Concepts for FAIR implementation: FAIR Digital Objects and technical components of the FAIR ecosystem - Pillar 1

Session date: 11th of May 2020
Chair: Gerard Coen (FAIRsFAIR)
Rapporteur: Anusuriya Devaraju (FAIRsFAIR)
Zoom Room Management & Tech support: Rita Meneses (FAIRsFAIR)

Information on the participants, the projects and working groups that they represent, and the spreadsheet used during the workshop can be found in the workshop report: https://doi.org/10.5281/zenodo.3953979

All recommendations and the action plan can be found on pp. 59-75 in *Turning FAIR into Reality*: https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf.

This session is about recommendations 1, 2, 3 (priority) and 16, 17.

Table of contents

Rec. 1: define FAIR for implementation

To make FAIR data a reality it is necessary to incorporate and emphasise concepts that are implicit in the FAIR principles, namely: data selection, long-term stewardship, assessability, legal interoperability and the timeliness of sharing.

1.1 In place

*What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.*

1.2 Planned

*What are the projects represented **developing or planning** to do? Again, this should build on the information in the spreadsheet: information about a planned deliverable, i.e. title, due date, short description*

FAIRsFAIR:
- D3.3 Policy enhancement recommendations https://zenodo.org/record/3686901
- D3.4 Recommendations for FAIR practice, May 2020
- D3.5 Description of transition support for repositories, Aug 2020
- D.2.1 D2.4 & D2.10 reports on FAIR requirements for persistence and interoperability, https://doi.org/10.5281/zenodo.3557381
- D2.3 Set of FAIR data repositories features, https://doi.org/10.5281/zenodo.3631527
- D2.7 Framework for assessing FAIR services https://doi.org/10.5281/zenodo.3688762
- M2.15 (Milestone) Assessment report on FAIRness of research software, 31st July 2020.

EOSC FAIR WG:
- The FAIR Practice Task Force examines FAIR practices in a variety of disciplinary fields. A draft report will be available for consultation around July 2020. There is also a plan to host a webinar for consultation.
- The Interoperability Task Force plans to examine legal interoperability.
- The Metrics & Certification TF are working on a recommendation for FAIR Metrics. An interim recommendation is available here: https://www.eoscsecretariat.eu/eosc-liaison-platform/post/interim-recommendations-fair-metrics-and-service-certification-apply It takes into account the output of the RDA FAIR Data Maturity

Model WG and of FAIRsFAIR WP2 and 4. There remain open questions for defining FAIR for other components of the FAIR ecosystem than data, e.g. software.

- Note the on-going creation of an RDA/Force11/ReSA WG FAIR 4 Research Software (FAIR for Research Software WG Case Statement submitted on 11/5 and open for review until 11/6, https://www.rd-alliance.org/sites/default/files/case_statement/2020_RDA_FAIR4RS_WorkingGroup_CaseStatement.pdf)

EOSC Nordic:

- Implementation of FAIR assessment based on the FAIR Evaluator (Wilkinson et al. 2019) and GO-FAIR. D4.1 and D4.3 will detail this including FAIR recommendations and the use of hackathons and M4M events to support communities to implement. Early (1st epoch) results presented 22 April at EOSC-Nordic workshop (recording available here: https://www.eosc-nordic.eu/first-online-event-for-eosc-nordic-project-is-a-big-success/). The evaluation is due to D4.1 (1 Sep 2020). M4Ms + hackathons will take place Aug 2020 - Feb 2022.

EOSC Pillar:

- D5.1 FAIR Research Data Management Tool set - first release 30 June 2020
- D5.2 FAIR Research Data Management Workbench Operation Report
- D5.4 FAIR-oriented Research Data Management: Support, Training and Assessment Activity Report (1 of 2), 31 Dec 2020
- D5.4 FAIR-oriented research data management: Support, Training and Assessment Activity Report (2 of 2), 30 June 2022
- D5.6 FAIR Research Data Management Tool set update (Month 24)
- FAIR federated data space (work-in-progress) to examine FAIR compliance

NI4OS-Europe:

