FAIR assessment framework for data services

Patricia Herterich (DCC) and Hylke Koers (SURF)
Agenda

• Welcome & introduction (10 min)
• Methodology & findings (15)
• Basic framework (10)
• Q&A (5)
• Your feedback (15; interactive)
• Wrap up (5)
Call: H2020-INFRAEOSC-5c
Budget: 10 million euro
Length: 36 months
Starting date: March 1 2019
6 core partners/WP leads
Our objective

To supply **practical solutions** for the use of the FAIR data principles throughout the research data life cycle. Emphasis is on **fostering FAIR data culture and the uptake of good practices** in making data FAIR.
The FAIR symphony

Image: European Union Youth Orchestra, euyo.eu
The FAIR symphony needs FAIR services

Priority Recommendation

Rec. 13: “Develop metrics to certify FAIR services”: More work is needed to extend the FAIR data principles for application to a wide range of data services, including registries, Data Management Planning tools, metadata standards and vocabulary bodies, identifier providers, software libraries and other cloud services...

(Also recently underlined by EOSC FAIR WG)
Guidance for service owners to enable FAIR

- FAIR is not an absolute, but rather a set of guiding principles that need further interpretation and definition.
- A lot of work on this has been done for datasets and other digital objects.
- For service owners, there is currently little guidance on how to make their service fit in the FAIR data ecosystem.
Objective

To deliver an *assessment framework* for data services that will help service owners to incrementally improve their services

→ stimulating an optimal interplay between digital objects and services

→ help realize the full potential of a truly FAIR ecosystem
Output 1: Case studies and methodology for ‘FAIR enablement’

### 4.2. Case Study 1: B2FIND

<table>
<thead>
<tr>
<th>Service Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2FIND is a metadata aggregator. The service harvests metadata from different community repositories and harmonizes them such that users and services can search through the combined metadata. B2FIND offers a rich faceted graphical search interface and a HTTP REST API that has been implemented in python for EUDAT’s B2FIND Training.</td>
</tr>
</tbody>
</table>

**Users**

- The service targets two types of user groups:
  - Scientific communities that can provide their metadata and integrate via the B2FIND service with other metadata
  - Scientists who can employ the service to search for interesting research data across different communities simultaneously.

**Services**

- Metadata harvesting and harmonisation to communities with a tool to search across the metadata for scientists.
- The relevant metadata of a DO is shown and a link to the metadata provenance is provided.

**Purpose**

B2FIND is a metadata aggregator. It gathers metadata from communities and repositories and integrates the different types of metadata. It provides a graphical user interface and an API to present the metadata and allows faceted searches across the metadata corpus.

**Adoption**

B2FIND hosts 324556 metadata entries harvested from 22 communities. We were unable to establish from the documentation how many users use B2FIND.

**Documentation**

EUDAT provides guidelines on how to use the B2FIND services as well as detailed guidelines for harvesting and mapping metadata.

**Examples**

- B2FIND entry (BONTROL 1984)
- DAI-PHR dataset’s metadata via B2FIND.

---

**FAIR enablement mapping:**

- **Enable**
- **Respect**
- **Reduce**

---

“M2.7 Assessment report on ‘FAIRness of services’”, available at: [https://doi.org/10.5281/zenodo.3688762](https://doi.org/10.5281/zenodo.3688762)
Output 2: Basic framework for FAIR service assessment

(intentionally left blank... for now...)

11
FAIRness of Services: Methodology and Findings

- Patricia Herterich
Building blocks

Structured literature review

Covering earlier work on service assessment such as CoreTrustSeal, the TRUST principles - but also recent work from EOSC-synergy, EOSC-nordic and other related projects.

