

Open Source Tools Driving New Services and Workflows at DataCite

Patricia Cruse (1) and Britta Dreyer (2)

(1) DataCite, orcid.org/0000-0002-9300-5278, (2) DataCite, orcid.org/0000-0002-0687-5460

The Open Science community has a strong preference for Open Source software. DataCite is a non-profit organization that provides DOI registration services for scholarly content with an emphasis on research. Over the past several months DataCite has undertaken a major service development cycle. An important part of this activity is engaging with our global community of users to make them an ever present part of our processes. We will present how DataCite is using and developing Open Source software, with a focus on the following topics: a) when to use Open Source software, workflows for Open Source product development, and collaboration in Open Source projects. All software developed by DataCite is released under an Open Source software license. We prefer Open Source solutions for the software we adopt, but also use commercial services when not strategic for DataCite's core mission (communication tools, etc.). DataCite's new development workflow is based on the ideas of transparency, feature development driven by users and user stories, and storing all data relevant for product development in a central open repository. We use GitHub issues and milestones, custom open APIs developed by DataCite, and waffle.io Kanban boards for this workflow, inspired by a similar approach used by the Dataverse Open Source community. The community of DOI registration service providers is small but DataCite is developing tools that are of interest to a broader research community. We focus on reusable libraries helpful to the scholarly community (e.g. metadata conversion), but have also built more generic tools relevant for infrastructure providers (e.g. a network client for REST APIs). Finally, we have started to register DOIs for our software, working with the Force11 software citation community.