- Implementation of FAIR is being done according to "The FAIR Guiding Principles for scientific data management and stewardship" by Wilkinson et al.
- Information on the FAIR uptake strategy of NI4OS-Europe is provided in the deliverable D6.1 "EOSC service and FAIR uptake strategy" which can be found on http://doi.org/10.5281/zenodo.3736152.
- Further information regarding FAIR implementation can be found in the deliverable D4.2 "Data repository integration and ORDM/FAIR compliance guidelines", https://doi.org/10.5281/zenodo.3736149

ExPaNDS:

- D2.1/2.3 Draft/Final extended data policy framework for Photon and Neutron RIs (Aug 2020/Aug 2021)
- T5.1.1 Workshop to promote FAIR principles will take place Sept/Oct 2020
  - Revision of RI data policies (with PaNOSC) to take FAIR principles into account; workshop on FAIR principles aimed at facility scientists and staff planned for Sept/Oct2020; liaising also with FsF Competence Centre WP6
  - Joy (FAIRsFAIR): we'd be very keen to ExPaNDS take a look at the FAIRsFAIR draft policy enhancement recommendations that came out in late February and see if these are relevant to your work and if there is potential for us to collaborate to support implementation.

ENVRI-FAIR:

- Report on overall ENVRI-FAIR project included:
  - Petzold, Andreas; Asmi, Ari; Vermeulen, Alex; Pappalardo, Gelsomina; Bailo, Daniele; Schaap, Dick; M. Glaves, Helen; Bundke, Ulrich; Zhao, Zhiming: ENVRI-FAIR – Interoperable environmental FAIR data and services for society, innovation and research, IEEE International Conference on eScience 2019 (eScience2019), San Diego, 24-27, Oct 2019, http://doi.org/10.1109/eScience.2019.00038
- Survey on FAIR maturity included in Rec.3

PaNOSC:

- Photon and Neutron Data policy update for FAIR D2.1 (May 2020) Common API for metadata catalogues
- D3.1 API definition
- D3.5 NeXus Metadata mapping schema and proposed new definitions.

SSHOC:

- The "The FAIR Guiding Principles for scientific data management and stewardship" by Wilkinson et al is followed in terms of definitions and implementation guidelines. Keeping the internal diversity in SSH, specific (data) communities addressing/translating what FAIRification would mean for them. Example: WP9 Data communities (esp. Ethnic and Migration Studies).

View from the Champions:

FAIR Champion 1: CSIC Institutional Open Access mandate entered into force on April 1, 2019, https://digital.csic.es/handle/10261/179077. The mandate covers both peer reviewed publications and research data produced by the CSIC research community. The mandate explicitly mentions that research data associated to publications must be FAIR and be deposited and made open access (unless exceptions may apply) in institutional repository DIGITAL.CSIC

Rec. 2: implement a model for FAIR Digital Objects

Implementing FAIR requires a model for FAIR Digital Objects. These, by definition, have a PID linked to different types of essential metadata including provenance and licencing. The use of community standards and sharing of rich documentation is fundamental for interoperability and reuse of all objects.

2.1 In place

*What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.*

4

2.2 Planned

*What are the projects represented **developing or planning** to do? Again, this should build on the information in the spreadsheet: information about a planned deliverable, i.e. title, due date, short description*

EOSC FAIR WG:
- PID TF: (Action 2.1) Second draft of PID Policy document (developed with the Architecture WG) available for consultation. Second draft in zenodo (but you have to use the link below the document window, https://doi.org/10.5281/zenodo.3574202
- Suggestion : Action 2.3 on automated compliance check: one has to consider the risks. More work needed on DO. RDA work being defined in the domain.
- More work needed on defining digital objects, for example how to apply semantics of data.

EOSC Nordic:
- Indirectly working on this recommendation via the FAIR Evaluator and FAIR recommendations (D4.1, due Aug2020). Follow the FAIR Digital Object framework (PIDs, explicit identifier for metadata and data) - follows the implementation by the evaluator tool by Mark Wilkinson.
- Training is ongoing since 2019 (but independently of the EOSC-Nordic project) and by the end of 2020 we will have trained more than a 100 Data Stewards via a 5-day course by GO-FAIR.

