Basic framework on FAIRness of services - iteration 3

Authors:

Koers, Hylke (ORCID: https://orcid.org/0000-0001-6538-7590)

Herterich, Patricia (ORCID: https://orcid.org/0000-0002-4542-9906)

Hooft, Rob (ORCID: https://orcid.org/0000-0001-6825-9439)

Gruenpeter, Morane (ORCID: https://orcid.org/0000-0002-9777-5560)

Aalto, Tero (ORCID: https://orcid.org/0000-0002-1748-1176)

Ramezani, Sara (ORCID: https://orcid.org/0000-0001-9526-302X)

This is an updated version of the initial basic framework on FAIRness of services as presented in detail in:

Koers, Hylke, Herterich, Patricia, Hooft, Rob, Gruenpeter, Morane, & Aalto, Tero. (2020). M2.10 Report on basic framework on FAIRness of services (Version 1.0). Zenodo. https://doi.org/10.5281/zenodo.4292598

The initial milestone was based on literature review, interviews and feedback collected at an EOSC-hub week session. A second version (March 2021) incorporated feedback from the report launch webinar (Jan 2021) and workshop (Feb 2021) gathering more detailed feedback on the assessment framework. That update did not affect the aspects laid out, changes covered minor rephrasing to clarify some recommendations, additions and deletions of recommendations and re-ordering of recommendations to reflect priorities expressed by attendees of the workshop breakout sessions.

This current version (iteration 3, May 2021) introduces an update to the overall aspects as Trustworthiness has been split into two aspects for clarification: Transparency and Longevity. Additional changes cover rewording of recommendations as well as deletions of duplicate recommendations. We also clarified FAIR-enablement by highlighting that it covers both augmenting FAIRness of a digital object and facilitating a certain FAIR principle.

FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Technically-oriented aspects

FAIR enablement

Objective:

The service enables FAIR data by elevating the FAIRness of digital objects and/or supporting the FAIRification process. FAIR enablement is actively driven through the implementation of community-supported standards and interoperability frameworks.

Recommendations:

- In consultation with the target community (or communities), identify which metadata schemas and other standards (e.g. technical and semantic aspects of data encoding) should be adopted. Consider in particular domain-specific standards and practices. Strive to include accessibility conditions in metadata. Where applicable, generate and capture metadata automatically and be transparent about the concepts the service can provide an answer to.
- Engage with both the user community and other service providers to improve interoperability between services. Of particular attention here are authentication & authorization infrastructure (AAI), PIDs, and data and metadata encoding specifications. Seek alignment with existing or emerging data type registries and interoperability frameworks, e.g. the EOSC interoperability framework.
- Consider both human and machine access to the service, specifically with a view towards supporting automated pipelines for the FAIRification of digital objects.
- Use automated tests that show how the service increments FAIRness of digital objects in a verifiable, measurable, repeatable and scalable way. Root such tests in community-supported methodologies that measure the FAIRness of digital objects in an objective way.
- Perform a self-assessment on how the function(s) of the service *enable, respect* or *reduce* each of the FAIR principles for the data that it operates on.[1] Make the results of the self-assessment publicly available, together with an outlook on the desired state for the service (including a cost/benefit analysis).[2]
- Use persistent identifiers to refer to data and metadata.

- ***Enable:***
    - ***Augment****:* the service provides elements improving FAIRness of the digital object- for example automatically minting a DOI;
    - ***Facilitate***: the service actively helps to realize a particular FAIR principle — for example by allowing the user to add metadata or enabling discoverability;
- ***Respect***: the service does not actively enable a particular FAIR principle, but also does not interfere with it — it can be said to respect the “FAIR-in-FAIR-out” principle;
- ***Reduce***: the service actually makes data less FAIR — at least for a particular principle — for example by detaching metadata or a PID when it acts on a digital object;

[1] The case studies presented in Ref. (3) offer a suggested format for this self-assessment. Of course other formats are acceptable as well, however we do recommend to include all of the aspects listed in the case studies (i.e.: Summary; Users; Purpose; Adoption; Services; Target Digital Objects; Examples; FAIR enablement mapping).

[2] Note that a service does not need to address all aspects of FAIR, and integration with other FAIR-enabling services (e.g. PID minting) is often preferable over developing your own solutions.

Quality of service

Objective

The service is delivered in a reliable, secure, high-quality way, consistent with its specifications.

Recommendations:

- Codify the service’s availability and other non-functional aspects in a public Service Level Agreement (SLA) which is easy to understand by users from different communities.
- Deploy the service on appropriate and well-supported hardware or virtual (cloud) infrastructure. Define operational-level agreements (OLA) with 3rd-party infrastructure services that enable service delivery.
- Take reasonable technical and non-technical measures to prevent, detect, and respond to cyber or physical security threats; securing the service and protecting sensitive information resources (e.g. only using secure HTTP connections). Organize security audits and pen-tests at regular intervals, ideally at least every two years.
- Assess whether the service deals with sensitive data (e.g. patient records) and, if so, take additional measures in line with both applicable legislation and expectations from the user community.
- Implement service management processes to bolster a reliable and predictable service delivery (including but not limited to capacity planning).
- Implement service management processes to govern changes in a controlled way. Make release notes and documentation publicly available. Announce maintenance breaks well ahead of time. Maintain backward compatibility when possible.
- Implement (ideally automated) testing procedures for every change to the service or a service (component) that it integrates with. Testing should ideally include not only functional testing, but also performance and stress testing.
- Consider service scalability, if applicable.
- Implement service management processes to deal with incidents or vulnerabilities in an effective and transparent way. Implement and test disaster recovery procedures. In case of service interruptions, aim to restore service as soon as possible even if that requires workarounds or other temporary measures.
- Implement a service monitoring system that generates alerts in case of unexpected behavior, including functional, performance and security-related issues.
- Implement and make available a set of metrics as indicators for the performance, stability and adoption of the service.
- In addition to single services, also consider service networks and interdependencies.

