Online meeting report

FAIR data maturity model Working Group

Online meeting #1 - 21 and 22 February 2019

Project	RDA FAIR data maturity model working group	Date & Time	21 February 2019 10:00—11:30 UTC 22 February 2019 15:00—16:30 UTC
Type	Online meeting	Location	Online GoToMeeting
Meeting Chair	Keith Russell (21 February), Edit Herczog (22 February)	Issue date	1 March 2019

Objectives

This first online meeting aimed at presenting the approach of the working group (i.e. presentation of the work methodology, a tentative timeline and the scope of the work). Moreover, some relevant initiatives and the results of the preliminary analysis were presented to pave the way for further discussions.

Agenda

1. Welcome, objectives of the meeting
2. Round table
3. Introduction to the Working Group
4. Survey results
5. Presentations from existing approaches
   a. DANS FAIR data assessment tool, FAIR checklist *[Eliane Fankhauser]*
   b. FAIR Metrics *[Luiz Bonino]*
   c. Data Stewardship Wizard *[Rob Hooft]*
   d. RDA SHARC IG *[Laurence Mabile & Romain David]*
   e. Dataset Fitness for Use *[Jonathan Petters]*
   f. ARDC FAIR self-assessment tool *[Keith Russell]*
6. Results of preliminary analysis
7. How to contribute
8. Logistics
9. Conclusion

Participants

The online meeting was attended by 33 participants on 21 February while 35 participants were present on 22 February. A total of 57 people attended, 11 attended both meetings and 46 attended one meeting.

Name		Affiliation(s)	21st	22nd
Alejandra Gonzalez-Beltran	UK	University of Oxford	•	
Alexandre Beaufays	BE	PwC, Editor team support	•	•
Angus Whyte	UK	DCC / H2020 FAIRsFAIR	•	
Athanasios Karalopoulos	BE	European Commission DG RTD	•	•
Barbara Sierman	NL	KB National Library of the Netherlands	•	
Brecht Wyns	BE	PwC, Editor team	•	•
Brian Matthews	UK	Science and Technology Facilities Council		•
Carole Goble	UK	University of Manchester		•
Christine Laaboudi	LU	EU Open Data Portal		•
Claire Austin	CA	Department of the Environment Canada		•
Christophe Bahim	BE	PwC, Editor team support	•	•
Edith Herczog	BE	Chair, Vision & values SPRL		•
Eliane Fankhauser	NL	DANS / H2020 FAIRsFAIR	•	•
Fabienne Meggetto-Pradelle	FR	CRCT		•
Fiona Murphy	UK	Independent Consultant in Research Data and Publishing	•	
Françoise Genova	FR	Observatoire de Strasbourg		•
Ge Peng	US	NOAA's Cooperative Institute of Climate and Satellites		•
Heinrich Widmann	DE	Deutsches Klimarechenzentrum		•
Hilary Goodson	NL	EOSC-hub	•	
Ingrid Dillo	NL	DANS / H2020 FAIRsFAIR	•	
Ines Hessler	US	Dalhousie University		•
Jane Wyngaard	FR	University of Notre Dame		•
John Chodacki	US	California Digital Library		•
Jonathan Petters	US	Virginia Tech		•
Juan Bicarregui	UK	Science and Technology Facilities Council		•
Julian Gautier	UK	Institute for Quantitative Social Science		•

