

## FAIR Ecosystem Components: Vision

The primary focus of work package four<sup>1</sup> in FAIRsFAIR<sup>2</sup> is (trusted) repositories that enable the curation of (FAIR<sup>3</sup>) objects. But to be integrated into an operational European Open Science Cloud (EOSC<sup>4</sup>) a wider vision of FAIR ecosystem dependencies and interconnections is required.

Data users and stewards of all kinds must be empowered to find, store and access data and metadata designed for interoperability and reuse. This draft presents a vision for the FAIR ecosystem components required to ensure FAIRness across the full data lifecycle.

Ensuring (meta)data FAIRness requires that digital objects are cared for by data stewards, now and in the future. The FAIRness of objects depends on their curation context including FAIR-enabling Trustworthy Digital Repositories (TDR) where (meta)data management can help assure that digital assets will remain FAIR over time (TDR + FAIR + time = preservation). This long-term perspective requires managed responses to changes in data infrastructure (people, processes and technology) and changes in the needs, composition and knowledge base of user communities are addressed.

Researchers, data stewards (including repository maintainers) and other ecosystem actors like funders need the results of automated and manual evaluation of objects' FAIRness. Objects may be evaluated inside or outside of the context of a TDR.

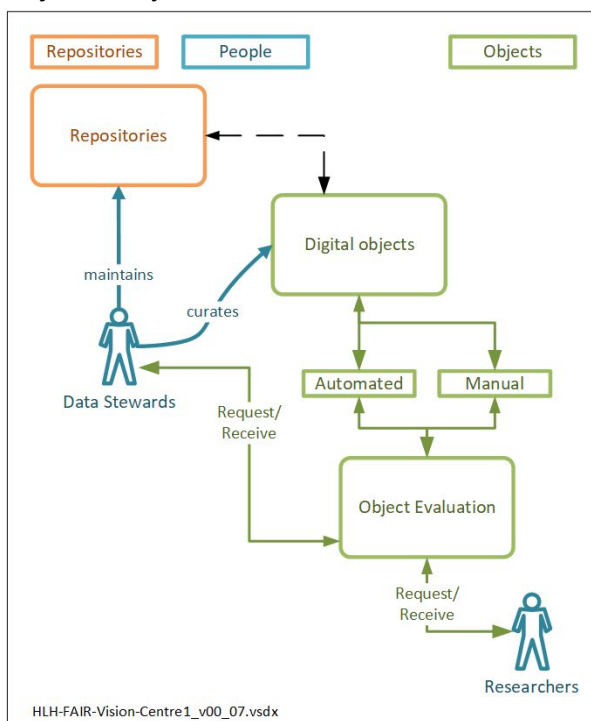


Diagram 1: **Repository Data Stewards and Researchers: FAIR Object evaluation.**

<sup>1</sup> <https://www.fairsfair.eu/fair-certification>

<sup>2</sup> <https://www.fairsfair.eu/>

<sup>3</sup> <https://www.force11.org/group/fairgroup/fairprinciples>

<sup>4</sup> <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

These actors, repositories, objects and processes do not exist in isolation. Metadata about digital objects is also stored in object registries (e.g. DataCite<sup>5</sup>) alongside their persistent identifiers. The trustworthiness of repositories is evaluated by comparing evidence to standard requirements using an assessment processes (e.g. CoreTrustSeal<sup>6</sup>). Evidence is also dependant on the software and the federations of services that support FAIR-enabling repositories of FAIR digital objects. Discovery and selection of repositories is supported by repository registries (e.g. re3data<sup>7</sup>) that provide relevant metadata including certification status.