EOSC Pillar:
- WP5.1 and WP5.2 will deliver a pilot where data repositories from some of the use cases of the project are federated. It will use existing tools (e.g. Cordra) to prove the concept which will include the FAIR Digital Object technology and which will deliver recommendations on integration of existing repositories. The work will be discussed and shared with other call 5b projects - in planning, no concrete outcomes yet.

NI4OS-Europe:
- Task T4.3 "Open and FAIR data in practice: tools for RDM following the data lifecycle" is focussing on this recommendation. This will lead to the deliverable D4.4 "Delivery of data management and certification tools" (type Demo) (July 2020) - IPR tool for digital objects.
- PID policy is being worked on (WP5); in addition, we monitor FAIR Data Maturity Model and FAIRsFAIR work on FAIR-aligned services in order to prepare with developing FAIR-aligned tools and services in WP4.
- Further information regarding FAIR implementation can be found in the deliverable D4.2 "Data repository integration and ORDM/FAIR compliance guidelines", https://doi.org/10.5281/zenodo.3736149

ExPaNDS:
- D2.2 Draft Recommendations for FAIR Photon and Neutron Data Management (Aug 2020)

- D2.7 Final Recommendations for FAIR Photon and Neutron Data Management (Feb 2022)
- D2.5 Advanced infrastructure for PIDs in Photon and Neutron RIs (Aug 2021)
- T5.1.1 Workshop to promote FAIR principles (to take place Sept/Oct 2020)
- Two main activities:
  - Metadata model for representing data at different stages in the experimental life cycle (from planning to publication, including descriptive, structural and administrative metadata); Final recommendation (D2.7) will be out in Feb 2022.
  - Working together with the FREYA project to look at what new PIDs (e.g. instrument, sample, software) would be useful to develop for Photon and Neutron RIs
- Joy: FAIRsFAIR WP3 is also looking at PIDs for instruments so we'd be keen to discuss this with ExPaNDS (Thanks, Joy - Useful to know! Abigail)

EOSC Synergy:
- D3.3 Intermediate report on technical framework for EOSC FAIR data principles implementation. Describes the evaluation of the recommendations for assessing data FAIRness and data repository features coming from FAIRsFAIR and provides details about architecture, requirements and a roadmap for implementation. (Due August 2020).
- D3.5 Final report on technical framework for EOSC FAIR data principles implementation. Implementation details for a technical framework to validate and monitor data FAIRness. Any change or addition to the information gathered in D3.3 will be reported. (Due Nov 2021)
- Observe existing demo across projects (1st report (summer 2020), second report Q4 2021)

ENVRI-FAIR:
- Completed training on ontology & interoperability, developing catalogue of training materials (https://training.envri.eu)

ESCAPE:
- Has training for data producers (the large RIs) and users on how to use the data in particular in the Virtual Observatory (Disciplinary Interoperability Framework)
  - See NI4OS-Europe Training Page page for trainings

PaNOSC:
- Common catalogue API implementation report D3.4
- Policy implementation guidelines defined D2.3
- Deliverable 5.1: Prototype simulation data formats as openPMD domain specific extensions including example datasets.

SSHOC:
- In the context of the SSHOC Open Marketplace, D7.1. (System Specification - SSH Open Marketplace) presents an implementation plan for FAIR digital objects, Oct 2019, https://doi.org/10.5281/zenodo.3547648

View from the Champions:

FAIR Champion 1: DIGITAL.CSIC is the institutional repository for the wide array of outputs by CSIC researchers (ranging from peer reviewed publications to research data and patents). The repository assigns handle PIDs to all items and in addition it mints DOIs for selected outputs: research data, software, and preprints. DIGITAL.CSIC item description handbook emphasizes the importance of including PIDs in selected metadata elements (use licence, research funders, related works)

Rec. 3: develop components of a FAIR ecosystem

The realisation of FAIR data relies on, at minimum, the following essential components: policies, Data Management Plans, identifiers, standards and repositories. There need to be registries cataloguing each component of the ecosystem, and automated workflows between them.