Name	Country	Affiliation		
Julie Arteza	IT	Trust-IT / H2020 FAIRsFAIR	•	
Keith Jeffery	UK	Keith G Jeffery Consultants	•	
Keith Russell	AU	Chair, ARDC	•	
Konstantinos Repanas	BE	European Commission DG RTD	•	•
Laurence Mabile	FR	Paul Sabatier University - Toulouse III	•	•
Luiz Bonino	NL	GO FAIR	•	•
Marialetizia Mari	IT	Trust-IT / H2020 FAIRsFAIR	•	
Maggie Hellström	SE	ICOS ERIC / ENVRI-FAIR	•	
Makx Dekkers	ES	Independent Consultant, Editor team	•	•
Mao Tsunekawa	JP	Chiba University	•	
Marios Meimaris	GD	Athena Research Center		•
Marieke Willems	ES	Trust-IT / SSHOC	•	
Mark Wilkinson	ES	GBGP, UPM – INIA	•	
Matthew Viljoen	UK	EGI / ESOC-hub	•	
Mohamed Yahia	FR	Inist-CNRS / Datacite	•	
Mustapha Mokrane	NL	DANS	•	•
Natalia Manola	GR	Athena Research & Innovation Center / OpenAIRE		•
Nataliya Rozbroj Jasinskaja	LU	EU Open Data Portal		•
Nick Juty	UK	University of Manchester		•
Nikolaos Loutas	BE	PwC, Editor team	•	
Paolo Manghi	IT	Istituto di Scienza e Tecnologie dell'Informazione		•
Pete McQuilton	UK	University of Oxford	•	
Ricardo de Miranda Azevedo	NL	University of Maastricht / GO-FAIR	•	
Rob Hooft	NL	Dutch Techcentre for Life Sciences		•
Romain David	FR	SHARC	•	•
Ronald Cornet	NL	University of Amsterdam	•	
Sara Garavelli	IT	Trust-IT		•
Sarah Jones	UK	Digital Curation Center	•	
Simon Lambert	UK	UKRI-STFC / H2020 FREYA	•	
Siri Jodha Khalsa	US	University of Colorado		•
Yoann Moranville	DE	Centre Marc Bloch		•
			33	*35*

Content

The online meetings were designed to be as interactive as possible; to do so, the attendees were invited to come up with their questions after several agenda items and anytime through the chatbox. There were continuing parallel discussions during the presentations. The participants took part in discussions to understand the practical implications of what was presented. A part of the online meeting was dedicated to present initiatives associated to the objectives of the working group in order to give a first taste of the work that will be done.

1. The Chairs opened the first online meeting, welcomed the participants and addressed the agenda.

2. The Chairs and the editor team introduced themselves; the other participants were kindly invited to share their names and affiliations via chat.

3. The Chairs outlined the problem, the approach and the intended result. The editor team introduced the work methodology with the proposed steps and a tentative timeline for the next six months.

Key points from the discussion
● How will this working group link with non-RDA initiatives? *According to the Working Group approach, existing models and initiatives were considered and analysed in order to know how they overlap and how they can be implemented. Further connections can be proposed.*   ● People's resistance to FAIR principles *People usually prefer their old way of working by keeping their data private and not sharing them. This attitude is an obstacle to implement FAIR and should be taken into account.*   ● Data life cycle assessment (and history record of evaluation)   - *The FAIR assessment does not take into account the data life cycle. Assessment should be done from the planning phase and continue through the life cycle.*   - *The maturity model could consider how FAIRness can be improved throughout the data life cycle.*   ● How will the possible interdependencies between the principles be managed? *The discussion around this question emphasised the approach that was outlined by the editor team: look first at the FAIR principles separately, then as whole in order to better see what is overlapping and how overlaps and gaps can be addressed.*   ● Whose maturity is modelled? (e.g. researchers, institutions, etc.) *There will be discussions until the meeting in Philadelphia to scope whose maturity will be modelled.*

- **What results will be presented and discussed in online meeting #3?**
  The second meeting at RDA13 in Philadelphia will serve to validate the proposed methodology and approach. Concerning online meeting #3, a first set of consolidated criteria, based on the input from the working group, may be presented and discussed.

- **Do FAIR principles also cover services and other assets in addition to datasets?**
  Automated metrics can consider other assets such as software tools.

- **Are there some prerequisites to fulfil FAIRness objectives?**
  FAIR principles have to be considered as guidelines, inspiring concepts but not strict rules. Furthermore, there might be ambiguity behind some concepts related to FAIR principles.

