https://doi.org/10.51408/issi2025_203

Analysis of compliance with the FAIR principles in Education Science

Andrea Sixto-Costoya[1M], Adolfo Alonso-Arroyo[2], Luiza Petrosyan[3], Rafael Aleixandre-Benavent[4], Rut Lucas-Domínguez[5]

[1]*andrea.sixto@uv.es*
UISYS Research Group, Unit of Information and Social and Health Research, University of Valencia
Associated Unit to INAECU. Interuniversity Institute for Advanced Research on Science and University Evaluation. UC3M-UAM (Spain)
Department of Social Work and Social Services, University of Valencia (Spain)

[2]*adolfo.alonso@uv.es*
UISYS Research Group, Unit of Information and Social and Health Research, University of Valencia
Associated Unit to INAECU. Interuniversity Institute for Advanced Research on Science and University Evaluation. UC3M-UAM (Spain)
Department of History of Science and Documentation, University of Valencia (Spain)

[3]*luipet@inf.upv.es*
Polytechnic University of Valencia. University Institute of Pure and Applied Mathematics (IUMPA) (Spain)

[4]*rafael.aleixandre@uv.es*
UISYS Research Group, Unit of Information and Social and Health Research, University of Valencia
Associated Unit to INAECU. Interuniversity Institute for Advanced Research on Science and University Evaluation. UC3M-UAM (Spain)
Ingenio (CSIC-Polytechnic University of Valencia) (Spain)

[5]*rut.lucas@uv.es*
UISYS Research Group, Unit of Information and Social and Health Research, University of Valencia
Associated Unit to INAECU. Interuniversity Institute for Advanced Research on Science and University Evaluation. UC3M-UAM (Spain)
Department of History of Science and Documentation, University of Valencia (Spain)
CIBERONC (Spain)

Introduction

The roles of accountability in sharing Research data and the ability to reproduce experiments have already been widely pointed out. To operationalize the practice of data sharing, the so-called FAIR principles, which stand for "findable," "accessible," "interoperable" and "reusable," were published in 2016 (Wilkinson et al., 2019). Without compliance with these principles, data quality can be so low that it becomes useless due to the difficulty of understanding it. As was pointed out, data quality is commonly conceived as a construct that is defined by the extent of its usefulness (Brennan, 2017).

In the field of education, raw research data are crucial because they allow for a better understanding of research on educational interventions and learning, which is considered one of the fundamental pillars of human, social and economic development. Their quality must be guaranteed to avoid the risk of misinterpretation or bias. Therefore,

our objective is to assess the quality of a set of educational data sets.

Methods

The methodology used in this study consists of three stages:

1. Capturing datasets on Education Sciences. A search equation was designed to retrieve datasets related to Education. The search was conducted in OpenAlex database where the term "Education" appeared in the Subfield OR Keywords fields.
2. Downloading the records. The total recovered records (datasets) were N=65,199. Looking at the repositories in which they are included, there are 223 different repository variables. For this study, the generalists Zenodo and Figshare repositories were selected. These records were downloaded in.txt format and processed with our *Bibliometricos* software. Once the information was parsed and organized, it was necessary to know the unique identifier (DOI) of the total records. This identifier is needed in the next step of the methodology.
3. FAIR evaluation of datasets with the F-UJI tool. The FAIR assessment for these datasets was performed with the F-UJI tool (https://www.f-uji.net/), that evaluates research data objects, which is a REST API using OpenAPI Specification from a remote server, published under an open-source MIT licence. It is based on aggregated metadata, including metadata embedded in the landing page and metadata retrieved from a DOI. The outcomes of such evaluations yield diverse scores pertaining to the metadata of data and datasets, with 16 metrics distributed across four principles: findability (5 metrics), accessibility (3 metrics), interoperability (3 metrics), and reusability (5 metrics) (Devaraju et al., 2022). This methodology was used and validated previously in Petrosyan et al. (2023) and Sixto-Costoya et al. (2025).

Results

After the analysis through the F-UJI tool, we obtained information about the level of FAIRification of the 4,642 DOI belonged to datasets in the Education Sciences area. Of them, 1,772 belonged to Zenodo and 2,483 to Figshare.

Through the report obtained by the F-UJI tool, we can observe the degree of compliance with the FAIR principles of the Education datasets in the two repositories studied. The first thing we can observe is that it is in Figshare where the best FAIR percentage was obtained in terms of average (Table 1). Furthermore, it is also observed that not only does Figshare achieve better mean FAIR compliance than Zenodo, but when looking at the distribution of scores, none of the Zenodo datasets achieve even 50% compliance (Figures 2 and 3).

Table 1. Average of the percentage of compliance with the FAIR principles obtained by the two repositories analysed.

