Recommendation S0

Recommendation for implementing harmonized semantic concepts in data infrastructures and products

Description

Status: Under development, Date: 2025/05/07 10:18, Version: 001

Motivation for this Recommendation

Using shared, community-endorsed vocabularies enables unambiguous naming, supports the alignment and integration of heterogeneous datasets, and facilitates data discovery and reuse. Within the framework of the FAIR Guiding Principles (Wilkinson et al., 2016), such annotation practices are essential not only for improving interoperability — by enabling machines to interpret and connect data across domains — but also for enhancing findability and reusability through the use of persistent, well-defined, and widely recognized semantic references.

Recommendation

Data infrastructures should ensure the annotation of the large majority of metadata using standardized terms within metadata systems — such as data repositories, sensor registries, electronic lab notebooks, or other platforms that manage or reference data, including descriptions of files stored outside formal repositories — by applying terms from established and, where applicable, FAIR-compliant controlled vocabularies (e.g., ontologies, taxonomies, or standardized terminologies) to promote semantic consistency, clarity, and interoperability.

Binding Convention

	mandatory	conditional	optional
Helmholtz FAIR Principle		Annotation is mandatory when appropriate controlled vocabularies, expert recommendations on their use, and the necessary domain expertise are available, and when the systems support annotation technically.	

Precondition for Implementation

Comprehensive metadata annotation is only effective if there is consensus within a research community about which controlled vocabularies and semantic resources best meet the community's

needs, and if these resources have clear governance, provenance, and documentation. Furthermore, they should be available and maintained over the long term (at least 5 years) and cover the vast majority of requirements.

Contributors

Content

1. Explanation of the Background and Benefits of the Recommendation

About

History and structure

Current Use of ...

Motivation

2. Possible alternative solutions

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

(quality of content, limitations, interoperability, sustainability: expected future dissemination / technical availability / funding)

4. The Recommendation

Data infrastructures should ensure the annotation of the large majority of metadata using standardized terms within metadata systems — such as data repositories, sensor registries, electronic lab notebooks, or other platforms that manage or reference data, including descriptions of files stored outside formal repositories — at the time of metadata creation or management, by applying terms from established and, where applicable, FAIR-compliant controlled vocabularies (e.g., ontologies, taxonomies, or standardized terminologies) to promote semantic consistency, clarity, and interoperability.

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

Relevant Community Recommendations

9. History of this document