Priority issues for the RDA FAIR Data Maturity Model Working Group

3rd of December 2020

International FAIR Convergence Symposium

Agenda

- Presentation of the agenda and objectives of the meeting [5 minutes]
- History and achievements of the work in the FDMM WG [5 minutes]
- Results of the survey on FAIR assessments [5 minutes]
- Moderated discussion [90 minutes]
- Conclusions and agreement on further work [15 minutes]

Tour de table

In the chat window, please type…

- Your name
- Your affiliation
- Your role
  - Researcher
  - Librarian
  - Service provider
  - Policy maker
  - Funder

Context

The principles are **NOT** strict

- **Ambiguity**
- Wide range of **interpretations** of FAIRness

Different **FAIR Assessment** Frameworks

- Different metrics
- No comparison of results
- No benchmark

SOLUTION is to bring together **stakeholders** to build on **existing approaches** and **expertise**

- Set of **core assessment criteria** for FAIRness
- FAIR **data maturity model & toolset**
- FAIR data **checklist**
- RDA recommendation

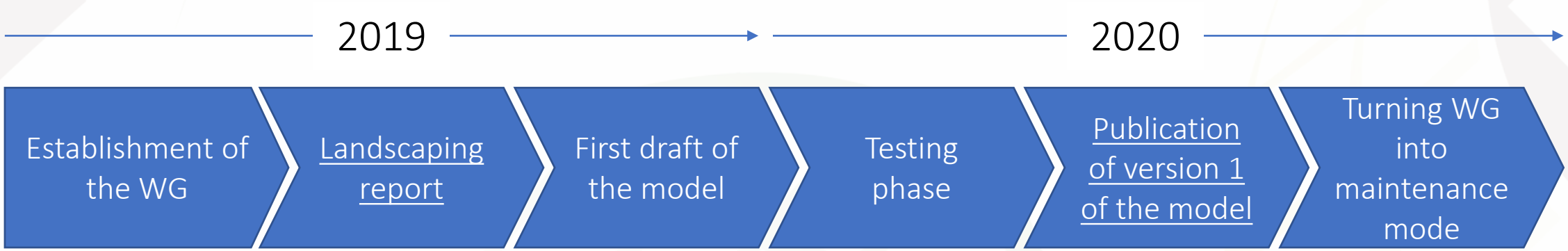
Join the **RDA** Working Group: [RDA WG web page]

Scope

BUT the Working Group does **NOT** have the purpose to …

- **develop yet-another-evaluation-method**: the core criteria are intended to provide a common 'language' across evaluation approaches, not to be applied directly to datasets.
- **define how the core criteria need to be evaluated**. The exact way to evaluate data based on the core criteria is up to the owners of the evaluation approaches, taking into account the requirements of their community
- **revise and re-design the FAIR principles**

History of the FAIR data maturity model WG



Remaining milestones

State of play

- Publication of the FAIR Data Maturity Model: Specification and Guidelines as an RDA recommendation, 25 June 2020
- EOSC-SYNERGY Intermediate report on technical framework for FAIR principles implementation, 2 Sept. 2020
- Publication of The FAIR Data Maturity Model: An Approach to Harmonise FAIR Assessments as a paper in the Data Science Journal, 27 Oct. 2020

Roadmap (remainder 2020)

- CODATA FAIR symposium
- Establish work plan for 2021

Survey on FAIR assessments

Problem statement

Improve the understanding of benefits and challenges of the FAIR assessments from the perspective of the funders and communities.

Objectives

1. Formulate conclusions and recommendations on the level of policy, (i.e. better understanding of the perspectives of both sides)
2. Finding out how the research community and the funders' community might want to use the model and what changes they would want to see

Format and output

- E-survey
- Voluntary basis
- 8-point questionnaire

Anonymise, compile and analyse the results, draw conclusions and propose recommendations for further actions to improve understanding

Survey on FAIR assessments - results

Policy & adoption

- Funders
  - Funders to set requirements for assessment
  - Professional societies and infrastructure providers to drive adoption of FAIR assessment methodologies
  - Cross-discipline interoperability through awareness raising across communities with community-specific approaches
- Communities
  - Funding for global standards, regional implementation
  - Communities to share best practices and develop community policies and competence centres
- Both:
  - RDA as a neutral platform to bring stakeholders together and create cross-community understanding

Future work

- Funders
  - Set the bar for 'FAIR enough'
  - Select most relevant indicators in specific context
- Communities
  - Gather experience and success stories that show impact
  - Pay attention to choices to be made before data is produced
- Both
  - Consider scoring in context of community targets and practices
  - Make assessment approach more practical
  - Create actionable guidelines

Topics for the next version of the model

63%
- Metadata practices:
  - Role of generic platforms / repositories in improving domain-specific metadata
  - Metadata at several levels (collection level, individual level)
  - Consensus on minimum level for 'Rich' metadata
  - Shared understanding of knowledge representation

50%
- Data granularity (collection, dataset, data item)

44%
- Approaches toward evaluation of FAIR assessment tools and services, taking into account community aspects

44%
- PID practices across communities (identifiers for metadata, data, separately, combined)

31%
- Cross-community interoperability and evolution of standards

31%
- Role of landing pages and human readable documentation

25%
- Consensus within the communities for the priorities of the indicators, respecting the different targets for FAIRness and speeds of implementation

Discussion items

Evolution of metadata practices to improve FAIRness within and across communities

Challenges of different data granularities for FAIRness (collections, datasets, data items)

Aspects to be considered for evaluation of FAIR assessment tools and services

Discussion item #1

Evolution of metadata practices to improve FAIRness within and across communities.

