

The HMC FAIR Data Dashboard: A Data-Driven Approach to Monitor and Improve FAIR Data in the Helmholtz Association

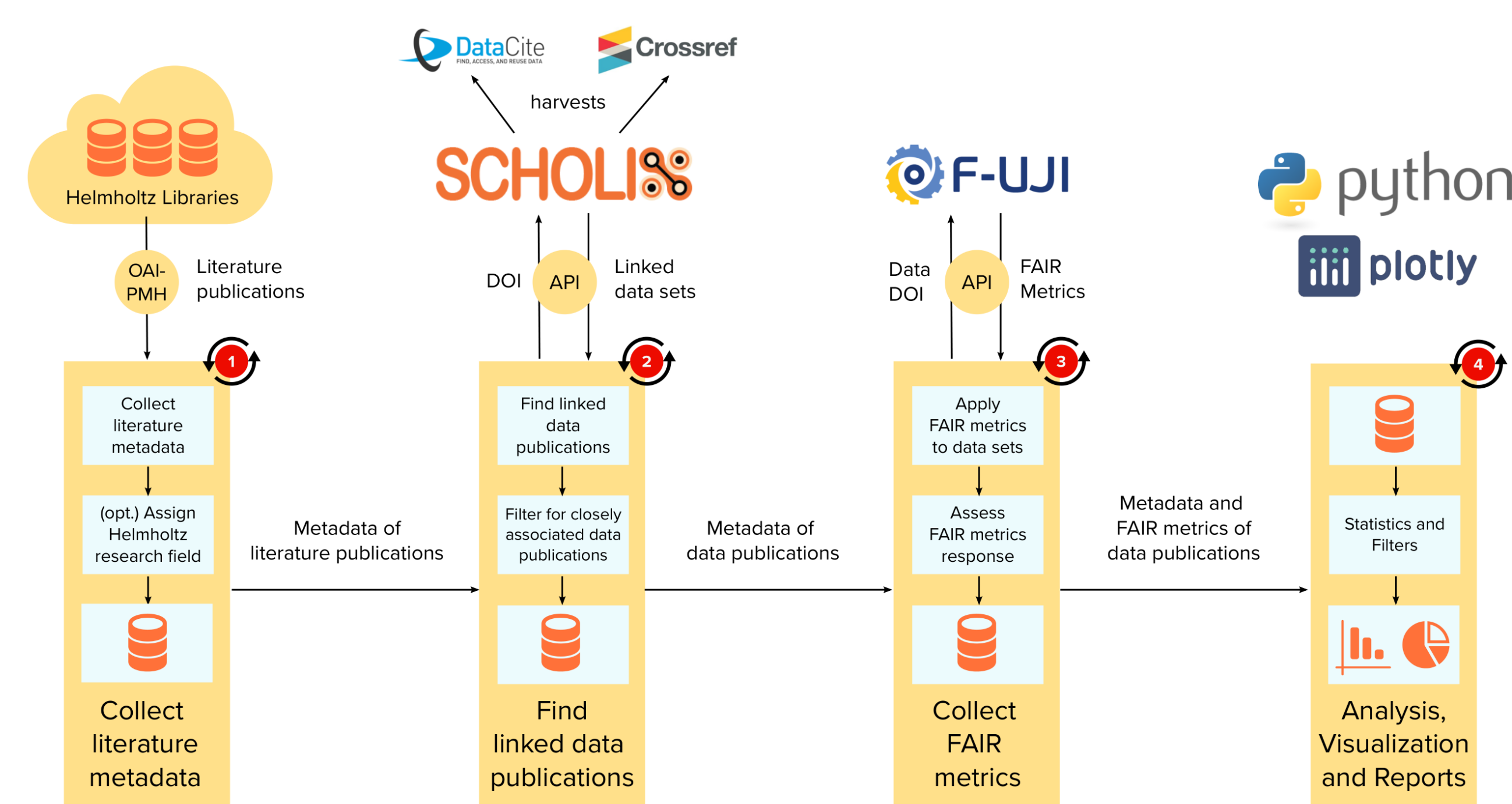
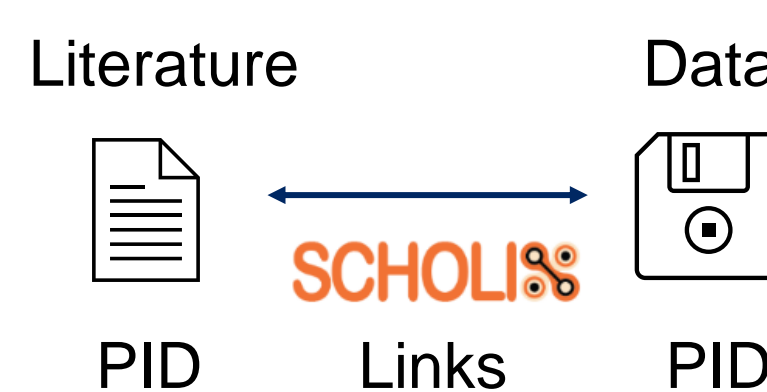
Mojeeb R. Sedeqi^{1,2}, Pascal Ehlers^{2,3}, Alexander Schmidt¹, Vivien Serve^{1,2}, Astrid Gilein¹, Tempest Glodowski¹, Gabriel Preuß^{1,2}, Oonagh Mannix^{1,2}, Markus Kubin^{1,2,*}

