Community Building for Research Data Repositories in Helmholtz

Nina Leonie Weisweiler[1,2] // Constanze Curdt[1] // Roland Bertelmann[2]

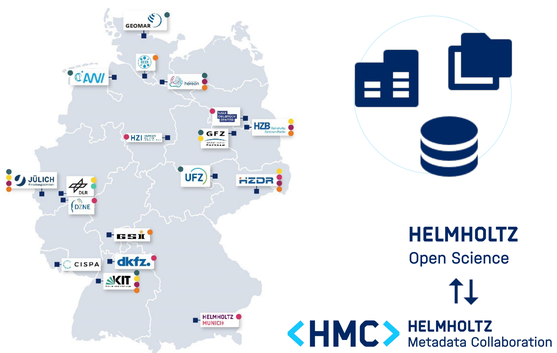
[1] Helmholtz Metadata Collaboration (HMC) //
[2] Helmholtz Open Science Office

Networking research data repository providers in the Helmholtz Association

The Helmholtz Metadata Collaboration (HMC) and the Helmholtz Open Science Office have launched a joint initiative to strengthen and connect research data repositories in the Helmholtz Association, and to increase their visibility in the international research landscape.

Which research data repositories can be mapped to the Helmholtz Centers?

With the joint initiative, we target repositories operated under the auspices of Helmholtz Centers. After searching re3data, the Registry of Research Data Repositories, we found nearly 100 such infrastructures associated with Helmholtz Centers, varying widely in size, purpose, and community practices.

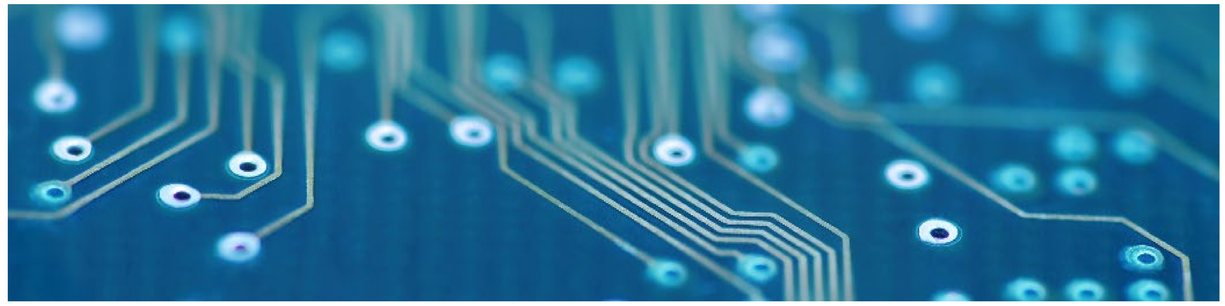


Connecting infrastructure providers across the Helmholtz Centers

What can we learn about Helmholtz-based repositories?

Tests have shown that only about 20% of these Helmholtz-associated repositories can be harvested automatically. According to re3data, about 40% of them support one or more persistent identifier (PID) systems and provide one or more API options. About 60% offer some level of quality management and 8 of them are certified, mostly via CoreTrustSeal.

Through assessment methods, such as the F-UJI Automated FAIR Data Assessment Tool, systematic gaps can be identified and communicated to repository operators (e.g., missing license information). This information can be used to effectively counsel data infrastructures towards implementing FAIR data management practices.



How can we reach and support these repositories to successfully connect them to the Helmholtz FAIR data space?

As part of the joint initiative, selected repositories will be contacted and offered support and community services, such as workshops and consulting. An important goal is to network repository operators with each other and to promote the exchange of experiences and best practices.

With this initiative, we aim to support research data repositories in providing persistent identifiers, appropriate APIs, and high-quality metadata so that they can connect to the Helmholtz FAIR data space and other national and international data ecosystems and increase the visibility of their services and data.

Research data repositories

- Support research data management and data curation, e.g. by supplying a persistent identifier and structured metadata to deposited datasets.
- Are central to the further development of an internationally networked research data infrastructure, taking into account the FAIR principles (Findable, Accessible, Interoperable, Reusable).
- Preserve valuable research data for the long term.
- Can vary widely in terms of their organizational structures, funding mechanisms, purposes, curation levels and scope.