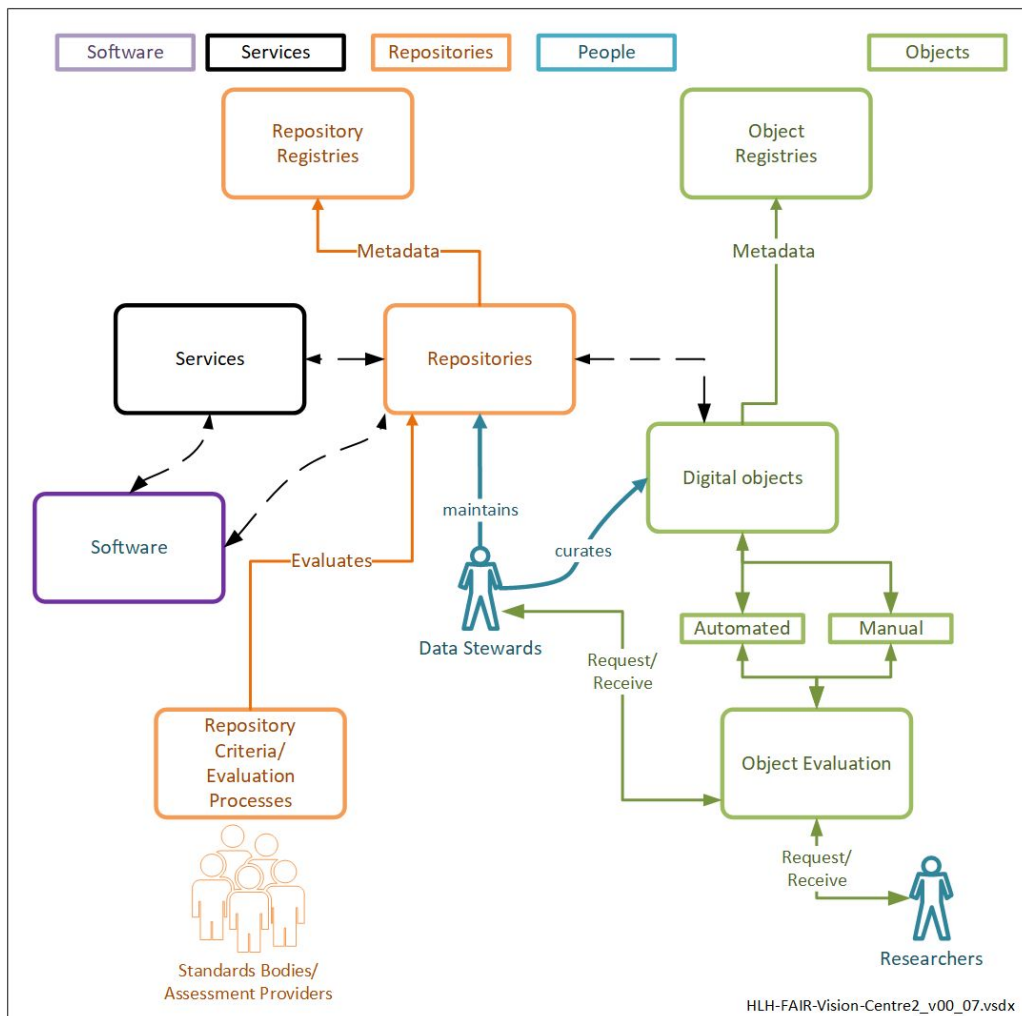


Diagram 2: **Repositories, Software, Services, Assessment & Registries.**

But a mature FAIR ecosystem depends on a wider range of interdependent and coordinated components. The Turning FAIR into Reality report<sup>8</sup> envisions the need for standards and assessments across a range of ecosystem components that are persistently identified and indexed in registries. The development of FAIR skills is also a clear goal.

<sup>5</sup> <https://datacite.org/>

<sup>6</sup> <https://www.coretrustseal.org/>

<sup>7</sup> <http://re3data.org/>

<sup>8</sup> <https://doi.org/10.2777/54599> (figure 6)

Any evaluation of FAIRness, from self-assessment to formal audit and certification, depends on FAIR standards and processes. Those processes seek to compare characteristics of the entities (repositories, objects etc) being assessed against standard indicators of compliance with the FAIR principles. Some aspects of the FAIR principles (E.g. richness, accuracy and relevance) can only be clarified and then evaluated based on the local context of the (meta)data, including domain and disciplinary expectations.

Registries can be conceptualised as curated repositories of metadata. The alignment of digital object repositories and metadata registries is to be defined. Repository evaluation focuses on the stewardship of research data objects: data and associated metadata and documentation created by, collected by, or of interest to researchers. A functioning ecosystem depends on a wider range of curated objects including publications, standards, policies, procedures and data management plans (DMP<sup>9</sup>). To be trustworthy over time these artefacts must also be made FAIR through good information governance.

In the diagram below a high-level vision of these ecosystem components and their interactions and dependencies is presented. Dotted lines with double-headed arrows represent a wide range of possible operational interactions.