1 Helmholtz-Zentrum Berlin für Materialien und Energie
2 Helmholtz Metadata Collaboration
3 German Aerospace Center
Contact: * markus.kubin (at) helmholtz-berlin.de

- Modular approach to find, assess, and analyze data publications in a federated research organization like the Helmholtz Association.
- Open data is identified by harvesting literature metadata from library databases and by finding linked datasets via SCHOLIX-links. The F-UJI framework is used as a first approach to FAIR assessment.
- All code is open source and reusable by the public. Please contact us if you are interested in contributing to the project!

Find and assess open data

- Harvest literature metadata from center libraries via OAI-PMH.
- Find linked data publications via SCHOLIX. [1]
- Automatized FAIR assessment with F-UJI. [2]
- Complementary manual FAIR assessment based on the RDA FAIR Data Maturity Model. [3]
- Data-driven community and infrastructure counselling.



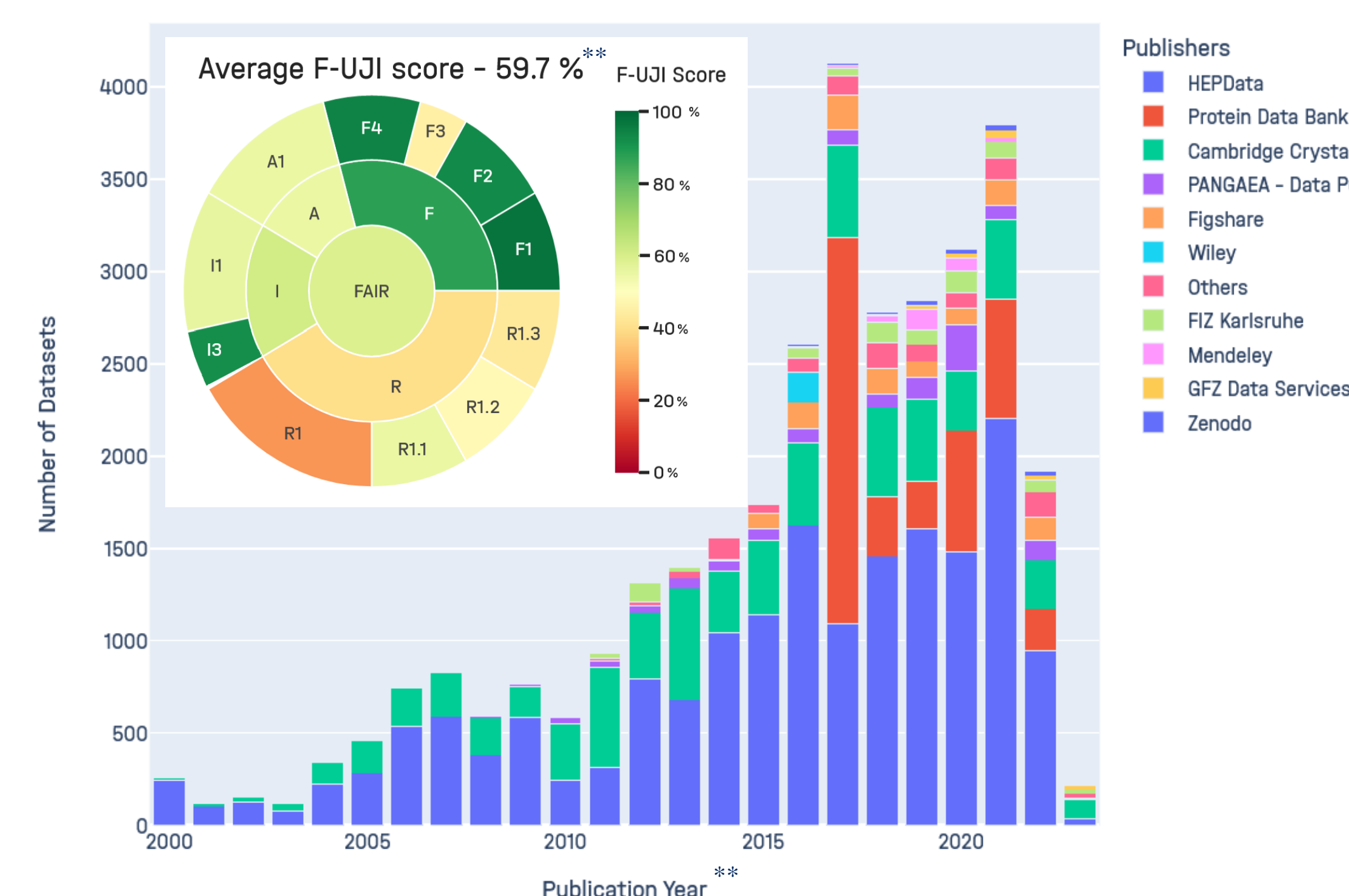
Dashboard workflow and activities

Dashboard version 1.0

- Monitoring Open and FAIR data in a federated research organization. [2]
- Analyze open and FAIR data by center / research field / repository
- Interactive analysis and self-learning tool
- Descriptive with multi-language support

Dashboard version 2.0

- Updated F-UJI Scoring
- Expanded Assess My Data feature
- Centralized F-UJI Server at DESY
- Improved User Interface
- Include toolbox version 2.0 with:
 - Management of database structure
 - Automatic weekly updates
 - Performance improvement (~10X)
- Expanded Coverage
- Implement E2E testing
- Complies with the standard best practices (OpenSSF, REUSE)
- Standard Documentation and Guidelines