There has been much discussion about how to increase the amount of metadata for discovery and reuse. So far, there has not been much progress in identifying a minimum set of descriptors that could enable reuse within and across communities.

1. If there are different approaches to ‘rich’ metadata, is there a need for mapping services to enable wider interoperability? If so, who would be best placed to develop such services?

2. What could be the role of repositories in improving FAIRness of domain-specific metadata?
How can repositories pave the way towards cross-domain metadata exchange?

3. What could be the role of Semantic Web/Linked Data/Knowledge Graph approaches in understanding knowledge representation?

Discussion item #2

Challenges of different data granularities for FAIRness (collections, datasets, data items)

There may be a mismatch between what data publishers think are the reusable entities and what the potential reusers would expect to find.

1. Is the granularity something that individual data creators decide, or is it standardised?
Should there be domain-specific or cross-domain agreements?

2. How do decisions on granularity affect interoperability and wider reuse? Should the FAIR
Data Maturity Model include examples of best practices for this?

Discussion item #3

Aspects to be considered for evaluation of FAIR assessment tools and services

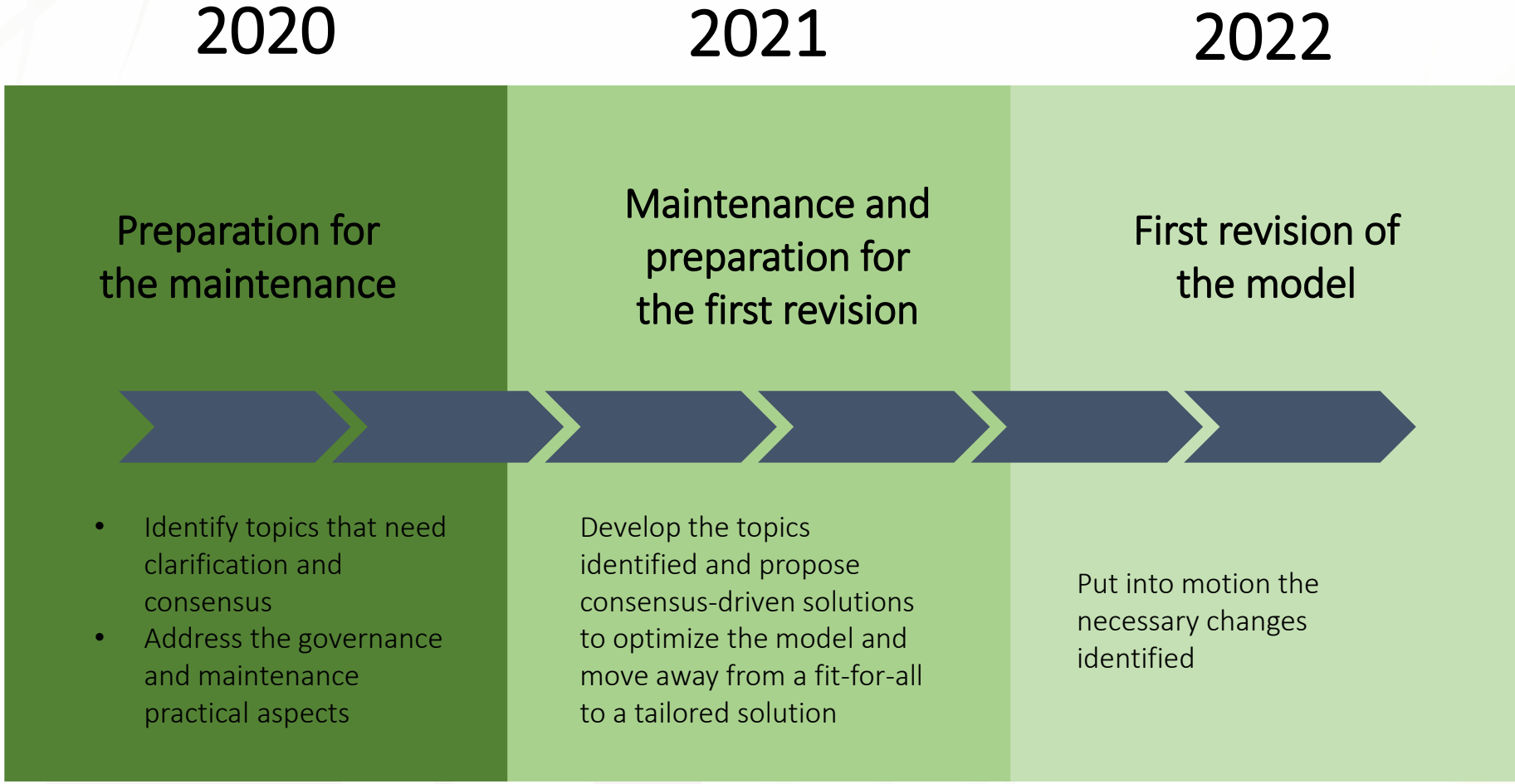
An assessment approach like the FDMM is necessarily focused on a high, cross-domain level. There was agreement in the WG that there is a need for more specific community-driven derivations of the high-level approach to take into account community-specific resource types and targets, by defining a tailored set of indicators and priorities for a particular environment.