3.1 In place

*What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.*

3.2 Planned

*What are the projects represented **developing or planning** to do? Again, this should build on the information in the spreadsheet: information about a planned deliverable, i.e. title, due date, short description*

FAIRsFAIR:
- D3.3 Policy enhancement recommendations https://zenodo.org/record/3686901
- D3.4 Recommendations for FAIR practice (May 2020)
- D3.5 Description of transition support for repositories (Aug 2020)
- D2.3 Set of FAIR data repositories features, https://doi.org/10.5281/zenodo.3631528
- D2.6 1st reference implementation of the data repositories features, Feb 2021
- D2.9 Framework for assessing FAIR services, March 2022
- Draft recommendation for machine readable policies planned. Early discussions with FAIRsharing and RDA groups but keen to work with others who are working on this.

EOSC FAIR WG:
- Interoperability FT: Recommendations on the EOSC Interoperability framework prepared with the Architecture WG. Includes discussion of technical, semantic, organisational and legal interoperability. Draft for consultation should be published soon.
- PID TF: PID Policy prepared with the Architecture WG. Second draft available for comments.
- Metrics & Certification TF: Examines certification of FAIR enabling services.

- Likely more work to do on registries (Action 3.1). The need for testbeds to continuously evolve and evaluate is a good point (Action 3.4).

EOSC Nordic:

- Supported through the testing and development of the FAIR Maturity Evaluator on a 100+ data repository sample, highly focussed on machine actionability. During the project we will support FAIRification activities aimed towards service providers (e.g. recommendations to repositories on metadata, hackathons + M4M). Results from the first epoch evaluations for FAIR Maturity will be published in D4.1 (Aug 2020) and D4.3 (Feb 2022).
- Using the Maturity evaluator tool to evaluate repositories (put emphasis on machine actionability)
    - GO-FAIR one of the partners of the project
    - Offer trainings for data stewards through GO-FAIR

NI4OS-Europe:

- D4.1 "Incentives for supporting ORDM and FAIR" (https://doi.org/10.5281/zenodo.3736147). This deliverable describes the incentives and rewards that can be employed to improve the uptake of ORDM and FAIR in general as well as their integration in EOSC through country-tailored systems of rewards and incentives.
- D4.2 "Data repository integration and ORDM/FAIR compliance guidelines", (https://doi.org/10.5281/zenodo.3736149) This deliverable provides the results of the analysis of guidelines for open science.
- D4.6 Second set of data management and certification tools (type Demo) (October 2021); also work on FAIR monitoring/validation for repositories and potentially other services.
- D3.2 First report on pre-production environment (June 2020), D3.5 Final report on pre-production environment (April 2022).

ExPaNDS + PaNOSC:

- ExPANDS: D2.4/8 (Active) DMPs for Photon and Neutron RIs (Aug 2021/Aug 2022)
- PaNOSC: WP 3 Common API to deliver interoperable search between facility meta data catalogues. WP4 Development of data analysis service infrastructure. WP8 Training tools for EOSC & FAIR data for P&N
- ExPaNDS activities (1) develop a template for DMP that is populated automatically wherever possible (?link), (2) piloting these active DMPs at RIs, towards the end of the ExPaNDS project
    - *FAIRsFAIR WP3 developing some work around DMPs with different communities. Angus Whyte had some initial discussions with Brian Matthews for PaNOSC- is there a possibility to join up?*

EOSC Synergy:

D4.1 – Best Practices Elicitation including Data Management Plans (Currently in review)

ENVRI-FAIR:

- WP4 Common FAIR Policies, D4.1 Organisation of PWG – Membership, procedures for operation, WP5 – Common requirements and testbed for (meta)data services, community standards and cataloguing, D5.1 Requirement analysis, technology review and gap analysis of environmental research infrastructures, https://envri.eu/deliverables/
- WP4 on common FAIR policy, WP5 common requirements and testbed for metadata services (related deliverables of both in excel sheet)

SSHOC:
- D4.14 Policy API Tool, D4.14 Policy API Tool, https://sshopencloud.eu/d414-policy-api-tool
- D4.16 Specification of the new features of the Aïoli platform https://sshopencloud.eu/d416-specification-new-features-a%C3%AFoli-platform
- D7.1 System Specification - SSH Open Marketplace: https://sshopencloud.eu/d71-system-specification-ssh-open-marketplace