- **Governance of metrics**
  For instance, a system to inform the community that a new metric has been proposed. After a review process, the metric could be approved.

- **Can curation be considered as a *'Reusable'* component?**
  Quality of curation processes are critical for FAIRness (i.e. there is no FAIR data without curation)

- **FAIR has not only technical aspects but also discipline-specific aspects**
  Possible risk to lose existing disciplinary achievements if this assessment remains too high-level (i.e. discipline specific knowledge for the assessment).

4. The editor team mentioned the methodologies that were entered in the online survey and highlighted discussion items derived from these methodologies.

5. A number of experts presented their experience with existing approaches, highlighting the lessons learnt and open issues. The editor team summarised common issues across the presentations.

Key points from the discussion
- **FAIRness of physical objects**   *Physical things certainly need metadata for discovery (i.e. physical entities can be part of the FAIR environment through their FAIR – and digital – metadata)*   - **Automated verification of FAIRness**   *A challenge for automatic verification tools is the consideration of end-user needs and point of view.*   - **Assessing and scoring FAIRness**   - *Essential criterion to be FAIR*

> - *Definition of 'passing' a test (e.g. YES/NO approach too black and white)*
> - *To help measure progress / efforts*
> - *History of evaluation*
> - *Re-assessment over time because the data can decay*
> - *Careful consideration of the results (e.g. the validity of a PID does not ensure its validity over time)*
>
> ● Generic versus discipline-oriented FAIR self-assessment
> *Approach towards a generic base level with domain specific considerations. (E.g. findability might not be a priority for a certain discipline if the community knew where to search for relevant data.)*

6. The editor team reported on the preliminary analysis carried out in a desk research and pointed to the resulting material that can be discussed on GitHub.

> **Key points from the discussion**
>
> ● Certification as a way to mitigate human interpretation
> - *Human interpretation to assess FAIRness might be misleading because of subjective perception. That is why a certification, which tends to be more objective, could mitigate this issue.*
> - *Yet, certification shouldn't only be to run a software tool to test the FAIRness of a digital object.*
> - *Certification should have a clear purpose, and clearly define the context.*
>
> ● FAIRness application to different communities
> - *Standards might be interpreted differently in different communities*
> - *How to demonstrate meeting community standards*
> - *How can machine interpret the above*
> - *FAIR is contextual, it should be trusted by its target audience (i.e. every community doesn't have the same minimal FAIRness level)*

7. The editor team described how the participants can contribute to the Working Group.

8. The editor team reminded to the participants where the different information was located.

9. The Chairs concluded the meeting by presenting the next steps and action items.

Outstanding issues and open questions

In this section are presented some outstanding questions and issues to be addressed first in preparation for the plenary session in Philadelphia.

Outstanding questions and issues that require further attention for the next online meetings
- Approach proposed   *Does the WG agree with the methodology proposed? Are there any suggestions for improvement?*   *Does the WG agree with the timeline proposed?*   - Scope of the assessment   *What will be the assessed entity (e.g. metadata, dataset, digital resource, other)?*   *When will FAIRness be assessed (e.g. creation, publication, periodically)?*   *What will be the nature of the assessment (e.g. generic versus domain-specific)?*   *What will be the profile of the respondent (e.g. researcher, data librarian, data steward)?*   *To whom would the results be targeted (i.e. target audience)?*   *What could be the format (e.g. automated versus self-assessed, what proportion could be automated?)*   - FAIRness evaluation   *What would be the essential criteria to be FAIR?*

Follow-up Action Plan

- Continue the discussion about outstanding issues and open questions on the GitHub foreseen
- Prepare for the [face-to-face meeting](#) at the 13th RDA plenary in Philadelphia, Pennsylvania, USA, on 3 April 2019, 12:00-13:30 EDT (17:00—18:30 UTC).

Useful links
[RDA FAIR data maturity model WG](#) [RDA FAIR data maturity model Case Statement](#) [Online meeting #1 presentation](#) [RDA FAIR data maturity model GitHub](#)