HMC Dashboard on Open and FAIR Data in Helmholtz

Mojeeb Sedeqi[1,2], Alexander Schmidt[1], Vivien Serve[1,2], Astrid Gilein[1], Tempest Glodowski[1], Gabriel Preuß[1,2], Oonagh Mannix[1,2], Markus Kubin[1,2,*]

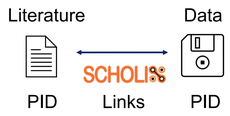
[1] Helmholtz-Zentrum Berlin für Materialien und Energie, [2] Helmholtz Metadata Collaboration.

* markus.kubin (at) helmholtz-berlin.de

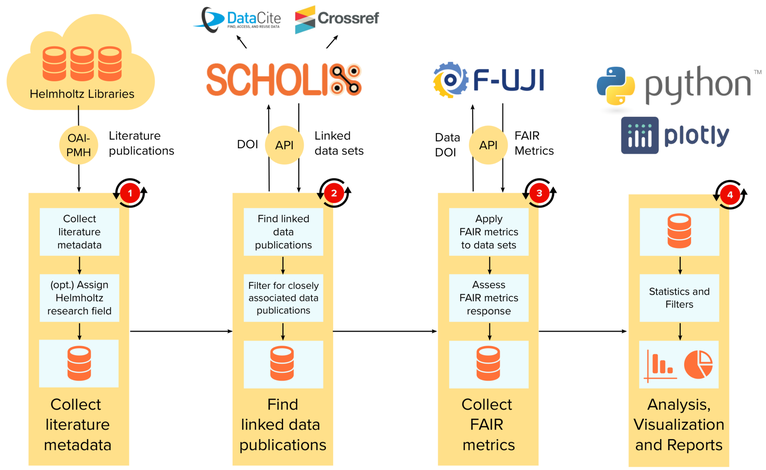
- Modular approach to find, assess and analyze data publications in a federated research organization like the Helmholtz Association.
- Open data is identified by harvesting literature metadata from library databases and by finding linked datasets via SCHOLIX-links. The F-UJI framework is used as a first 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!

Modular approach: Find and assess open data

- Publication-based approach using SCHOLIX [1]



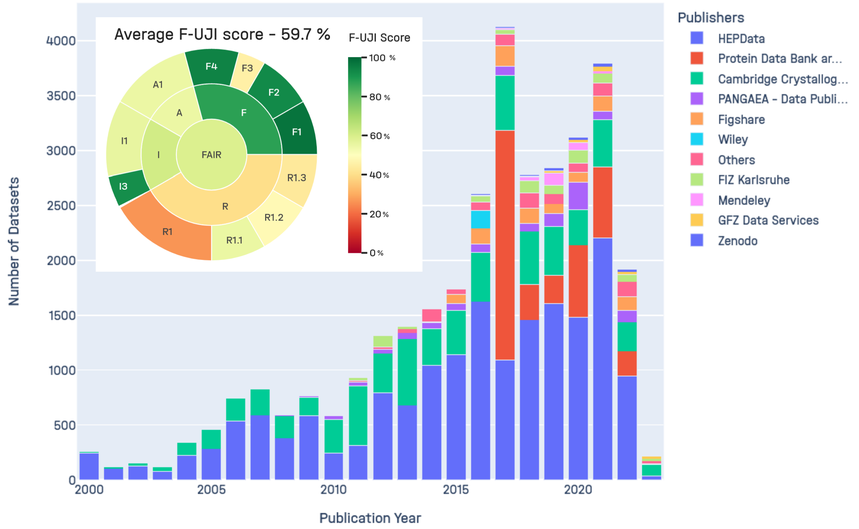
- Pros and cons of this publication-based approach [2]
  - ✓ Harvest from well-curated, federated literature databases via OAI-PMH.
  - ✓ Find associated data publications with SCHOLIX. (“*IsSupplementedBy*”)
  - ✓ Find data repositories used by the research communities.
  - – SCHOLIX approach is biased, partially erroneous and incomplete.



- Automatized FAIR assessment with F-UJI. [3]
  - Probing 13/15 FAIR principles [4] with 17 FAIRsFAIR metrics. [5]
  - Subject of continuous development (metrics and tools)
  - Focus on machine-actionable aspects.
  - Provide guidance to systematically improve FAIR data and infrastructure.

A dashboard to engage communities

- 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



Outlook

- 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!

- **15** centers connected to the dashboard
- **>300k** literature publications harvested
- **>32k** data publications found using SCHOLIX
- **>80** 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

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