Open & Connected

Objective:

The service is operated in a low-barrier and inclusive way; seeking integrations and connections with other services; and championing principles of openness consistent with Open Science and Open Research.

Recommendations:

- Publish clear, inclusive and non-discriminatory licences and/or terms of use. Enable wide access to the service.
- Provide guidance about the service licensing to better understand the limitations in usage.
- Seek integrations with other services rather than replicating functionality, especially for common reusable infrastructure components. Provide documentation to ensure better sustainability for the network of integrations. Adopt EOSC architectural components and standards as enablers for deep interoperability with other services in the EOSC portfolio.[3]
- Adopt well-documented and community-supported open standards and specifications, in particular for API's and other interfaces to better understand the service's usage.
- Make the service and all documentation available online through URLs that are fully qualified domain names and assign PIDs where applicable.
- Offer the service with the lowest possible entry barrier for end-users (which does not preclude monetization or cost-recovery models)
- Use community-supported PIDs to integrate with other services; keep data, metadata and PID's tightly connected. Consider implementing the FAIR Digital Object model to enable interoperability with other data services.
- Where possible, make any source code that is used to run the service available under a common open-source licence.[4]
- Seek inclusion in relevant service catalogs, ideally obtaining and using a PID for the service.

[3] Part of the EOSC interoperability framework, the EOSC Profiles (https://data.d4science.net/13af) specify common data models for EOSC entities (Providers, Resources, etc) which helps drive interoperability of resources within EOSC.

[4] See e.g. https://spdx.org/licenses/ for a list of relevant software licences.

Socially-oriented aspects

User centricity

Objective

The service is managed such that it serves the (possibly evolving) goals of the user community, and maximises usability while minimizing burden.

Recommendations:

- Invest in user training and outreach activities to help users understand the service's value proposition and how to effectively use it.
- Ensure the service provider organization has adequate support staff available to assist users where needed.
- Determine and monitor your target user community to understand how the service fits within its data management norms and expectations.
- Ensure that there is an ongoing, consistent dialogue between the service and its user community, such that users can optimally make use of the service and influence its development.
- Ensure that sufficient documentation is available for users and organize a process to regularly review and update (at least with every change to the service). Documentation should cover functional aspects, a description of the various service components and their relationship, and explain which phases of the data life cycle and data management processes are supported by the service. Ideally documentation should be version-controlled, have a PID and an (open) licence.
- Strive for continual improvements to the user experience. In addition to making use of data and service usage statistics, actively work with the community to understand and improve usability, for example through user tests or design studios.
- Include multi-lingual support and accessibility features[5], both for the service and its documentation, to the extent relevant for the service's (potential) user base. Key information must be available in English if the service is intended to be included within EOSC.
- Engage the user community in establishing and prioritizing the service's backlog and roadmap.

[5] For accessibility on the web, we specifically recommend the Web Content Accessibility Guidelines (WCAG) overview: https://www.w3.org/WAI/standards-guidelines/wcag/

Transparency

Objective:
The service provider communicates with its user community in a transparent manner.

Recommendations:

- **Clearly communicate** the service's core value proposition and any pertinent (technical or non-technical) features, as well as its limitations.
- Be open and **transparent** about the organisational mission, business model, legal status and target user communities. Be transparent and accountable about costs, profits and cost-recovery models.
- For services that are meant to preserve research objects over a longer period of time (such as data repositories), state a clear minimum preservation timeframe and provide a contingency and/or preservation plan.
- Implement an appropriate and **transparent** governance structure that includes representation of the service's target user community.
- Be clear about how the service implements community standards.
- Seek to attain certification where relevant community-endorsed certification mechanisms exist.

Longevity

Objective:
The service provider designs the service with a timeframe for the maintenance and sustainability of the service in mind and implements measures accordingly, considering the researchers' necessity for reproducible research.

Recommendations:

- Take reasonable measures to ensure a sustainable long-term operation — including both financial and organisational aspects. Aim to reduce long-term operational dependencies on short-lived project funding. If available, provide clear information to indicate how long the service will minimally be available and maintained.
- Implement technical measures to safeguard the continuity of the service, and the longevity and integrity of any (meta)data that is stored as part of the service. This includes keeping backups on independent systems, implementing fail-over mechanisms and exercising proper life cycle service management.
- Ensure that the service provider organization has sufficient staff with knowledge to operate the service, now and in the future.

Ethical & Legal

Objective:

The service complies with all applicable legal and ethical guidelines, in a transparent and auditable way.

Recommendations:

- Take reasonable measures to manage the intellectual property rights of data producers.
- Define, publish and adhere to a code of conduct that is in accordance with commonly agreed principles regarding the conduct of research in the service's user community.
- Take reasonable measures to ensure data is handled in compliance with disciplinary and ethical norms, and that data licences are clearly defined and respected within global and local legislation.
- Provide clear and user-friendly information about the extent of the data usage/access, in addition to data licences.
- Maintain a publicly available privacy policy.
- Clearly communicate a contact address for security issues including hacks, vulnerabilities and privacy breaches. Ensure the address is actively monitored by multiple staff members.
- Implement auditable measures to ensure that the service respects all applicable legislation and regulations around user privacy and sensitive data (including but not limited to GDPR in Europe). In particular, when processing personal data, roles and responsibilities must always be well-defined and data subjects must be provided with the name and contact details of the data controller and of the Data Protection Officer.