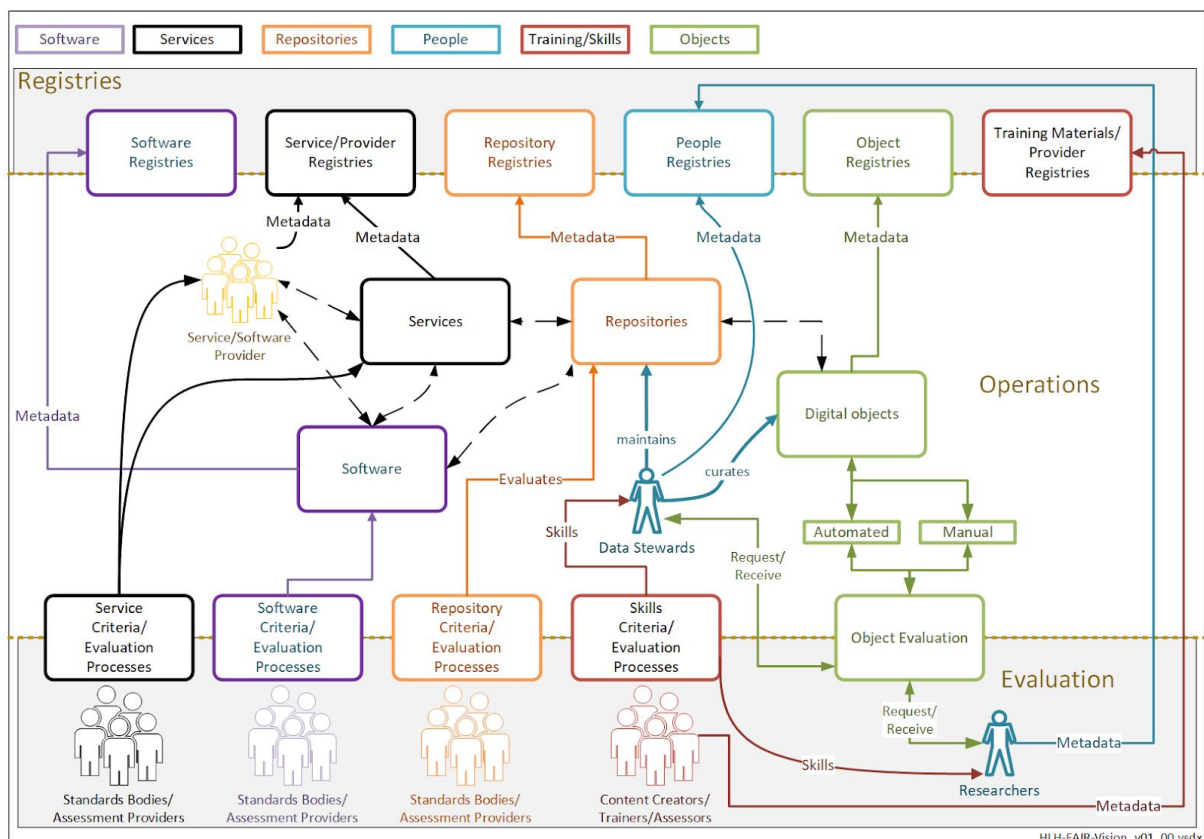


Diagram 3: FAIR Vision: Ecosystem components.

<sup>9</sup> <http://www.dcc.ac.uk/dmponline>

- Standards-compliant operations are supported by evaluation processes and registries.
- Operations, registries and evaluation systems must be standardised, assessable and sustainable.
- FAIR objects consist of uniquely, persistently identified FAIR data, metadata and documentation.
- Objects may be research data or other objects necessary to a FAIR ecosystem including policies or publications
- FAIR-enabling repositories are trustworthy digital repositories (TDR) that can provide the environment necessary to ensure the FAIRness of objects over time.
- Registries provide common-reference points for persistently identified objects and entities (people, organisations, grants, instruments) which are curated with relevant metadata. Including:
  - Repository registries (e.g. Re3data) of repository metadata, including certification status.
  - (Meta)data standards registries and terminology registries support FAIR (meta)data
  - People registries (e.g. ORCID<sup>10</sup>) support the unique identification of actors and roles.
  - Registries of FAIR training and qualifications support FAIR skills development.
  - Registries of FAIR services provide insight into FAIR across federations of partners and throughout the data lifecycle. The scope of and approach to FAIR services have not yet been fully defined.
  - Research data registries (e.g. DataCite) provide metadata which include persistent identifier resolution to objects in repositories.
  - Registries of policies, including metadata about their machine-actionability support the governance of a FAIR ecosystem by funders, service/software providers, research producing organisations and repositories.

A FAIR operational lifecycle, supported by FAIR registries, with FAIR-trained actors using FAIR-services and software would streamline the certification of FAIR-enabled Trustworthy Digital Repositories and their FAIR-data collections.

This vision for ecosystem components which support findable, accessible, interoperable, and reusable metadata and data has direct relevance to the European Open Science Cloud (EOSC). It is also more widely applicable to the global trend toward defragmentation and integration of data infrastructure into more efficient, interoperable infrastructure for data collection, creation, analysis, use and long-term preservation.

Although it is impractical to expect such a fully integrated data and metadata ecosystem at this stage of development an agreed vision will support design and realisation.

---

<sup>10</sup> <https://orcid.org/>