

Lessons Learned: RDM Tools in CRC 1430

Tuesday, September 30, 2025 2:15 PM (1h 30m)

Research data management (RDM) in the scope of collaborative research centers (CRC) faces a very specific set of challenges. CRCs usually comprise researchers from different disciplines leading to very heterogeneous requirements with regard to RDM tools, storage capacity and data types involved. This is especially true for consortia in the field of biomedical research which include a plethora of different specialties, such as bio-informatics, medicine, structural biology and cell biology. Each of these fields has its own workflows, data formats and even scientific traditions. Furthermore, data is often stored in disparate locations. From this, the necessity arises to integrate and consolidate all information in a centralized platform and to establish a meaningful data structure. Learning from experiences with other CRCs, carefully selected tools have been introduced to address this issue and to ensure data within the CRC 1430 follows FAIR principles [1].

A modular framework of RDM tools with focus on free and open-source software has been established based on previous experiences [2] and comprises of the **electronic lab notebook** (ELN) *eLabFTW* [3, 4], the **collaborative cloud storage** solution *Nextcloud* and an **internal repository** based on *Dataverse* [5]. This framework already allows for an efficient data management workflow including primary documentation of experimental results and storage and sharing of raw data within the project. Due to its biomedical scope with focus on cell-state transitions, microscopic images play a significant role in the overall data of the CRC 1430. However due to their large storage requirements and very diverse file format situation a central platform for storage, file format conversion, annotation and sharing of microscopic data was urgently needed. Hence, the **imaging management platform** *OMERO* was introduced [6, 7].

Finally, the internal repository *Dataverse* was envisioned to integrate information from all of the above-mentioned sources in a single platform to allow easy and intuitive access to research data. To this end, a **custom metadata schema** specifically tailored to the data produced within the scope of CRC 1430 was developed [8]. This allows researchers to directly and easily reference related data in *OMERO*, *eLabFTW* and *Nextcloud* from within the internal repository as well as to search the CRC 1430 data based on relevant keywords, experimental methods or project collaborations. The metadata schema is continuously refined based on the involved researcher's feedback. Somewhat unexpectedly, the repository roll-out triggered an intense and constructive discourse within the CRC 1430 about publication practices, especially with regard to data publication. We therefore conclude that *Dataverse* as an internal repository not only serves as staging area for later data publication but actively catalyzes discussion and awareness for this topic.

To encourage upload of research data predating the launch of the internal repository, full-day **data upload events** were held. During these events, RDM staff was available on-site to assist with any issues and questions that arose in the process and to engage in discussions about possible enhancements to the repository and the underlying metadata schema. As a socializing and networking component, food and drinks were provided throughout the event which encouraged participants to stay and discuss even after their data uploads were finished. These events prove to be very successful and will be a corner stone in upcoming on-boarding processes of research groups and individual researchers.

Another successful initiative is the introduction of **Data Champion Awards** [9]. These are awarded to early-career researchers within the CRC for outstanding RDM practices, including innovative use of the above-mentioned tools. The overall goal is to create more incentives that will contribute to the establishment of good RDM as a routine for a new generation of researchers.

References

1. Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). doi: 10.1038/sdata.2016.18
2. Rehwald, S. Research Data Services for INF Projects. in *Posters des SFB-INF Workshops: Let's share! Gute Praktiken und Fallstricke im Datenmanagement in SFBs und Forschungsgruppen* (2025). doi: doi:10.5281/ZENODO.14904441.
3. Carpi, N., Minges, A. & Piel, M. eLabFTW: An open source laboratory notebook for research labs. *JOSS* 2, 146 (2017). doi: 10.21105/joss.00146
4. Timm, H., Wittkamp, A. & Beyer, S. Automatisiertes Deployment von Elektronischen Laborbüchern mit Ansible. in *E-Science-Tage 2021* (heiBOOKS, 2022). doi: 10.11588/heibooks.979.c13744
5. King, G. An Introduction to the Dataverse Network as an Infrastructure for Data Sharing. *Sociological Methods & Research* 36, 173–199 (2007). doi: 10.1177/0049124107306660

6. Burel, J.-M. et al. Publishing and sharing multi-dimensional image data with OMERO. *Mamm Genome* 26, 441–447 (2015). doi: 10.1007/s00335-015-9587-6
7. Lama, S., Koch, J., Schulze, N. & Rehwald, S. Sharing, storing and analysing cell image data in a Collaborative Research Centre using OMERO. (2023) doi: 10.5281/ZENODO.10069503
8. Minges, A. & Rehwald, S. CRC 1430 Metadata Block Configuration for Dataverse. *DuEPublico: Duisburg-Essen Publications online, University of Duisburg-Essen, Germany* (2025). doi: 10.17185/DUEPUBLICO/83389
9. Minges, A. & Rehwald, S. OMERO and Data Champions in CRC 1430. *Informationsmaterialien der Research Data Services an der Universitätsbibliothek Duisburg-Essen* (2025) doi: 10.17185/DUEPUBLICO/83183.

Abstract

Poster

Author: Dr MINGES, Alexander (Universität Duisburg-Essen)

Co-author: REHWALD, Stephanie (Research Data Services (RDS), University Library, University of Duisburg-Essen, Duisburg, Germany)

Session Classification: Poster Session