Recommendation M1.3

Data Infrastructures

Recommendation to Helmholtz data stewards, repository maintainers and developers to implement personal ORCID as a reference to people in technical infrastructures

Description

[Status: Under development, Date: 2025/12/16 14:39, Version: 003]

Motivation for this Recommendation:

The Helmholtz Association is determined to make their data available according to the FAIR principles, thus making it findable, accessible, interoperable and reusable. In order to achieve interoperability of datasets among various data infrastructures (DIS) within the Helmholtz Association, a common and agreed procedure to **refer to people** within and across the DIS is needed.

In order to be able to uniquely and sustainably identify both researchers and employees in data infrastructures and repositories in the Helmholtz Association, the respective person should always be referenced with a persistent identifier (PID) (see recommendation M0).

For the Helmholtz Association we recommend to use ORCID to refer to people and contributors to resources in data infrastructures and repositories of the Helmholtz Association wherever possible (see recommendation M1.0).

To be able to implement this measure, several activities need to be conducted by different stakeholder groups. This recommendation M1.3. calls for activity of the data stewards, repository maintainers and developers.

Recommendation

It is recommended that data infrastructures (data repositories, data bases) should:

1. record an ORCID with any person registered in conjunction with the metadata of datasets, publications, instruments and alike where possible.
2. treat ORCID metadata as the primary source of truth and update their own metadata accordingly.
3. Optionally inform persons if they think the metadata registered with the ORCID is not accurate, or request permission to update the ORCID metadata (see also M1-1).
4. consider ORCID as an attribute in identity and access management systems (IAM), authentication and authorisation infrastructures (AAI) or community attribute services

Binding Convention:

	mandatory	conditional	optional
Helmholtz FAIR Principle	if ORCID is available		

Precondition for Implementation:

[Precondition 1]: The ORCID Registry is available for all researchers, maintained and further developed.

Related Recommendations

Parent: M1.0

Dependent: none

Other: M1.1, M1.2

Contributors

Martin Abbrent ORCID: 0000-0003-1252-9107 UFZ,

Emanuel Söding ORCID: 0000-0002-4467-642X GEOMAR (lead)

Content

1. Explanation of the Background and Benefits of the Recommendation

The Helmholtz Association is determined to make their data available according to the FAIR principles, thus making it findable, accessible, interoperable and reusable. In order to achieve interoperability of datasets among various data infrastructures (DIS) within the Helmholtz Association, a common and agreed procedure to **refer to people** within and across the DIS is needed.

In order to be able to uniquely and sustainably identify both researchers and employees in data infrastructures and repositories in the Helmholtz Association, the respective person should always be referenced with a persistent identifier (PID) (see recommendation M0).

For the Helmholtz Association we recommend to use ORCID to refer to people and contributors to resources in data infrastructures and repositories of the Helmholtz Association wherever possible (see recommendation M1.0).

To be able to implement this measure, several activities need to be conducted by different stakeholder groups. This recommendation M1.3. calls for activity of the data stewards, repository maintainers and developers.

2. Possible alternative solutions

see recommendation M1.0

3. Consideration of the advantages and disadvantages of implementing the recommendation

Implementing ORCID requires updating metadata schemas and workflows to support consistent recording and validation of ORCID iDs. Synchronizing metadata with ORCID records can be technically demanding, especially when dealing with discrepancies or outdated profiles. Infrastructures must also handle user consent, as not all researchers have ORCID iDs or keep them current. Integrating ORCID into identity and access management (IAM/AAI) systems adds complexity around authentication flows, attribute handling, and privacy policies. Finally, relying on an external service like ORCID demands ongoing technical maintenance, monitoring of API changes, and alignment with broader PID infrastructure.

When an ORCID is not available, data infrastructures should still record the person’s full name and leave the ORCID field empty or marked as pending, allowing for future updates. To support post-hoc linking, systems should enable users to add ORCIDs later, for example through a “claim your record” function. As a fallback, institutional identifiers can be used to reference the person within identity and access systems. Finally, infrastructures should implement regular quality control processes to identify missing ORCIDs and enrich metadata over time, either manually or through automated tools such as the ORCID Public API.

4. The Recommendation

It is recommended that data infrastructures (data repositories, data bases) should:

1. record an ORCID with any person registered in conjunction with the metadata of datasets, publications, instruments and alike where possible.
2. treat ORCID metadata as the primary source of truth and update their own metadata accordingly (see note below).
3. Optionally inform persons if they think the metadata registered with the ORCID is not accurate, or request permission to update the ORCID metadata (see also M1-1).
4. consider ORCID as an attribute in identity and access management systems (IAM), authentication and authorisation infrastructures (AAI) or community attribute services.

Note on 2. + 3.: treating ORCID metadata as the primary source of truth is a conceptual decision on the issue, who is ultimately responsible for personal data. In this model we assume, that each person / contributor, identified by an ORCID is solely responsible for their own data and should only made

aware of incorrect information, not forced to update it. Technically, institutions could maintain their own records of these personal data. This, however, leads to several problems: 1. the maintenance of this data is very difficult and keeping it current almost impossible; 2. by keeping personal data of people, data privacy becomes an issue, as people cannot control, what data about them is shared with others. This can be avoided by delegating the responsibility for personal information to the persons themselves, at the cost of not being in full control of that data.

Note on 4.: Identity and access management systems (IAM), authentication and authorization infrastructures (AAI) or upcoming community attribute services like defined in the AARC blueprint architecture are suitable environments for enriching user information, for example with an ORCID attribute. The mechanisms to control forwarding of private and personal user information allows the users to decide if they agree with that. As an attribute of the user objects the usage of ORCID within scientific software applications would be simplified and encouraged.

5. Naming of communities that have already implemented the recommendation

6. Documentation of the test to validate correct implementation

7. Examples of Instances

8. Further Information

References

[1] ORCID Terms of use: https://info.orcid.org/terms-of-use/

Relevant Community Recommendations

9. History of this document

From:
https://earth-and-environment.helmholtz-metadaten.de/wiki/ - **HMC Earth and Environment Community Wiki**

Permanent link:
https://earth-and-environment.helmholtz-metadaten.de/wiki/doku.php?id=wiki:1_fair-building-blocks:0_pid:0_orcid:m1.3&rev=1765895953

Last update: **2025/12/16 14:39**