FAIR principles and benchmarking of research software as essential data

Authors:
Jörg F. Unger[1]*, Tomislav Maric[2], Jan Linxweiler[3] and Bernd Flemisch[4]
*Lead presenter
[1]joerg.unger@bam.de, Bundesanstalt für Materialforschung und -prüfung
[2]maric@mma.tu-darmstadt.de, TU Darmstadt
[3]j.linxweiler@tu-braunschweig.de, TU Braunschweig
[4]bernd.flemisch@iws.uni-stuttgart.de, Universität Stuttgart

Abstract:
Software is an integral part of research data. The RDA working group 'FAIR for Research Software (FAIR4RS)' has extended the FAIR principles, particularly tailored to research software [1]. The challenge of integrating these general principles into software for numerically solving physics based model is solved by each community and often even each individual researcher with a different approach.
A particular challenge of research software in the simulation domain is its complex dependencies because it is usually built upon many other software contributions. Another challenge for a FAIR implementation is the integration of these modules into simulation workflows that define a sequential order for the execution of individual simulations tasks usually described by a graph, as well as the compute environment. In NFDI4Ing, a comparison of selected workflow tools has been performed [2]. When making research software FAIR, all these levels (individual module, simulation workflow, result of a specific run of the workflow) are to be considered. Several approaches such as RO-Crate with an extension Workflow Run Crate [3], the Citation File Format [https://citation-file-format.github.io/], the common workflow language with CWLProv for describing workflow provenance [4], and many more are currently being developed. In NFDI4ING, there are related activities such as a search engine to find research software matching certain attributes [5]. The goal of a working group on the BaseNFDI level would be to establish standards on how FAIR simulation workflows can be developed and shared within the communities, related to both the workflows as well as the results of an individual run.
Another requirement is the interoperability with domain-relevant community standards for data exchange. Even though ongoing efforts in our community aim to build standardized file formats, e.g., the VMAP standard for Finite Element Models, implementations are often slightly different, making it a challenge to exchange one interoperable software tool with another one while being able to reproduce the results. From our perspective, benchmarking of tools with well-defined test data is thus mandatory, but a general platform for this purpose is missing. This includes e.g. building an editorial board of experts for defining benchmarks, a platform for hosting codes with appropriate administrative structures to ensure code quality, a tool independent and machine-readable description of benchmark problems, an execution platform for running the benchmarks and tools and resources (e.g. a jupyterhub) to publish/visualize and explore the results within the community. For this purpose, a tool for comparing field data [6] has been developed, and a platform for integrating benchmark tests is planned to be developed. Those challenges are not specific to the engineering domain, and as such, we hope to establish a working group in the base community to build a software benchmarking framework for simulation software to address these challenges in a cross-disciplinary manner.

[1] N. P. Chue Hong et al., „FAIR Principles for Research Software (FAIR4RS Principles)“, Research Data Alliance, 2021, doi: 10.15497/RDA00068.
[2] P. Diercks et al., „Evaluation of tools for describing, reproducing and reusing scientific workflows“, Universitäts- und Landesbibliothek Darmstadt, 2023, doi: 10.48694/INGGRID.3726.
[3] S. Leo u. a., „Recording provenance of workflow runs with RO-Crate“. arXiv, 2023. doi: 10.48550/ARXIV.2312.07852.
[4] F. Z. Khan et al., „Sharing interoperable workflow provenance: A review of best practices and their practical application in CWLProv“, GigaScience, Bd. 8, Nr. 11. Oxford University Press (OUP), Nov. 01, 2019. doi: 10.1093/gigascience/giz095.
[5] V. Seibert, A. Rausch, S. Wittek, „Betty's (Re)Search Engine: A client-based search engine for research software stored in repositories“, Universitäts- und Landesbibliothek Darmstadt, 2023, doi: 10.48694/inggrid.3953.
[6] D. Gläser et al., „fieldcompare: A Python package for regression testing simulation results“, Journal of Open Source Software, Bd. 8, Nr. 81. The Open Journal, S. 4905, Jan. 31, 2023. doi: 10.21105/joss.04905.

Keywords: [workflows, research software, benchmarking]