Good FAIR Practices

Erik Schultes, Barbara Magagna, Andrea Tarallo

4 December 2024

IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-Mission 4 “Education and Research” - Component 2: “From research to business” - Investment 3.1: “Fund for the realisation of an integrated system of research and innovation infrastructures”

Finanziato dall'Unione europea NextGenerationEU
Ministero dell’Università e della Ricerca
Italiadomani

FAIR Data Policy vs FAIR Practice

13:40-14:00
(Barbara)

FAIR Principles as FAIR Data Policies

- Ensure data are described with standardized metadata (F2)
- Store data in repositories with robust search capabilities (F4)
- Provide APIs and use standard data exchange protocols to support automated data sharing and integration (A1.1)
- Provide open access to datasets where legally and ethically permissible (A1.2)
- Implement secure mechanisms for controlled access where necessary (A1.2)
- Store data using open, well-documented formats that facilitate integration with other datasets (I1)
- Use common vocabularies and ontologies to describe data, enabling semantic interoperability (I2)
- Provide comprehensive documentation to describe the context, methodology and provenance (R1.2)

FAIR Data Policy

- A set of **guidelines** that **organizations, research institutions, or funding bodies establish** to ensure that data is **FAIR**. It includes specific requirements and best practices for data management, sharing, and preservation, and promotes data stewardship in line with the FAIR principles

- And some FAIR sub-principles address:
  - Metadata preservation policy (I2)
  - Data usage license (R1.1)

- Specific policies e.g. for F1: EOSC compliant PID Policy

- Infrastructures related: PARC FAIR Data Policy (see ITINERIS platform)

- Or criteria, like https://www.gofair.foundation/criteria

EOSC Compliant PID Policy

eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

ITINERIS

4 Guidelines for formulating PID policies for PID Managers

4.1 Information sources of the guidelines

4.2 16 Guidelines applicable to PID Managers

Guideline 1. “Select a PID Stack that is globally unique and persistently resolvable”
Guideline 2. “Manage Persistence”
Guideline 3. “Manage Versions”
Guideline 4. “Involve Stakeholders”
Guideline 5. “Conform to a PID Stack checklist”
Guideline 6. “Select an appropriate scale”
Guideline 7. “Select appropriate identifier schema and structure”
Guideline 8. “Consider resolution options”
Guideline 9. “Maintain resolution integrity”
Guideline 10. “Manage Metadata”
Guideline 11. “Consider implementation of Machine-Actionable Extensions”
Guideline 12. “Monitor Resolution Integrity”
Guideline 13. “Take sensitive metadata into consideration”
Guideline 14. “Consider periodic resolvability sampling”
Guideline 15. “Develop and implement sustainability and continuity mechanisms”
Guideline 16. “Adopt a level for maturity and availability of Services”

Work Package	**WP 3, Persistent Identifiers**
Lead Author (Org)	**René van Horik (DANS-KNAW)** **Wim Hugo (DANS-KNAW)**
Contributing Author(s) (Org)	**Joy Davidson (DCC)** **Josefine Nordling (CSC)** **Liisa Marjamaa-Mankinen (CSC)** **Lassi Lager (CSC)** **Gabriela Mejias (DataCite)** **Elizabeth Newbold (STFC)**
Due Date	**2024-05-31**
Date	**2024-05-30**
Version	**V1.0 - DRAFT NOT YET APPROVED BY THE EUROPEAN COMMISSION**

Guideline 2. “Manage Persistence”

Guaranteeing persistence requires effort - usually from the registry (Authority) and from the Manager. Managers must develop policies and procedures to guarantee maintenance of the correct link between the identifier and the resolution target, and make sure the responsibilities are well defined in their agreements or contracts with the PID owner. The two main manifestations of lack of persistence are “link rot”(the weblink does not resolve to a resource) and “content drift” (the original link does not refer to the resource it was initially connected to).