- Now with contributions of Hub AST!



(**) Data and analyses based on data collected in Q1/2-2023

Outlook for version 3.0

- Further improve dashboard features and user experience.
- Complementary data sources beyond SCHOLIX: harvest repositories [6].
- Complement FAIR assessment, e.g., FAIR enough or analog methods [7].
- Communication with libraries and infrastructure: harmonize metadata.
- Information basis for counseling towards FAIR data and infrastructure.

Join us!

- 17 centers connected to the dashboard
- >360k literature publications harvested
- >40k data publications found using SCHOLIX
- >85 publishers (repositories) identified
- >30 commits/month on Gitlab since v.1.0.0

<https://fairdashboard.helmholtz-metadaten.de>

<https://codebase.helmholtz.cloud/hmc/hmc-public/FAIR-dashboard>

<https://doi.org/10.5334/dsj-2024-041>



References

- [1] A Burton et al.: The Scholix Framework for Interoperability in Data-Literature Information Exchange (2017). doi: 10.1045/january2017-burton
- [2] M Kubin et al., Launch Meeting: HMC Dashboard on Open and FAIR Data in Helmholtz (2023). doi: 10.5281/zenodo.7693377
- [3] A Devaraju and R Huber: F-UJI - An Automated FAIR Data Assessment Tool (2020). doi: 10.5281/zenodo.4063720
- [4] M Wilkinson et al.: The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). doi: 10.1038/sdata.2016.18
- [5] A Devaraju et al.: FAIRsFAIR Data Object Assessment Metrics (2020). doi: 10.5281/zenodo.6461229
- [6] P Videgain Barranco et al., How FAIR is my data? Benefits and pitfalls of quantitative assessment of FAIRness (2022). doi: 10.5281/zenodo.7313153
- [7] M Kubin, G Günter: Assessing the FAIRness of a prototypical PaN instrument at BESSY II (2022). doi: 10.5281/zenodo.6059994