EOSC-hub week workshop

Feedback from FAIRsFAIR workshop on FAIR certification of repositories and other data services during EOSC-hub week in May 2020

Interviews with service owners

Semi-structured interviews with a range of service owners
Literature review

- Covering 18 papers and reports
- Analysis
  - Extracting statements (249 in total)
  - Coding by type of statement (recommendation, requirement, principle, action, etc.) and applicability to a wider range of data services

Findings from literature review

● Various projects working on FAIR data object assessment, less focus on FAIR-enabling services

● Work that has been done is characterized by:
  ○ Broad diversity in type of statements → difficult to compare and harmonize
  ○ Some resources focus on certain aspects potentially resulting in an unbalanced view
  ○ Unclear process to implement recommendations and requirements
  ○ Lack of case studies reporting benefits and challenges
  ○ Unclear or abstract relation to EOSC work (e.g. EOSC rules of participation or service management onboarding requirements)
Interviews

• 5 interviews with service owners, approx. 60 minutes each
• Semi-structured interviews covering the following parts:
  1. Understanding the service, its users and context;
  2. Understanding the service maturity;
  3. Understanding affinity and familiarity with FAIR;
Analysis

- Key statements from each interview
- Coded by

Findings from interviews

• Services have been supporting aspects of FAIR before the term was coined
• They have a good understanding of their communities and awareness of the wider policy context
• Most find that automated FAIRness assessment tools could be helpful
• Training for users will be needed
• Sustainable funding will be crucial
• External and community-endorsed assessment framework is preferred over self-assessment
Findings from interviews

• Services have been supporting aspects of FAIR before the term was coined
• They have a good understanding of their communities and awareness of the wider policy context
• Most find that automated FAIRness assessment tools could be helpful
• Training for users will be needed
• Sustainable funding will be crucial
• External and community-endorsed assessment tools are preferred over self-assessment

“It's like a driving instructor. They are allowed to pass things, you know, they're allowed to grant to FAIR badges. But it's not the car company giving you FAIR. It's not the car company saying here's a car, and we'll give you a driver's licence. It's: here's a car, you can pass your driver's licence using this because it passed its MOT and it's all the level that it's safe to drive on the roads. We need the FAIR data equivalent of driving instructors” (M. Hahnel)

“We want a test that we are able to pass, but we nevertheless want it to be quite serious. Nothing is more horrible to an honest operator than a test that doesn't actually test. Self-assessments with free form questions [...] tend to select for people who know how to advertise themselves well.” (G. Aben)
EOSC-hub session

• Feedback on:
  • Scope, i.e. which services are seen to be essential in enabling FAIR data and should thus be in scope for the assessment framework;
  • Desired qualities, i.e. what are seen to be important properties for a data service to be an enabler for FAIR data;
  • Form of the assessment framework, specifically on the continuum from descriptive (sharing recommendations and good practices) to prescriptive (formal certification)
• 90 active contributors, 44 % of which identified as service providers
Findings - scope

Please name three types of data services that you consider essential to enable FAIR

- PID Services
- Repositories
- Metadata Services
- Search & Discovery Services
- Data Access Services
- DMP Services
- Storage, Preservation & Archiving Services
- Registries
- Ontology Services
- Stewardship & Support Services
Findings - desired qualities

What do you consider to be the most important qualities for a data service to enable FAIR data?
Findings - form

How important is it for you that

(i) there are shared ‘good practices' and recommendation for FAIR-enabling services;

(ii) There is a self-assessment tool for FAIR-enabling services;

(iii) There is a formal certification process for FAIR-enabling services
FAIRness of Services: Basic Assessment Framework

Hylke Koers
Proposing: A basic framework for FAIR service assessment

- Technically-oriented
  - FAIR enablement
  - Quality of service
  - Open & Connected

- Socially-oriented
  - User centricity
  - Trustworthiness
  - Ethical & Legal

“M2.10 Report on basic framework on FAIRness of services”, available at: https://doi.org/10.5281/zenodo.4292599
Each aspect has a high-level objective with actionable recommendations

**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:**

- 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.\(^\text{12}\) 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).\(^\text{13}\)
- 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.
- In consultation with the target community (or communities), identify which metadata
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.
Quality of service

**Objective:**

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

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

Objective:

➢ The service is perceived by the user community as **reliable** and **trustworthy**, both in terms of its utility and its warranties, **now and in the future**.
Ethical & Legal

Objective:

➢ The service complies with all applicable legal and ethical guidelines, in a transparent and auditable way.
A basic framework for FAIR service assessment