View from the Champions:

FAIR Champion 1: DIGITAL.CSIC research data policy and related services for CSIC community is at http://digital.csic.es/dc/politicas/politicaDatos.jsp (Spanish language only). The policy explicity indicates the types of research data accepted, main data management issues and recommendations and top data related services including DOI mining, support/review of DMPs, compliance with journals data sharing policies, compliance with FAIR Data Principles and aggregation into broader research data infrastructures

FAIR Champion 2: A number of building blocks are being developed in the context of the digital transformation initiatives for the Austrian academic sector including repositories, maDMPs; data citation support; controlled vocabularies for metadata search; APIs to CRIS systems or for funders; modules for data aggregation supporting k-anonymity, l-diversity, t-closeness; fingerprinting of data; reference set-up of a highly secure data infrastructure supporting data visiting based entirely on open-source components.

FAIR Champion 3: FAIRsharing both as a registry of standards and repositories, but also content provider for other tools: e.g. collaborating with DMPonline tool (more info in excel), and FAIRevaluation tools: https://fairsharing.org/communities#activities
FAIRsharing is already working to connect its content with DMPonline tool, and FAIRevaluation tools to use the standards in the FAIR assessment. As a registry, it is already used and recommended by various EOSC reports and projects, as well as EC reports and guidelines to researchers, and also by publishers.
Some links for adopters https://fairsharing.org/communities#adopters and activities https://fairsharing.org/communities#activities

Rec. 16: apply FAIR broadly

FAIR should be applied broadly to all objects (including metadata, identifiers, software and DMPs) that are essential to the practice of research, and should inform metrics relating directly to these objects.

16.1 In place

*What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.*

16.2 Planned

*What are the projects represented **developing or planning** to do? Again, this should build on the information in the spreadsheet: information about a planned deliverable, i.e. title, due date, short description*

FAIRsFAIR:
- D3.3 Policy enhancement recommendations *https://zenodo.org/record/3686901*

EOSC FAIR WG:
- To keep in mind the recommendations, although not all components are at the same stage for the definition of what FAIR means for them and of FAIRness criteria.
- The need to tailor the FAIR principles to the relevant context, in particular the research field, has to be kept in mind all the time (cf comments on inclusiveness below).
- The FAIR Practice TF is surveying the FAIR practices of different communities (Actions 16.2 and 16.4).

EOSC Nordic:
- We aim to introduce recommendations for FAIR services (in collaboration with WP3) and for data repositories (including publishing of code).

EOSC Pillar:
- D5.3 Training Plan (Dec 2019)

NI4OS-Europe:
- Training the trainers programme focuses on Open and FAIR RDM; promoting those principles at the national level by the trainers is expected. Webinar on FAIR data and principles was delivered in Feb 2020 to https://training.ni4os.eu/enrol/index.php?id=16 to support the wider application of FAIR.

ExPaNDS:
- D2.2/2.7 Draft/Final Recommendations for FAIR Photon and Neutron Data Management (Aug 2020/Feb2022)

EOSC Synergy:

- Working with FAIRsFAIR to see where the overlap is between Software-QA and Service-QA and what that means for FAIR. Not an expressed task but happening as part of collaboration activities.

PaNOSC:
- All WP aim to develop capability that allows phonon and neutron facilities to apply fair principles to their data chain. With a common approach across each domain. Data policy work and update to metadata schema WP2 & Wp3 are directly relevant to facilitating the application of FAIR principles.

SSHOC:
- D6.1 SSHOC Community Engagement Strategy, https://doi.org/10.5281/zenodo.3592243
- There will be SSHOC webinars with LIBER and Trust-IT on 19 May about FAIRIfication in the scholix framework, https://sshopencloud.eu/training
- Ensuring broad application of FAIR: DARIAH Research Data Management WG (https://www.dariah.eu/activities/working-groups/research-data-management/)
- Future plans: establishing stronger ties with the cultural heritage sector as they are important partners in humanities data workflows providing source materials for their work.