REPOSITORIES	ANALISED DOIS (num)	TOTAL FAIR %
FIGSHARE	2,483	56.86
ZENODO	1,772	43.30

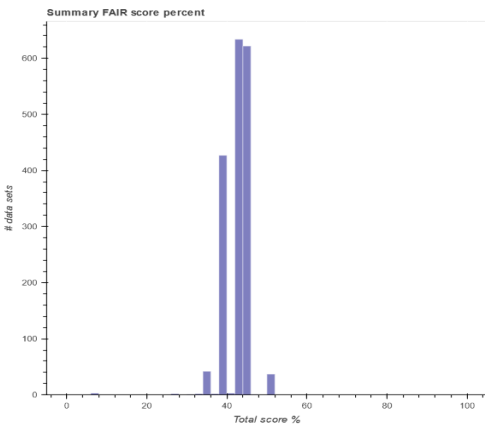

Figure 1. Percentage of compliance with FAIR principles in Zenodo repository.

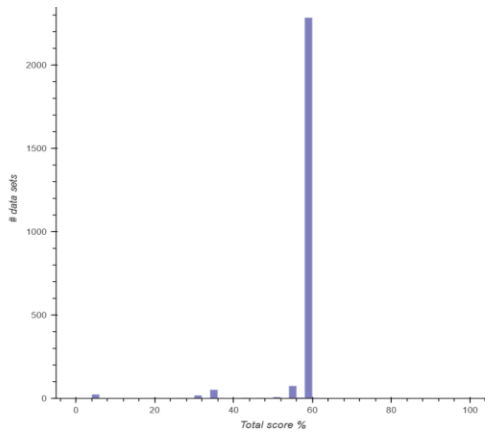

Figure 2. Percentage of compliance with FAIR principles in Figshare repository.

A similar result is observed when the percentage of compliance with the FAIR principles is looked at separately (Table 2). Overall, Figshare is better on three principles, only narrowly beaten by Zenodo on the Accessible principle. It is noteworthy that this Accessible principle is the lowest scoring of the two repositories, and it should be noted that only Findable achieves more than 50% compliance.

Table 2. Average of the percentage of compliance with each of the FAIR principles in a differentiated manner obtained by the two repositories analysed.

REPOSITORIES	F	A	I	R
FIGSHARE	84.3	33.1	47.8	48.42
ZENODO	65.4	33.3	31.3	35.6

Conclusions

The preliminary results of our study showed the issues that remain to be resolved, especially in relation to the FAIR principle of Accessibility, but also to Interoperability and Reusability. However, it is important to note that Findable is a principle that, at least in the two repositories studied, is acceptable.

Further in-depth analysis of the causes and possible solutions for the improvable score of the other three principles is crucial for the development of data sharing practices in Educational Sciences. This is an area that has a direct impact on the well-being of citizens and whose improvement in terms of research is necessary to make faster progress.

Acknowledgments

This work has benefited from an aid from the Ministry of Science and Innovation of the Government of Spain. State Research Agency. Ecological and Digital Transition Projects 2021. TED2021-131057B-I00.

References

Brennan, R. (2017). Challenges for Value-driven Semantic Data Quality Management. *Proceedings of the 19th International Conference on Enterprise Information Systems - Volume 1*. https://doi.org/10.5220/0006387803850392

Devaraju, A., Huber, R., Mokrane, M., Herterich, P., Cepinskas, L., de Vries, J., L'Hours, H., Davidson, J. & White, A. (2022). FAIRsFAIR Data Object Assessment Metrics (0.5). In *Zenodo* (Issue October). https://doi.org/10.5281/zenodo.6461229

Petrosyan, L., Aleixandre-Benavent, R., Peset, F., Valderrama-Zurián, J. C., Ferrer-Sapena, A. & Sixto-Costoya, A. (2023). FAIR degree assessment in agriculture datasets using the F-UJI tool. *Ecological Informatics*, *76*(May). https://doi.org/10.1016/j.ecoinf.2023.102126

Sixto-Costoya, A., Ferrer-Sapena, A., Aleixandre-Benavent, R., Peset, F., Valderrama Zurián, J. C. & Petrosyan, L. (2025). The compliance to FAIR principles of shared data in addiction. *Scientometrics*, *0123456789*. https://doi.org/10.1007/s11192-024-05227-5

Wilkinson, M. D., Dumontier, M., Sansone, S. A., Bonino da Silva Santos, L. O., Prieto, M., Batista, D., McQuilton, P., Kuhn, T., Rocca-Serra, P., Crosas, M. E. & Schultes, E. (2019). Evaluating FAIR maturity through a scalable, automated, community-governed framework. *Scientific Data*, *6*(1), 1–12. https://doi.org/10.1038/s41597-019-0184-5.