Technically-oriented
- FAIR enablement
- Quality of service
- Open & Connected

Socially-oriented
- User centricity
- Trustworthiness
- Ethical & Legal

“M2.10 Report on basic framework on FAIRness of services”, available at: https://doi.org/10.5281/zenodo.4292599
## Next steps

<table>
<thead>
<tr>
<th>Time</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec – Jan 2020</td>
<td>Public consultation on assessment framework as proposed in this work – including webinar to kick off the public consultation (perhaps with RDA FAIR maturity WG)</td>
</tr>
<tr>
<td>Feb 2021</td>
<td>Workshop with different stakeholder groups to validate completeness</td>
</tr>
<tr>
<td>March 2021</td>
<td>Dissemination of workshop findings, including next iteration of the assessment framework</td>
</tr>
<tr>
<td>April 2021</td>
<td>Second workshop for service owners to validate utility and actionability</td>
</tr>
<tr>
<td>May 2021</td>
<td>Draw up final framework for assessing FAIR services</td>
</tr>
<tr>
<td>June 2021</td>
<td>“Last chance” external expert review</td>
</tr>
<tr>
<td>June 2021</td>
<td>Internal review within FAIRsFAIR</td>
</tr>
<tr>
<td>July – Aug 2021</td>
<td>Finalization and publication of framework for assessing FAIR services (deliverable D2.7)</td>
</tr>
</tbody>
</table>
Next steps

<table>
<thead>
<tr>
<th>Dec – Jan 2020</th>
<th>Public consultation on assessment framework as proposed in this work – including webinar to kick off the public consultation (perhaps with RDA FAIR maturity WG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb 2021</td>
<td>Workshop with different stakeholder groups to validate completeness</td>
</tr>
<tr>
<td>March 2021</td>
<td>Dissemination of workshop findings, including next iteration of the assessment framework</td>
</tr>
<tr>
<td></td>
<td>Second workshop for service owners to validate utility and actionability</td>
</tr>
<tr>
<td></td>
<td>Draw up final framework for assessing FAIR services</td>
</tr>
<tr>
<td></td>
<td>“Last chance” external expert review</td>
</tr>
<tr>
<td></td>
<td>Internal review within FAIRsFAIR</td>
</tr>
<tr>
<td>July – Aug 2021</td>
<td>Finalization and publication of framework for assessing FAIR services (deliverable D2.7)</td>
</tr>
</tbody>
</table>

February 19
Let us know (e.g. via chat) if you’re interested to join!
Please visit www.menti.com → use code 92 94 76 6
FAIRness of Services: Next Steps

Hylke Koers
Next steps

<table>
<thead>
<tr>
<th>Time</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec – Jan 2020</td>
<td>Public consultation on assessment framework as proposed in this work – including webinar to kick off the public consultation (perhaps with RDA FAIR maturity WG)</td>
</tr>
<tr>
<td>Feb 2021</td>
<td>Workshop with different stakeholder groups to validate completeness</td>
</tr>
<tr>
<td>March 2021</td>
<td>Dissemination of workshop findings, including next iteration of the assessment framework</td>
</tr>
<tr>
<td>April 2021</td>
<td>Second workshop for service owners to validate utility and actionability</td>
</tr>
<tr>
<td>May 2021</td>
<td>Draw up final framework for assessing FAIR services</td>
</tr>
<tr>
<td>June 2021</td>
<td>“Last chance” external expert review</td>
</tr>
<tr>
<td>June 2021</td>
<td>Internal review within FAIRSFAIR</td>
</tr>
<tr>
<td>July – Aug 2021</td>
<td>Finalization and publication of framework for assessing FAIR services (deliverable D2.7)</td>
</tr>
</tbody>
</table>
Finally, thanks to “team 2.4”!

- Tero Alto (CSC)
- Morane Gruenpeter (INRIA)
- Patricia Herterich (DCC)
- Rob Hooft (DTL)
- Hylke Koers (SURF; task lead)

And earlier contributions from:

- Christine Staiger (DTL)
- Roberto Di Cosmo (INRIA)
- Sarah Jones (DCC)
- Jessica Parland-von Essen (CSC; work package lead)
- Jonas Tana (CSC)