Contribution ID: **75**   Type: **Poster**

DataLad: 10+ years of academic software development

Wednesday 26 February 2025 19:40 (20 minutes)

DataLad (Halchenko et al., 2021 [1]) is free and open source software for managing digital objects and their relationship built on top of Git and git-annex. Its initial commit in 2013 marked the beginning of a more than 10 year long academic software history so far, supported by various grants, institutions, and underlying research endeavors. Over time, the software became an extendable ecosystem, addressing a broad range of data logistics challenges in a core library and many extension packages, growing both in features and contributor community. In turn, it also sparked development and grant support in git-annex, a crucial software with a bus factor of 1. Navigating the research software waters of changing affiliations, developer churn, research obligations, and a modular architecture that offers flexibility, but also bears a potential for complexity and fragility, has never been easy.

In this contribution, we want to give a case-study-like overview of the lifetime of this research software so far, reflect on the design and development decisions we have made over the years and their advantages or shortcomings, share lessons learned, and give an outlook into the future of the software ecosystem.

[1] https://joss.theoj.org/papers/10.21105/joss.03262

I want to participate in the youngRSE prize

Primary authors: WAGNER, Adina Svenja (INM-7); WAITE, Alex (Forschungszentrum Jülich, INM-7); POLDRACK, Benjamin (INM7); MÖNCH, Christian (Research Center Jülich/INM-7); WAITE, Laura; Dr WIERZBA, Małgorzata (Nencki Institute of Experimental Biology); HANKE, Michael; Dr SZCZEPANIK, Michał (Forschungszentrum Jülich, INM-7); HEUNIS, Stephan; Dr HALCHENKO, Yaroslav (Dartmouth College, NH, USA)

Presenter: HANKE, Michael

Session Classification: Poster and Demo Session together with Reception

Track Classification: Research Software: software sustainability