

# Workshop report

## 13th RDA Plenary - Workshop #2

Project	RDA FAIR data maturity model working group	Date & Time	3 April 2019 16:00-17:30 UTC
Type	Workshop	Location	Philadelphia, PA, USA
Chairs	Edit Herczog, Keith Russell		
Reporting	Kathleen Gregory Christophe Bahim Makx Dekkers Brecht Wyns	Issue date	06/05/2019

## Objectives

The primary objective of the second workshop was to approve the approach of the Working Group (i.e. presentation of the work methodology, a tentative timeline and the scope of the work), presented at the first workshop on 21/22 February 2019. In doing so, the foundations would be established allowing the Working Group to focus on developing a set of core assessment criteria. Furthermore, the second objective of this second workshop was to give a hands-on demonstration on how to develop the core assessment criteria.

## Agenda

1. Welcome, objectives of the meeting
2. Round table
3. Approval of methodology and scope
4. Report on discussions on GitHub
5. Hands-on exercise
6. Action items and next steps

### ***Useful links***

[RDA FAIR data maturity model WG](#)  
[RDA FAIR data maturity model Case Statement](#)  
[Workshop #2 presentation](#)  
[RDA FAIR data maturity model GitHub](#)

# Participants

More than 52 people physically attended the workshop while 8 people attended remotely. Below you can find the names of a number of persons that attended. Please note that this list is not complete.

<b>Name</b>		<b>Affiliation</b>	<b>Presence</b>	
			On site	Online
Amanda Lindsay	US	UC Davis	•	
Anne Lambon-Thomsen	FR	CNRS	•	
Ardrian Burton	AU	ARDC	•	
Ari Asmi	FI	University Helsinki	•	
Astrid Verheusen	NL	LIBER	•	
Athanasios Karalopoulos	BE	European Commission DG RTD	•	
Ben Schaap	NL	Godan	•	
Bonnie Carroll	US	Information International Associates	•	
Brecht Wyns	BE	PwC, Editor team	•	
Brian Matthews	UK	STFC	•	
Chris De Loof	BE	Belspo		•
Christophe Bahim	BE	PwC, Editor team		•
Claire Austin	CA	Department of the Environment Canada		•
Daniel Jaquette	US	LDC	•	
David Carr	UK	Wellcome	•	
David Groenewegen	AU	Monash University	•	
Dawei Lin	US	NIH/NIAID	•	
Edith Herczog	BE	Chair, Vision & values SPRL	•	
Ewa Zegler-Poleska	US	Indiana University	•	
Fabrizio Gagliardi	ES	BSC	•	
Fernando Aguilar	-	N/A		•
Fiona Murphy	UK	Independent Consultant in Research Data and Publishing		•
Françoise Genova	FR	CDS	•	
Françoise Pearlman	US	FourBridges	•	
Ge Peng	US	North Carolina State University/NCEI	•	

Hannah Calckins	US	Children's Hospital of Philadelphia	•	
Ian Bruno	UK	CCDC	•	
Jake Carlson	US	University of Michigan	•	
Jane Wyngaard	US	University Notre Dame	•	
Jay Pearlman	US	FourBridges	•	
John Watkins	UK	UKRI	•	
Jonathan Petters	US	Virginia Tech	•	
Juan Bicarregui	UK	UKRI-STFC	•	
Kathleen Gregory	NL	DANS	•	
Kathleen Shearer	CA	COAR	•	
Keith Jeffery	UK	Keith G Jeffery Consultants	•	
Keith Russell	AU	Chair, ARDC	•	
Kiera McNeice	UK	Cambridge University	•	
Konstantinos Repanas	BE	European Commission DG RTD		•
Laurence Mabile	FR	Toulouse University	•	
Laurents Sesink	NL	Leiden University	•	
Maggie Hellström	SE	Lund University		•
Magnus Eriksson	SE	Swedish Research Council	•	
Max Dekkers	ES	Independent Consultant, Editor team		•
Mercè Crosas	US	Harvard University	•	
Mike Brown	UK	CEM	•	
Mustapha Mokrane	NL	DANS	•	
Nick Juty	UK	University of Manchester	•	
Nikolaos Loutas	BE	PwC, Editor team	•	
Paolo Manghui	IT	CNR-ISTI	•	
Pascal Suppers	NL	DataHub Maastricht	•	
Pradeep George	SE	INCF	•	
Romain David	FR	INRA	•	
Shelley Stall	US	AGU	•	
Sophie Aubin	FR	INRA	•	
Susan Gregurick	US	NIH	•	
Susanna-Assunta Sansone	UK	University of Oxford	•	

Tobias Weber	DE	LRZ	•	
Xenia Specka	DE	ZALF	•	
Zalmary Trauit	US	NIST	•	
			<b>52</b>	<b>8</b>
			<b>60</b>	

# Content

The workshop was designed to be as interactive as possible, giving the attendees the opportunity to contribute their thoughts. As a result, the meeting was fruitful and enabled lively discussions.

1. The Chairs opened the workshop, welcomed the participants and addressed the agenda. The approach to the Working Group was presented:
  - a. Challenges rising from the different interpretations of FAIRness
  - b. Bringing together the relevant stakeholders to discuss and build on existing expertise and different approaches
  - c. Intended results

The Chairs outlined that there is already a potential first adopter of the *to-be* RDA recommendation, which is the European Open Science Cloud (EOSC).

