Promoting digital object sharing practices through data stewards

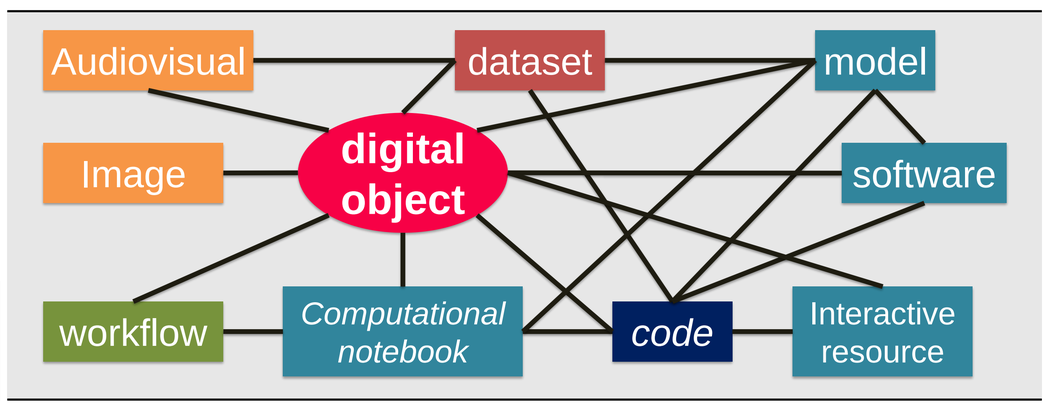
QR: Online version of this poster

Claire Jean-Quartier, Alexander Bardel, Birgit Söser, Ilire Hasani-Mavriqi

Research Data Management, Graz University of Technology, Austria

Sharing of digital objects in research remains essential for scientific reproducibility and integrity[1]. While making research data available is increasingly recommended by funders and institutions[2], there is still a variety in types of research output including computational materials such as software artefacts, models and the associated documentation necessary for the reproducibility of scientific studies[3,4]. Data stewards play a critical role in fostering related open science practices by promoting the formalized use of repositories for all types of digital research output.

Research Results - Digital Objects (DO)



Digital objects in science *- exemplary resource types as defined by datacite: general resources types (e.g. dataset, software, audiovisual, model, workflow, computational notebook, interactive resource) and customizable subtypes (e.g. code)[5]*

DOs which may comprise a simple data entity or a complex collection represent structured sequences of bits that are stored in repositories, described by metadata and ideally referenced by a persistent identifier (PID), and currently discussed more specifically as FAIR digital objects (FDO) for self-contained, typed, machine-actionable data packages[6].
Modern data-driven research is based on fully digital scientific methods and workflows comprising several steps of research planning, data collection, processing, analysis, communication and publication or archiving.
Data stewards still encounter misconceptions of the term “data” according to the view of some researchers. Researchers who have no or little experience with the concept of research data management (RDM) are frequently unaware of the research data lifecycle and do not recognize all necessary components of reproducible experiments. In particular, computational methods in natural, life, social or related engineering sciences often lack comprehensive documentation limiting the reusability of both computational methods and their outputs. Such research teams without trained computer scientists require support in managing, describing, and publishing research-related digital objects.

Reproducible computational results

Emphasis on practical reproducibility, not only data or code availability:

Reproducibility Checklist

- ❑ Methodology link between data, code, input, and output
- ❑ Dataset references & availability
- ❑ Code/software availability
- ❑ Readme explains purpose, input, output; installation/use steps
- ❑ Reuse conditions
- ❑ Analysis & transformation steps
- ❑ Software versions, dependencies, packages, compute environment
- ❑ Computing infrastructure/energy
- ❑ Version/release with PID

Packaging methods and structuring tools in practice can maximize the utility of digital objects as reusable resources. Further, recommendation checklists guide researchers using computational analyses, such as:

- FAIRER Aware Interactive Checklist https://fairerdata.github.io/FAIRER-Aware-REPRODUCIBILITY-Assessment/
- REFORMS: Consensus-based Recommendations for ML-based Science https://doi.org/10.1126/sciadv.adk3452
- Research code quality and execution https://doi.org/10.1038/s41597-022-01143-6
- FAIR Principles for research software https://doi.org/10.1038/s41597-022-01710-x
- Ten simple rules for training scientists to make better software. https://doi.org/10.1371/journal.pcbi.1012410

Data steward activities supporting DO sharing

Data stewardship has been introduced to improve RDM practices and promote data sharing for purposes of reuse in the interests of the scientific community and the public[7].

