. (2015). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLOS ONE*, 15(8), e0237468. <https://doi.org/10.1371/journal.pone.0134826>