2. The Chairs and the editor team introduced themselves. The Chairs asked the participants in the room to raise their hands according to their profiles.

3. The Chairs presented the methodology, timeline and scope that were proposed during the first workshop. The methodology is articulated in four phases; *definition - development - testing - delivery*. With regard to the timeline, the Chairs put forth that the third workshop will be held in June 2019, during which the results of the development phase will be presented and discussed. Then, the scope was narrowed down to the following aspects; *entity - nature - format - time - respondent - audience*.

### **Key points from the discussion**

Many discussions happened about the scope of the assessment. Three different aspects were significantly debated: *nature – format – audience*.

*The audience reminded that it is important to consider the context in which the dataset and data-related objects will be assessed, it was mentioned that repositories, communities and disciplines might have different requirements.*

*Part of the audience was in favor of a manual assessment because an automated assessment is too complicated. Moreover, sometimes the only possibility is to have an entity certified by a human. The advocates of the automated assessment put forth the difficulty to remain objective using a manual assessment. In addition, a manual assessment is hindering FAIR for machines and scalability. Removing the automated part of the to-be assessment would mean a great loss.*

*Concerning the audience of the assessment, it was put forth that ‘developers’ (i.e. people who would build automated assessment tools) need to be involved as well.*

The outcomes of these discussions are the following;

*Nature: the focus will be on cross-domain aspects, but for future extensions, different domains will be considered as they have different requirements or objectives.*

*Format: the focus will be on the manual assessment for the time being but the potential for automated assessment may be considered later; scalability is an important issue, and automated assessment may scale better than manual assessment, although some aspects may continue to require human intervention.*

*‘Developers’ need to be involved in the process of developing a common set of core criteria.*

*The rest of the elements presented in the scope, namely entity – time – respondent - were not challenged and consequently accepted.*

The Chairs reminded the audience that the overarching goal is to define core assessment criteria, the question whether criteria could be implemented in an automated or manual way can be further debated. Additionally, the timeline is limited; consequently, some discussions should take place once the objectives set in the case statement are reached.

4. The Chairs summarized the discussions that have occurred on GitHub:
  - a. Definition of Findability
  - b. Definition of Accessibility
  - c. The different flows beyond FAIR (e.g. data flow, legal aspects, financial flow)
5. The editor team walked the audience through the hands-on exercise. The purpose of this exercise was to give the Working Group a taste about “*how to develop a set of core assessment criteria*” (i.e. development phase from the methodology).

The European Commission reminded that the goal of the Working Group is to identify a minimum, commonly agreed, set of indicators that can be used. The goal is not to create yet another maturity model, but rather to build on existing approaches to create comparability across communities and repositories.

#### ***Key points from the discussion***

- Are the communities deciding which indicators to implement?  
*The Working Group will decide which indicators are part of the set of core assessment*
- Indicators need to be realistic and assessable. The current technology at our disposal only allows to verify the presence of a URI and not what is beyond (i.e. can machines make the assessment?)
- Different points were raised with regards to R.1.1
  - *There is a need to verify that a licence is valid for a particular object (e.g. software, data)*
  - *R.1.1 should be different for data object(s) and metadata object(s)*
  - *To correctly understand a licence, a set of ontologies is needed to tell you what lies behind a licence*

The outcomes of these discussions are as follows; each FAIR principle will be further analysed and discussed by the WG on the GitHub repository and other means offered. The chairs closed the discussion by indicating that the editor team will compare the existing FAIR methodologies and propose a core baseline for discussion.

6. The editor team closed the workshop by first raising two questions to the audience:
  - What can be the nature of the RDA recommendations & outputs?
  - How to keep the Working Group involved?

Secondly, they put forth the action items (i.e. development of the core assessment criteria) and called for volunteers. Lastly, the date of the next online workshop was announced to take place on 18 June 2019, taking place twice to allow for participation from different time zones, at 07:00 UTC and at 15:00 UTC.

### ***Key points from the discussion***

- The concept of FAIR keeps on evolving and changing  
*The FAIR principles and their understanding have moved beyond what was initially intended by its authors. The community has adopted the FAIR principles and in the process, created a range of interpretations. In consequence, the community now has to define how they want to measure FAIRness (i.e. the WG should define a common meaning of the FAIR principles through a good description of the assessment criteria).*
- Different stakeholders may have different interpretations of what a minimum indicator will be  
*The WG should aim reach and document a common understanding as much as possible*
- Need to have explanatory documentation for the different stakeholders as well as tools to satisfy the indicators

The Chairs closed the discussion mentioning that it is out of scope for the Working Group to deliver extra documentation and tools other than what is foreseen by the case statement.



# Follow-up Action Plan

- Development of the core assessment criteria on GitHub and on [the collaborative spreadsheet](#)
- **Volunteers** leading the development of the core assessment criteria per specific principles
  - Analysis of all the FAIR principles
    - [FAIR – Findable](#)
    - [FAIR – Accessible](#)
    - [FAIR – Interoperable](#)
    - [FAIR – Reusable](#)
    - [Beyond FAIR](#)
  - Comparison and consolidation of the metrics per principle
  - Identification of levels per metric
  - Pathways of improvement per metric

The next and **third workshop** will take place online

**18 June 2019 at 07:00 UTC (local times)**

**18 June 2019 at 15:00 UTC (local times)**

**Location:** <https://global.gotomeeting.com/join/560494093>