The Helmholtz Road-to-FAIR Strategy

The Helmholtz Metadata Collaboration pursues the vision of a harmonized FAIR research data space. However, the FAIR Guiding Principles define the desired properties of research data and metadata without prescribing how these properties should be realized in practice. They do not determine which standards, technologies, responsibilities, or organizational processes should be adopted. A complementary implementation strategy is therefore required to translate the general objectives of FAIR into coordinated and actionable measures.

The **Road-to-FAIR Strategy** provides such a framework. It establishes a common narrative through which technical, semantic, organizational, and procedural activities can be related to one another, prioritized, and evaluated. This is particularly important because research data management encompasses a broad range of interdependent topics, including metadata quality, semantic interoperability, institutional responsibilities, technical infrastructures, and readiness for emerging forms of data-intensive research.

The strategy addresses four closely connected areas. It:

- defines and prioritizes implementation objectives through a common set of **Road-to-FAIR building blocks**;
- identifies the stakeholder groups involved in research data management and clarifies their respective responsibilities;
- establishes workflows through which relevant information can be created, maintained, and transferred between responsible actors and systems; and
- specifies the technical data ecosystem required to support these responsibilities and workflows.

The Road-to-FAIR Strategy consequently treats research data management as a distributed and coordinated activity. Scientific information is generated at different stages of the research process and by a wide range of actors, many of whom may not primarily identify themselves as participants in research data management. The strategy seeks to ensure that this information is captured close to its point of origin, maintained by appropriate authoritative sources, and made available for repositories, aggregation services, scientific models, simulations, and machine-assisted analysis.

1. Defining the FAIR Building Blocks

The FAIR building blocks define the principal implementation objectives of the Road-to-FAIR Strategy. They provide a structured means of assessing which elements of FAIR are already supported and where further coordinated action is required.

Within the Helmholtz Research Field Earth and Environment, Findability and Accessibility are comparatively well supported for data deposited in established institutional or disciplinary repositories. Repository registries and discovery services, including re3data, FAIRsharing, and OpenAIRE, assist users in identifying appropriate repositories and locating relevant datasets. Open-science policies further support access to research data, while the provision of metadata for restricted datasets enables their discovery and communicates the conditions under which access may be obtained.

The principal remaining challenges concern the **Interoperability** and **Reusability** of heterogeneous

research data. Repositories, projects, research programmes, disciplines, and observational networks frequently apply different standards and procedures when describing their data. As a result, information may be difficult to combine across repositories and, in some cases, even within individual infrastructures. Projects that integrate heterogeneous datasets must therefore devote substantial effort to data cleaning, semantic harmonization, metadata enrichment, and the reconciliation of incompatible structures.

The Road-to-FAIR building blocks define practical measures intended to reduce these barriers. Measures supporting Interoperability include:

- the consistent use of persistent identifiers for recurring and systematically described entities;
- the coordinated application of agreed semantic concepts and vocabularies;
- the adoption of standardized interfaces, protocols, schemas, and exchange formats; and
- the provision of data and metadata through uniform, machine-actionable representations.

Measures supporting Reusability include the systematic documentation of provenance, licences, access conditions, data quality, and other contextual information required to assess whether data are suitable for a particular purpose.

The building blocks do not constitute a single technical specification. Rather, they define areas in which institutions, communities, and infrastructures must establish and implement shared agreements.

2. Defining Stakeholders and Responsibilities

The Road-to-FAIR Strategy regards research data management as a distributed institutional responsibility. It cannot be assigned exclusively to individual researchers, data managers, or repositories.

The information required to describe a research dataset is created and maintained by different stakeholder groups. These may include researchers, technicians, laboratory and field personnel, administrative units, libraries, data stewards, repository operators, infrastructure providers, and organizational management. No single stakeholder normally possesses all information required to produce a complete and accurate description of a digital research object.

Responsibilities should therefore be assigned to those actors or systems that are best positioned to create, verify, and maintain the relevant information. Institutions must communicate these expectations clearly and, where appropriate, formalize them through policies, role descriptions, and agreed procedures. The identification of stakeholder groups and their responsibilities thus provides the organizational foundation for subsequent implementation.

3. Establishing Coordinated Workflows

Distributed responsibilities require coordinated workflows. The Road-to-FAIR Strategy therefore defines procedures through which information is captured close to its point of origin, maintained by an appropriate authoritative source, and transferred to downstream information systems.

This approach differs from workflows in which researchers are asked to reconstruct all required

metadata only at the point of data publication. Information concerning people, organizations, projects, instruments, samples, methods, licences, or administrative conditions may already exist in dedicated systems and should, wherever possible, be obtained from these authoritative sources.

Relevant information should therefore move continuously through interfaces connecting research processes, administrative systems, technical services, and data infrastructures. Such workflows reduce redundant data entry, improve consistency, and distribute the effort of metadata creation across the research data lifecycle.

Their implementation requires both institutional support and technical services capable of exchanging, validating, aggregating, and enriching metadata. Workflows must also define how information is updated, who is responsible for correcting errors, and how changes are propagated across connected systems.

4. Providing an Interoperable Data Ecosystem

The final component of the Road-to-FAIR Strategy is the definition and provision of a coordinated technical data ecosystem. This ecosystem comprises the tools, services, registries, interfaces, and infrastructures required to support the responsibilities and workflows established in the preceding steps.

Relevant components may include persistent identifier services, electronic laboratory and field notebooks, sample and instrument management systems, institutional information systems, metadata editors, vocabulary services, validation tools, repositories, interfaces, and aggregation services.

These components should not operate as isolated applications. They must form an interoperable environment in which data and metadata can be exchanged, validated, enriched, and reused across organizational and disciplinary boundaries. The technical ecosystem thereby enables information to be captured at its point of origin, maintained by responsible actors or authoritative systems, and transferred reliably to repositories and other downstream services.

The data ecosystem provides the operational foundation through which assigned responsibilities and agreed workflows are translated into sustainable practice.

Operationalizing FAIR

The Road-to-FAIR Strategy complements the FAIR Guiding Principles by translating their high-level objectives into a coordinated implementation framework. FAIR defines the properties that digital research objects should exhibit, whereas the Road-to-FAIR Strategy addresses the standards, responsibilities, processes, and technical services required to realize those properties.

A central component of the strategy is the development of shared community agreements. Persistent identifiers, metadata schemas, semantic vocabularies, exchange protocols, provenance models, licences, and quality procedures must be selected and applied consistently across infrastructures. FAIR implementation is therefore treated as a collaborative and incremental process rather than as a binary state. Institutions and communities can identify comparatively mature areas, prioritize remaining gaps, and progressively improve interoperability, reusability, and machine actionability.

The strategy also recognizes that the provision of rich metadata alone is insufficient. Metadata must employ harmonized semantics, explicit relationships, and interoperable structures if information is to be reliably interpreted and combined across disciplinary and organizational boundaries. At the same time, FAIRness is distinguished from unrestricted openness: data may remain subject to legal, ethical, contractual, or institutional access controls, provided that their metadata, access conditions, and access procedures are clearly documented.

The Road-to-FAIR Strategy thus establishes the socio-technical conditions required to operationalize FAIR. By connecting strategic objectives with community agreements, institutional responsibilities, coordinated workflows, and an interoperable data ecosystem, it provides a structured pathway from a general commitment to FAIR towards a progressively harmonized and reusable Helmholtz research data space.

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_road-to-fair-strategy:start&rev=1786013683

Last update: **2026/08/06 10:54**