View from the Champions:

FAIR Champion 1: Being a multidisciplinary research performing institution, CSIC policy needs to address research data specificities of each of its 8 broad research areas. DIGITAL.CSIC produces supporting material (e.g. http://digital.csic.es/handle/10261/184580) addressing best practices and recommendations by research area, and its bulletin CSIC Abierto http://digital.csic.es/handle/10261/150210 includes interviews with CSIC researchers across several disciplines to showcase their motivations/challenges when sharing data and making them FAIR

Rec. 17: align and harmonise FAIR and Open data policy

Policies should be aligned and consolidated to ensure that publicly-funded research data are made FAIR and Open, except for legitimate restrictions. The maxim 'as Open as possible, as closed as necessary' should be applied proportionately with genuine best efforts to share.

17.1 In place

*What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.*

17.2 Planned

*What are the projects represented **developing or planning** to do? Again, this should build on the information in the spreadsheet: information about a planned deliverable, i.e. title, due date, short description*

FAIRsFAIR:
- D3.3 Policy enhancement recommendations *https://doi.org/10.5281/zenodo.3686900*

EOSC FAIR WG:
- It would be useful to work on guidance to provide to researchers to guide them to find the balance between open and legitimate restrictions (Action 17.8).

EOSC Nordic:
- Harmonising policies in the region and FAIR incentives is on the agenda of WP4 in the last year of the project (2021-22).

NI4OS-Europe:
- WP2 works on building national OSC initiatives aiming at stimulating discussions regarding national research settings and EOSC policies (incl FAIR) taking stock of OpenAIRE NOADs experience and resources that facilitate policy alignment.

ExPaNDS:
- D2.1 Draft extended data policy framework for Photon and Neutron RIs (Aug 2020)
- D2.3 Final data policy framework for Photon and Neutron RIs (Aug 2021)

EOSC Synergy:
- Looking at the SLAs in place for service providers related to data sharing, storage and transfer - making sure DMPs are available.
- Looking at the AAI policies in place and defining changes which might be recommended to improve/allow transnational data sharing by service providers.

ENVRI-FAIR:
- WP3 Strategy for alignment with national and international stakeholders, community development and innovation activities. "ENVRI-FAIR EOSC Position Paper", https://doi.org/10.5281/zenodo.3666805

PaNOSC:
- WP2 D2.1 PaNOSC data policy framework updated

View from the Champions:

FAIR Champion 1: DIGITAL.CSIC Template to prepare DMPs under the framework of H2020 projects https://digital.csic.es/handle/10261/207866 is an attempt to harmonize DMPs across the institution. On another front, there remains a lot to accomplish to raise more awareness about making data FAIR and handling patents at the same time. Global registries and best practices in this latter issue would be most welcome.

FAIR Champion 2: While opening up access to "shareable" data is increasingly accepted we notice that the biggest barrier to open data is the (massive, majority?) of data that is sensitive

(due to privacy reasons or commercial interests). In the wake of COVID19 TU WIEN has thus set up a secure data infrastructure based entirely on open-source components that allows data owners to provide controlled access for researchers to subsets of their data needed to address research questions while maintaining full control of the data and preventing data download to the fullest extent possible via a combination of technical and legal mechanisms. (The infrastructure is based upon experienced gained with operating such an infrastructure for medical data for several years, adapting it now for the needs of a broader group of stakeholders) This infrastructure is currently also under review for broader adoption, supporting the vision of data visiting instead of data sharing, and thus allowing to open a pathway to access for highly sensitive data, no matter whether this is due to privacy or commercial reasons. Activities are on their way to identify necessary changes to legal regulations to support such access provisioning for sectors with sensitive data that should be opened up for research purposes.

Whole-Pillar.Q1 What's missing in the recommendations and actions in this pillar?

What do projects do - related to implementing FAIR in the context of the EOSC - that is not covered by the original recommendations? Should it be included in an updated action plan and revised set of recommendations? Please focus on this pillar.

EOSC FAIR WG: regarding Rec.16 -> It is essential to take the context into account for all recommendations, and it is too bad that it is lost in a supporting recommendation. Linked to inclusiveness.