1. Is it important to look at the distribution of effort across the four FAIR areas? E.g. not just F and A, but also sufficient attention to I and R.

2. How can tools tailored to a specific resource type or domain be compared? Does that stop at assessing whether they correctly implement a specific approach or should the evaluation also consider whether the tools and services contribute to cross-domain interoperability?

3. How can an initiative like the GO-FAIR FAIR Implementation Profiles help in understanding the results of assessments done by different tools?

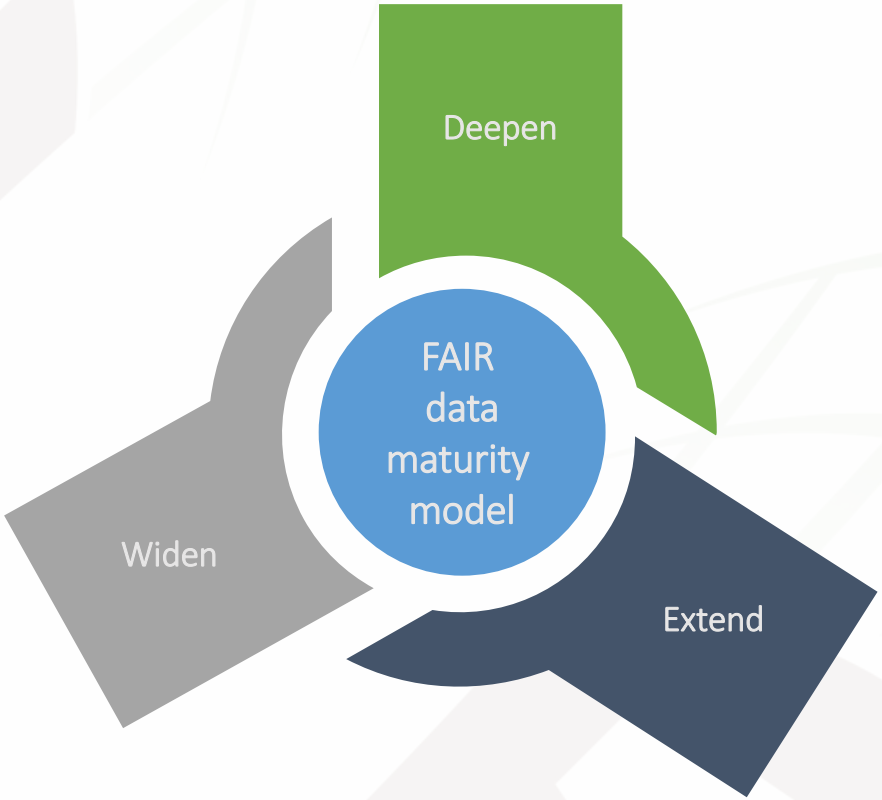
Maintenance and governance



Next steps

- Create further connections with other groups – having an interest in contributing to FAIR assessments – for instance;
  - FAIR for Research Software (FAIR4RS) WG
  - FAIRsharing Registry WG: connecting (meta)data standards, repositories and policies
  - CURE-FAIR WG (FAIR curation)
  - WDS/RDA Assessment of Data Fitness for Use WG
  - Etc.
- Identify which indicators of the FDMM are the ones to begin the FAIR journey with
- Map the FDMM indicators to the different stages of the lifecycle of (FAIR) data
- Look into the 'community of practice' (CoP) for the FDMM

Next steps



- *Deepen* – Work on different elements to understand them better (e.g. metadata, granularity etc.)
- *Widen* – Follow up on the FDMM indicators and the FAIR assessment methods in different domains, sectors (e.g. private and public research) and regions.
- *Extend* – Follow the FDMM indicators through the data lifecycle (e.g. from the creation to archiving)

FAIR data maturity model

Stay in touch!
https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

RDA Webpage

Evaluation mechanism

RDA recommendation

- RDA FAIR data maturity model WG – GitHub
- RDA FAIR data maturity model WG – Mailing list

images: Flaticon.com