Adoption of the FAIR Data Maturity Model

18 June 2020
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](https://rd-alliance.org) | [GitHub](https://github.com)
Public review period complete now to council

THANKS TO ALL REVIEWERS

3600+ page views

14 comments

Adoption examples
Early adopters – Experience sharing

- Ge Peng | NOAA
- Anusuriya Devaraju | FAIRsFAIR

... will share their relevant experience with regard to the adoption of the FDMM and answer to the following questions:

1. What is the level of adoption at your organisation? (E.g., pilot, production, ...)
2. Do you plan to continue to use the Recommendation?
3. Did you need to modify the Recommendation for your use?
4. Can you give an estimate of how much time / effort you have spent on the adoption so far?
5. What’s your overall experience? (E.g., Very Good, Good, Fair, Poor)
6. Would you do it again?
Evaluating the FAIRness of Environmental Data  
– Application of the RDA FAIR Data Maturity Indicators

Ge Peng, PhD

Cooperative Institute for Satellite Earth System Studies (CISESS) Between U.S. National Oceanic and Atmospheric Administration (NOAA) and North Carolina State University at NOAA National Centers for Environmental Information (NCEI)

#9 Workshop of the RDA FAIR Data Maturity Model Working Group, May 20–21, 2020
Purposes of Pilot Application

- Examine the relevancy of the RDA FAIR DMIs (v0.04)
- **Baseline the FAIRness of NCEI managed data**
  - In particular, *OneStop*-Ready datasets,
    - *OneStop* project was Initiated in 2015 to improve discovery and access services for NOAA datasets.
  - What worked?
- **Identify potential gaps & define path forward in NCEI data sharing practices**
Adopting OAIS RM & DSMM Helped!

Mapping FAIR Data Principles to NCEI/CICS-NC Data Stewardship Maturity Matrix (DSMM)

<table>
<thead>
<tr>
<th>FAIR Data Principles (Wilkinson et al. 2016)</th>
<th>DSMM Key Components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Preservability</td>
</tr>
<tr>
<td>F1. (meta)data are assigned a globally unique and eternally persistent identifier</td>
<td>L3</td>
</tr>
<tr>
<td>F2. data are described with rich metadata (defined by R1 below)</td>
<td>L3</td>
</tr>
<tr>
<td>F3. metadata clearly and explicitly include the identifier of the data it describes</td>
<td>L3</td>
</tr>
<tr>
<td>F4. (meta)data are registered or indexed in a searchable resource</td>
<td></td>
</tr>
<tr>
<td>A1. (meta)data are retrievable by their identifier using a standardised communications protocol</td>
<td>L2 &amp; L3</td>
</tr>
<tr>
<td>A1.1. the protocol is open, free, and universally implementable</td>
<td>L3</td>
</tr>
<tr>
<td>A1.2. the protocol allows for an authentication and authorization procedure, where necessary</td>
<td></td>
</tr>
<tr>
<td>A2. metadata are accessible, even when the data are no longer available</td>
<td>L2</td>
</tr>
<tr>
<td>I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation</td>
<td>L3</td>
</tr>
<tr>
<td>I2. (meta)data use vocabularies that follow FAIR principles</td>
<td></td>
</tr>
<tr>
<td>I3. (meta)data include qualified references to other (meta)data</td>
<td>L3</td>
</tr>
<tr>
<td>R1. (meta)data are richly described with a plurality of accurate and relevant attributes</td>
<td>L3</td>
</tr>
<tr>
<td>R1.1. (meta)data are released with a clear and accessible data usage licence</td>
<td></td>
</tr>
<tr>
<td>R1.2. (meta)data are associated with detailed provenance</td>
<td></td>
</tr>
<tr>
<td>R1.3. (meta)data meet domain-relevant community standards</td>
<td>L3</td>
</tr>
</tbody>
</table>

Many data stewardship quality attributes are not explicitly addressed by the FAIR Data Principles.

- Most of data are open by default,
- Use agreements or use constraints,
- CC license not yet explicitly included.

* Can be easily implemented via relevant metadata entity and modified document template.

(Version: v00r01 20200403; POC: apeng@ncsu.edu; CC-BY 4.0)
Path Forward

Improving the FAIRness of NCEI & NOAA Data

- Explicitly include a data usage license, e.g. CC-BY 4.0; CC0, in the metadata record:
  - Discussions are on-going,
  - Procedure is under development.

Extending the Application Scope – under discussion

- Assess: 200+ additional NCEI datasets,
  - produced by NCEI’s Center for Weather and Climate, various stages of OneStop-ready.
- Examine the scalability of the evaluation.

Integrating Assessment Results - Fairly

- Community guidelines – consistently curating and representing dataset quality information,
- Virtual workshop on July 13, 2020 – bringing together international domain experts,
- Contact me at gpeng@ncsu.edu if interested in participating or contributing.
RDA FAIR Data Maturity Model Adoption
(Impression and Experience)
Anusuriya Devaraju & Hervé L'Hours
(on behalf of FAIRsFAIR)
Repository Certification

• CoreTrustSeal follows a self-assessment and peer review model

• FAIRsFAIR is offering support with a CoreTrustSeal+FAIR angle

• Map object characteristics to where repositories can enable FAIR
Repository Certification

• CoreTrustSeal follows a self-assessment and peer review model

• FAIRsFAIR is offering support with a CoreTrustSeal+FAIR angle

• Map object characteristics to where repositories can enable FAIR

Later:

• Integrate object evaluation outcomes
Overall Adoption Experience

- The recommendation should be used as a starting reference point for data FAIRness assessment.
- Presentation - specification and guidelines are well structured!
- ‘What’ aspect of FAIR assessment
  - Descriptions of indicators are very helpful!
  - Suggestion - Include priority level next to each of the indicators.
  - Essential I-indicators missing (needs further work or not important?)
- ‘How’ aspect of FAIR assessment
  - Context matters (e.g., practices, data types)
  - Assessment details not always provide sufficient detail to implement tests.
  - Potential supporting technologies and services should be described.
Next steps

- Reach out to your communities as for the publishing of the FAIR data maturity model: specification and guidelines (i.e. RDA recommendation)

- Continuously provide feedback to the Editorial Team and pass on information with regards to the use of the FAIR data maturity model: specification and guidelines (i.e. RDA recommendation)

The editorial team will look into a release calendar and change management schedule

WORKSHOP #10

Possibly September 2020
Thank you!
Background slides
Objectives

What are to be evaluated to determine FAIRness?

- Identify the indicators that can serve as core criteria
- Propose guidelines and a checklist
- Test the core criteria
- Enable the development of automated tools for evaluation
- Update the core criteria based on feedback
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**