



Contribution ID: 71

Type: **Presentation**

Sciebo RDS - reducing friction of FAIR data handling for researchers

Tuesday, March 7, 2023 10:30 AM (15 minutes)

Research Data Services (RDS) is a self-hosted cross-platform interoperability layer which allows research data to be curated, prepared and published directly from an EFSS solution such as Sciebo (ownCloud) or Sunet Drive (Nextcloud). It provides modular interoperability to external data repositories like the Open Science Framework (OSF), InvenioRDM (e.g., Zenodo), Harvard Dataverse, or Doris from the Swedish National Dataservice, each including domain-specific customizations.

Publishing data sets, corresponding metadata and persisting the information with a digital object identifier (DOI) is not only increasingly required by project funding entities such as the European Union or scientific journals, but also positively contributes to researchers' credibility and visibility. However, publishing research data often requires a specific data repository. This challenge is addressed by using Sciebo Research Data Services (RDS) as an interface between the enterprise file sync and share solution (EFSS) and the data repository. While research object crates (RO-Crate) are used as an intermediate lightweight package for the data and respective metadata, individual connector microservices ensure compliance with each publication service and easily developed support for additional data repositories. Metadata is added through Describo and based on schema.org annotations in JSON-LD and aims to make best-practice in formal metadata description accessible and practical for use in a wider variety of situations, from an individual researcher working with a folder of data, to large data-intensive computational research environments.

An integral part of the research data lifecycle process is the publication of research data, either by itself or as part of the supplementary information (SI) of a published article. Contributing to the Open Science movement, research data should be FAIR: findable, accessible, interoperable, and reusable. However, not only researchers' fears (e.g., fear of misuse of the data or fear of errors in the data or analysis) but also technical barriers (such as a lack of functionality or too long of a tool chain) are causing friction, often stopping researchers from adequately managing and sharing their research data.

Originally funded by the Deutsche Forschungsgemeinschaft (DFG) and implemented by WWU Münster in collaboration with the University of Duisburg-Essen, Sciebo RDS has sustainably evolved into a cross-platform solution that is used by several institutions and NRENs, notably SURF and Sunet. Its EFSS application has been ported from ownCloud to Nextcloud, thus also enabling cross-platform interoperability with the ScienceMesh.

To ensure that Sciebo RDS offers the highest possible benefit and ease of use for researchers, the development is accompanied by an iterative scientific evaluation process. This involves an extensive qualitative requirements analysis with researchers from various disciplines and qualitative as well as quantitative usability studies at different phases of prototyping.

To use Sciebo RDS, researchers can log on to their respective EFSS system and will find Sciebo RDS directly in the EFSS main menu. After connecting Sciebo RDS to appropriate data repositories (e.g. OSF or Zenodo), users are guided through a four-step data publication process, including:

1. the configuration of a research data project,
2. the collection and management of the data,
3. adding and editing the metadata, and
4. the transfer and publication of the data to external services.

Sciebo RDS integrates into academic enterprise file sync and share solutions and lets researchers collect data, collaborate on documents, and publish valuable scientific results directly from one simple solution.

Authors: HÖSELBARTH, Juri (University of Münster); HOFEDITZ, Lennart (University of Duisburg-Essen); FREITAG, Richard (SUNET)

Presenters: HÖSELBARTH, Juri (University of Münster); FREITAG, Richard (SUNET)

Session Classification: FAIR Data Management

Track Classification: Main session: User Voice: Novel Applications, Data Science Environments & Open Data