FAIR Metrics and Digital Objects Task Force Terms of Reference

1   TF Background and Scope

The FAIR Metrics and Digital Objects Task Force focuses on the issues related to FAIR (Findable, Accessible, Interoperable, Reusable)[1] data (hereafter interchangeable with "digital objects"[2]) in the context of the EOSC federation, exploring the limitations of the existing digital ecosystems, particularly the EOSC and EOSC Data Spaces, concerning exposure of FAIR digital objects for discovery and appropriate, for-purpose reuse. It will examine how significant repositories are helping data providers to expose provider-authored metadata, what technologies are available to assist and encourage providers to undertake this authoring task, and the degree to which their success can be objectively measured.

The Strategic Research and Innovation Agenda (SRIA) of EOSC strongly focuses on the fulfillment of the FAIR principles with the Strategic Objectives SO4 (increased FAIR by design) and SO5 (EOSC Interoperability Framework of FAIR DOs), as well as the Operational Objectives OO5 (specifications for FAIR DOs[3]) and OO6 (FAIR metrics). The TF is also aligned with several priorities of the Multi-Annual Roadmap, such as the liaison with HEU projects and the domain-specific environments (in objective 1) and fostering FAIR-enabled trusted repositories (in objective 2).

1.1   Main Objective

The main principles and delivery objectives of the TF are:

- Identify the limitations of the current FAIR assessment concerning the for-purpose reuse of digital objects.
- Regarding resources on the EOSC, guide tooling to achieve alignment with international initiatives such as GREI, Signposting, RO-Crate, ISA, and resources such as OpenAIRE,

---

[1] Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3 https://doi.org/10.1038/sdata.2016.18

[2] As in the FAIR Principles paper, we intend "data" to mean any kind of digital resource (in particular scholarly objects) made available through network access, including among others, dataset files, interactive databases, software, and workflows. Throughout this document, we will use "digital object" to ensure that this broad scope is recognized, and will use upper-case Digital Object only when referring to the project in footnote 3.

[3] "FAIR Digital Object" (https://doi.org/10.5281/zenodo.7824713) is a set of specifications from the FDO Forum for implementing digital objects as FAIR resources; however, in this document we consider more broadly the FAIR aspects of any digital object.

1
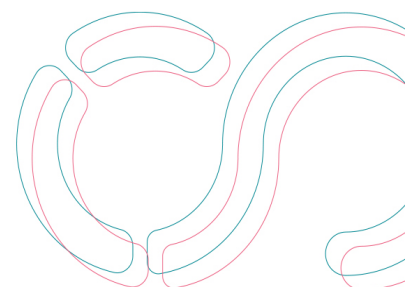
FAIRsharing, etc., to facilitate the definition of common metadata schemas and their interoperability.

- Identify common issues on data privacy, considering data usage, data access, and data licensing and specification for machine-actionable data usage policies (e.g., ODRL, DUO)
- Analyse current support and potential for maintaining the provenance record for digital objects, especially in the context of the EOSC federated environment.
- Identify synergies with the Data Spaces initiative and the FAIR Assessment and Alignment Opportunity Area Expert Groups.
- Examine the degree to which data quality measurements can be exposed in FAIR metadata following existing standards and if new standards need to be created.
- Explore the intersection and gaps between FAIR (Findable Accessible Interoperable Reusable) and the requirements for selecting a digital object for machine actionability and automatic processing (for example, bootstrapping AI applications).
- Align with and participate in the nascent FAIR Assessment Governance task established by the EOSC INFRA OSTrails project and the broader global stakeholder community.

1.2 Key Focus Areas

The set of measurable Key Focus Areas (KFAs) contributing to the main objective is listed next:

- KFA1: Survey and analyse the benefits, limitations, and adoption of the current metadata schema standards (DCAT, schema.org, etc.) and their interoperability, focusing on the reusability of data and provenance.
- KFA2: Identify issues and limitations when handling sensitive data in repositories, focusing on managing data access limitations.
- KFA3: Proposal of FAIR Metrics for evaluating the reusability of data, considering the analysis of the previous Key Focus Areas.
- KFA4: Identify synergies and complementarities between the Data Spaces, especially considering the Simpl Open middleware and its implementation in each space.

1.3 Expected Membership

The TF is expected to attract the interest of experts from research communities. Data repositories and Research Infrastructures will jointly gather the necessary skills required for achieving the TF objectives.
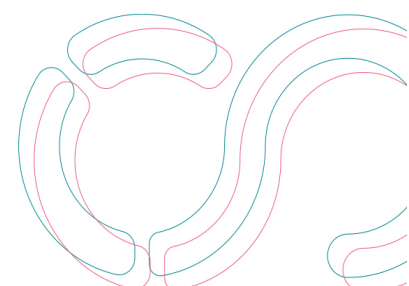
1.4   List of Deliverables

The Task Force anticipates two distinct deliverables that are associated with these KFAs:

<u>KFA1,2,3</u>

- **Report on Repository Support for Data-level Interoperability/Reusability**
  - Schema Interoperability Report (with a focus on data reusability and provenance in federated environments)
  - Data Privacy and Usage Policy Analysis (with a focus on existing machine-actionable policies)
  - Recommendations on metrics and a gap analysis

<u>KFA4</u>

- **Joint Workshop** (TF and other stakeholders) to explore data spaces and Simpl w.r.t. their ability to expose critical FAIR metadata and, more importantly, FAIR data.
  - **Output: Position paper, including**
    - Review (literature and technology) on middleware solutions for EOSC and how they relate to other international initiatives (e.g., GREI)
    - Community perspectives on Governance of FAIR assessment as we become increasingly able to assess the digital objects contained within a repository
    - Report on all facets that should be available in a metadata record (data quality, citation, context, license, and usage constraints), including an Intersection Analysis over domains/digital objects, as well as established repositories vs Data Spaces