Additional Activities

- Individual guidance of researchers

Consulting and coordination, support formalized work with scientific digital objects underlining inter-mediary aims, policy compliance, long-term utility

QR: RDM dashboard summarizing support at TU Graz

- Support of project consortia

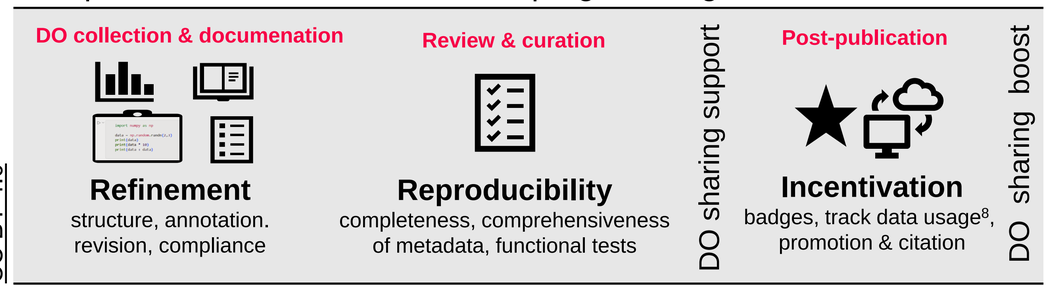
Training and methodical use of RDM tools increasing transparency and impact of research, respecting the legitimate interest of all stakeholders (licenses, startups, industry-funded research and patents)

- Curation of research metadata

Research data repository software offering a quality layer for record reviews for restricting the self-publishing of unreviewed drafts ensuring recommended file formats and essential metadata

QR: Demo of curation feature of repository.tugraz.at

Data stewards and institutional RDM teams support researchers by identifying, providing, maintaining, and promoting the use of RDM tools and approaches that enable the effective communication, sharing, and reuse of research results in accordance with FAIR principles. Their support may include computational infrastructures, documented strategies and guidelines, tailored workflows and applications, expert consultation and intermediation, awareness-raising events, and community exchange.
Above all, comprehensive documentation is prerequisite for the reuse of digital objects involving related concepts of strategic planning, structuring, annotation, testing, revision, and given the example of computational methods also literate programming.



Support and Boost of DO Sharing *– refinement and reproducibility support through data stewardship, additional incentivation by institutions & community*

References

1. Hanson, B., Sugden, A., Alberts, B. (2011) Making data maximally available. Science 331(6018), 649–649. https://doi.org/10.1126/science.1203354
2. Jean-Quartier, C., Kleinberger-Pierer, H., Zach, B., Hasani-Mavriqi, I., Pešec, L., & Schmikl-Reiter, C. (2024). Fostering Open Data Practices in Research-Performing Organizations. Publications, 12(2), 15. https://doi.org/10.3390/publications12020015
3. Jean-Quartier, C., Jeanquartier, F., Stryeck, S., Simon, J., Söser, B., & Hasani-Mavriqi, I. (2025). Sharing practices of software artefacts and source code for reproducible research. International Journal of Data Science and Analytics, 20(3), 2539-2550. https://doi.org/10.1007/s41060-024-00617-7
4. Jean-Quartier, C., Tscheppe, N., Millonig, S., Klambauer, L., Holzinger, A., Stryeck, S., & Jeanquartier, F. (2026). Reproducibility and Environmental Efficiency of Metabolomics Cancer Modeling. Applied Sciences, 16(2), 588. https://doi.org/10.3390/app16020588
5. Robinson-Garcia, N., Mongeon, P., Jeng, W., Costas, R. (2017) DataCite as novel bibliometric source. Journal of Infometrics 11(3). https://doi.org/10.1016/j.joi.2017.07.003
6. De Smedt, K., Koureas, D., Wittenburg, P. (2020) FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units. Publications 8, 21. https://doi.org/10.3390/publications8020021
7. Wendelborn, C., Anger, M., & Schickhardt, C. (2023). What is data stewardship? Towards a comprehensive understanding. Journal of biomedical informatics, 140, 104337. https://doi.org/10.1016/j.jbi.2023.104337
8. Jean-Quartier, C., Reichmann, S., Rey Mazón, M.& Hasani-Mavriqi, I. (2026). From Data Citations towards effective Valorization and Tracking of Data Reuse. Graz University of Technology. https://doi.org/10.3217/ybv6h-m2837

TU Graz – RDM Team
8010 Graz, Münzgrabenstraße 36/II
rdm.tugraz.at – mail: **rdmteam@tugraz.at**,