



Contribution ID: 84

Type: **Presentation**

## **iRODS and Sync&Share: Getting the best out of both worlds**

*Tuesday, January 28, 2020 11:00 AM (20 minutes)*

Within the Netherlands, iRODS is gaining substantial traction with universities and other research institutes as a tool to help manage large amounts of heterogeneous research data. In this context iRODS is usually used as middleware, providing value through data virtualization, metadata management and/or rule-driven workflows. This is then typically combined with other tools and technology to fully support the diverse needs of researchers, data stewards, IT managers, etc.

While integrations with other RDM tools are facilitated by iRODS' flexibility, a significant amount of work is usually still required to develop and test them with users in their specific context. For this reason SURF – as the collaborative ICT organisation for Dutch education and research – sees a role for itself to spearhead the development of such integrations as that effectively means pooling of resources which lowers the collective development cost and accelerates the pace of adoption.

In this contribution, we will focus on a recent project undertaken by SURF to explore the integration between OwnCloud and iRODS. OwnCloud is an open-source, “sync and share” solution to manage data as an individual or as a research team. OwnCloud is the technology behind two successful existing SURF products: SURFdrive and Research Drive. Offering a GUI, versioning, off-line sync and link-based sharing, it's functionality is in many ways complementary to iRODS. This makes integrating the two technologies attractive, yet there are several challenges in terms of file inventory synchronization, metadata management, and access control.

As an outlook into future work, this integration could be extended to support seamless publication of research data in trusted, long-term data repositories. Existing data publication workflows have many common tasks, but also significant variance in the “details” of how these tasks are strung together and how they need to be operationalized. To address this balance, we are exploring an approach that essentially abstracts data publication tasks into an overarching workflow framework, so as to allow for flexibility yet also benefit from standards and common patterns.

**Primary author:** WOLFSHEIMER, Stefan (SURFsara)

**Co-author:** Dr TROMPERT, Ron (SURFsara)

**Presenter:** Dr TROMPERT, Ron (SURFsara)

**Session Classification:** Fabric and platforms for Global Science

**Track Classification:** Scalable Storage Backends for Cloud, HPC and Global Science