PARC FAIR Data Policy

This PFDP is not intended as a practical operational document to help PARC in data management, which is within the DMP and project specific DMPs, but instead provides guiding principles on a.o.: aim and scope of PFDP, applicable laws and frameworks; data governance and management, including security, quality; transparency and openness; FAIR (Findable, Accessible, Interoperable, Reusable) including the foreseen ambition of PARC to which extent FAIR can be accomplished throughout PARC; intellectual property and sensitive personal data. This will be explained in more detail in the next Chapters.

Based on the PFDP, WP7 will set up specific methods and tools and guidance to enable partners to FAIRify their data and facilitate data reuse within the research and innovation (R&I) WPs of PARC, by regulatory agencies and for research outside the scope of PARC by the broader scientific community. In detail, this operationalisation of the PFDP has three important pillars:

5. Data governance
6. Data management
7. Data security
8. Repositories
9. Transparency and openness
10. FAIR
10.1. Findable
10.2. Accessible
10.3. Interoperable
10.4. Reusable
10.5. PARC's Ambition in relation to FAIR data

Lead Beneficiary/ Responsible AE	Lead Beneficiary: RIVM / Responsible AE: TNO
Contributing Participants	EAA (AT), UG-PL (PL), VITO (BE), UBA (DE), UU-IRAS (NL), ANSES (FR), MU (CZ), ISS (IT), EV-ILVO (BE), UZIS (CZ), AUTH (EL), BRGM (FR), JSI (SI), UL-LACDR (NL), KWR (NL), KI (SE) , UoB (UK)
Responsible author(s)	Fred van de Brug / TNO / fred.vandebrug@tno.nl Rob Stierum / TNO / rob.stierum@tno.nl
Co-authors	Tessa Pronk / KWR / Tessa.Pronk@kwrwater.nl Penny Nymark / KI / penny.nymark@ki.se Peter Von der Ohe / UBA / Peter.VonderOhe@uba.de Iseult Lynch / UoB / Iseult.Lynch@bham.ac.uk
Reviewers	Reviewers up to "PARC_D7.2_FAIR_Data_Policy_**V1.1**_20230623.pdf" Jan Theunis / VITO (BE) Sylvie Remy / VITO (BE) Erik Schultes / GFF (NL) Jane Richardson (EFSA)

GFF Criteria

- **Interpretation of the FAIR Principles** publicly accessible
- The **hourglass model** to promote freedom to operate with a minimal standard a the center
- **Openness**: As open as possible, as restricted as necessary
- **Distrution**: Distributed architecture should be the default, and should be centralizd only as necessary

FAIR Practice

- GFF definition: A method description detailing the use of a specific resource or a combination of resources as applied by a community to achieve FAIR processing of information.

Examples:

- GO FAIR Foundation uses Open Science Framework for publishing material for events
- GO FAIR Foundation uses nanopublications for documenting events

- A broader definition might be: any action, methodology, process that promotes FAIR

- Might also be used as synonym to FAIR Data Policy

Repositories for FAIR Practices

- FAIRsharing (https://fairsharing.org/)
- FAIR cookbook (https://faircookbook.elixir-europe.org/content/home.html)
- FAIR Implementation catalogue (https://catalogue.fair-impact.eu/)
- FAIR Connect (https://fairconnect.pro/search-fair-nanopublications/)

FAIR Connect

- A FAIR repository of resources that support the FAIRification of data
- Resource typed according to a typology
- Resources are formalized via nanopublications as representation language
- A growing list of resources as contributed by participating communities
- The platform will be presented and used in more detail in the FAIR Implementation Profile course

Breakout session

- 3 groups (FAIRsharing, FAIR cookbook, FAIR implementation catalogue)

Tasks, to be documented in the notes:

- Identify core characteristics of these repositories
  - What are they about
  - How are they described
  - How are they represented
  - Do they have a persistent identifier
- Identify 2 FAIR practices of your choice you did not know so far
- Present your findings to the whole group

FAIR Well!

THANKS!

Erik Schultes, Barbara Magagna, Andrea Tarallo
27 November 2024

IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-Mission 4 "Education and Research" - Component 2: "From research to business" - Investment 3.1: "Fund for the realisation of an integrated system of research and innovation infrastructures"

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani