

# The RDA FAIR Data Maturity Model Licences in Astronomy

Françoise Genova

Work done with François Bonnarel (CDS), Marco Molinaro (INAF),

Mark Allen (CDS)

**IVOA Virtual Interoperability Meeting** 

**May 2020** 

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.





### The Research Data Alliance



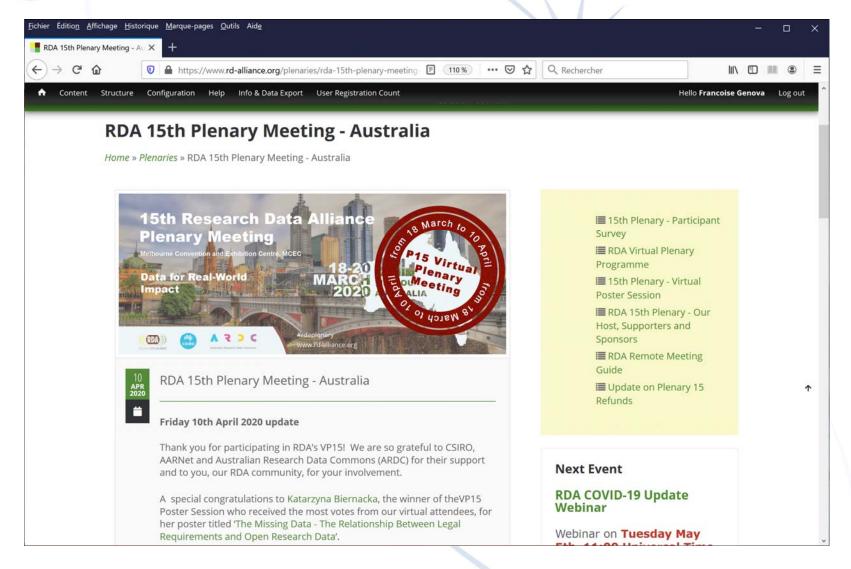
- Created in 2013 by the Australian government, the European Commission and the NSF (USA)
- A neutral, open, international forum to discuss all the aspects of scientific data sharing and produce recommendations and other outputs
- Yesterday 10 599 members from Today 9778 members from 144 countries
- 103 Working Groups and Interest Groups tackling many different subjects – technological and sociological
- https://www.rd-alliance.org/







# We are not the only ones...





Horizon 2020 - Grant N° 824064





## Many relevant activities, among which...

- Data Repository Audit and Certification Recommendation >>> Core Trust Seal
- CORE TRUST SEAL
- ➤ Discussed at Paris Interop, May 2019, DCP Session
- FAIR Data Maturity Model Working Group
  - Core criteria for FAIR data







#### The FAIR guiding principles: https://doi.org/10.1038/sdata.2016.18

#### To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

#### To be Accessible.

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
- A1.1. the protocol is free, open and universally implementable
- A1.2. the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

#### To be Interoperable:

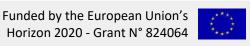
- 11. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
- (meta)data uses vocabularies that follow FAIR principles
- 13. (meta)data include qualified references to other (meta)data

#### To be reusable:

- R1. (meta)data are richly described with a plurality of accurate and relevant attributes
- R1.1. (meta)data are released with a clear and accessible data usage license
- R1.2. (meta)data are associated with data provenance
- R1.3. (meta)data meet domain relevant community standards

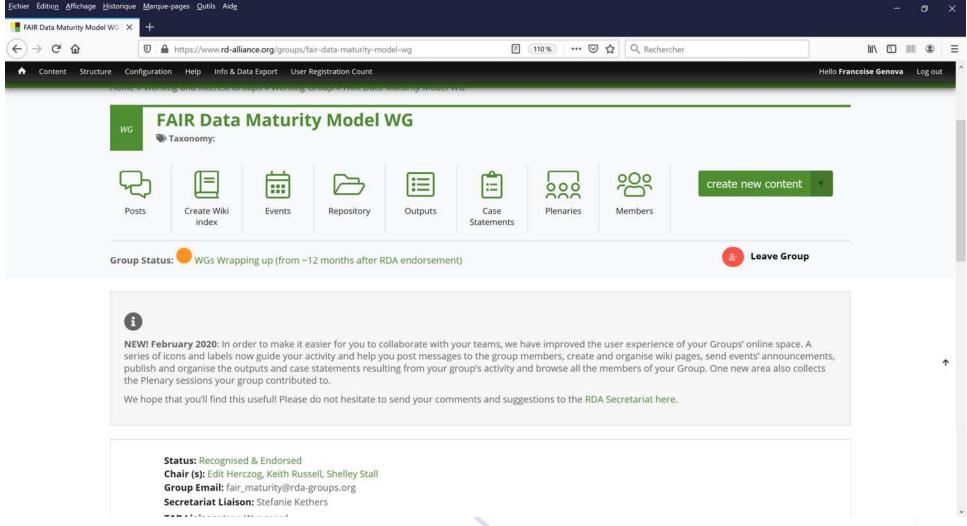
Figure 2. The FAIR guiding principles

Wilkinson et al. 2016, Scientific Data 3, 160018











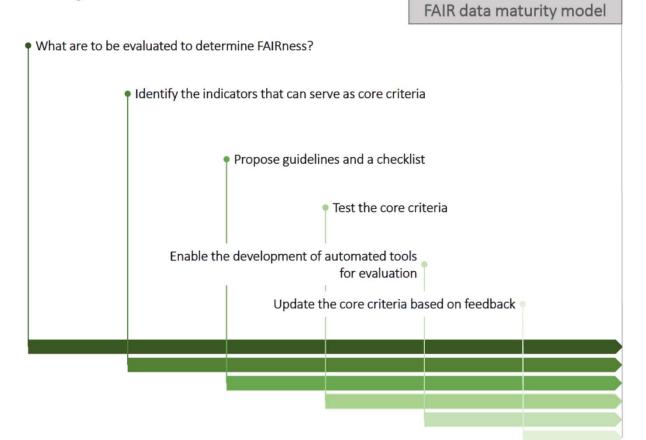




# The WG Objectives and timeline



### Objectives



23/03/2020

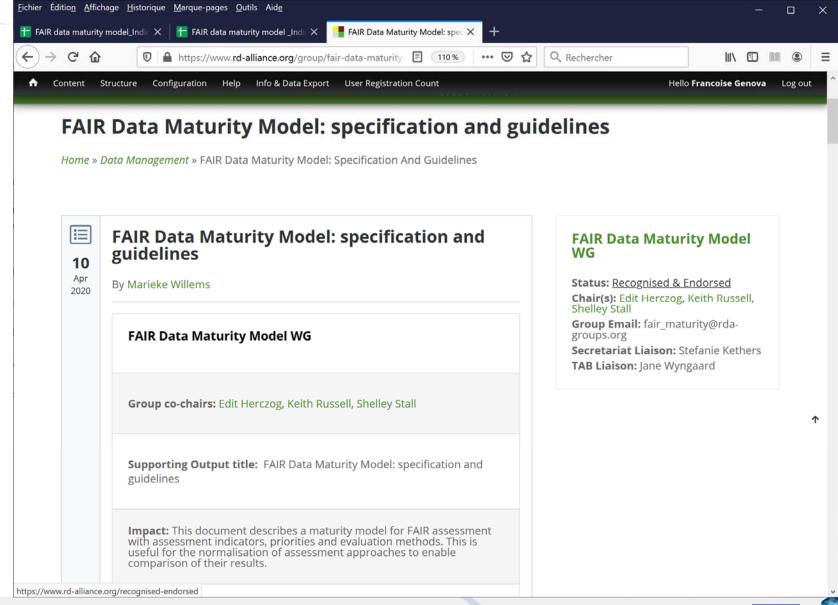
www.rd-alliance.org - @resdatall













# RDA Request for Comments

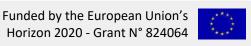
Recommendations and guidelines

https://www.rd-alliance.org/group/fair-datamaturity-model-wg/outcomes/fair-data-maturitymodel-specification-and-guidelines

RfC open until 13 May 2020

Three levels of importance are defined:

- Essential: such an indicator addresses an aspect that is of the utmost importance
  to achieve FAIRness under most circumstances, or, conversely, FAIRness would be
  practically impossible to achieve if the indicator were not satisfied.
- Important: such an indicator addresses an aspect that might not be of the utmost importance under specific circumstances, but its satisfaction, if at all possible, would substantially increase FAIRness.
- Useful: such an indicator addresses an aspect that is nice-to-have but is not could indicator MAY be satisfied, but not necessarily indispensable.









# FAIR Data Maturity Model Specification and Guidelines 2020

Proposed RDA Recommendation
Produced by: FAIR Data Maturity Model WG, 2019-2020

https://www.rd-alliance.org/groups/fair-data-maturity-model-wg







FAIR	ID	Indicator	Priority	
F1	RDA-F1-01M	Metadata is identified by a persistent identifier	000	Essential
F1	RDA-F1-01D	Data is identified by a persistent identifier	000	Essential
F1	RDA-F1-02M	Metadata is identified by a globally unique identifier	000	Essential
F1	RDA-F1-02D	Data is identified by a globally unique identifier	000	Essential
F2	RDA-F2-01M	Rich metadata is provided to allow discovery	000	Essential
F3	RDA-F3-01M	Metadata includes the identifier for the data	000	Essential
F4	RDA-F4-01M	Metadata is offered in such a way that it can be harvested and indexed	•••	Essential
A1	RDA-A1-01M	Metadata contains information to enable the user to get access to the data $% \left( 1\right) =\left( 1\right) \left( 1\right)$	00	Important
A1	RDA-A1-02M	Metadata can be accessed manually (i.e. with human intervention)	000	Essential
A1	RDA-A1-02D	Data can be accessed manually (i.e. with human intervention)	000	Essential
A1	RDA-A1-03M	Metadata identifier resolves to a metadata record	000	Essential
A1	RDA-A1-03D	Data identifier resolves to a digital object	000	Essential
A1	RDA-A1-04M	Metadata is accessed through standardised protocol	000	Essential
A1	RDA-A1-04D	Data is accessible through standardised protocol	000	Essential
A1	RDA-A1-05D	$\label{eq:definition} \mbox{Data can be accessed automatically (i.e. by a computer program)}$	00	Important
A1.1	RDA-A1.1-01M	Metadata is accessible through a free access protocol	000	Essential
A1.1	RDA-A1.1-01D	Data is accessible through a free access protocol	00	Important
A1.2	RDA-A1.2-02D	Data is accessible through an access protocol that supports authentication and authorisation	•	Useful
A2	RDA-A2-01M	Metadata is guaranteed to remain available after data is no longer available	000	Essential
I1	RDA-I1-01M	Metadata uses knowledge representation expressed in standardised format	••	Important
I1	RDA-I1-01D	Data uses knowledge representation expressed in standardised format	••	Important
I1	RDA-I1-02M	Metadata uses machine-understandable knowledge representation	00	Important
I1	RDA-I1-02D	Data uses machine-understandable knowledge representation	00	Important
I2	RDA-I2-01M	Metadata uses FAIR-compliant vocabularies	00	Important







FAIR	ID	Indicator	Priority	
I2	RDA-I2-01D	Data uses FAIR-compliant vocabularies	•	Useful
13	RDA-I3-01M	Metadata includes references to other metadata	00	Important
13	RDA-I3-01D	Data includes references to other data	•	Useful
13	RDA-I3-02M	Metadata includes references to other data	•	Useful
13	RDA-I3-02D	Data includes qualified references to other data	•	Useful
13	RDA-I3-03M	Metadata includes qualified references to other metadata	00	Important
13	RDA-I3-04M	Metadata include qualified references to other data	•	Useful
R1	RDA-R1-01M	Plurality of accurate and relevant attributes are provided to allow reuse	000	Essential
R1.1	RDA-R1.1-01M	Metadata includes information about the licence under which the data can be reused	•••	Essential
R1.1	RDA-R1.1-02M	Metadata refers to a standard reuse licence	••	Important
R1.1	RDA-R1.1-03M	Metadata refers to a machine-understandable reuse licence	••	Important
R1.2	RDA-R1.2-01M	Metadata includes provenance information according to community-specific standards	••	Important
R1.2	RDA-R1.2-02M	Metadata includes provenance information according to a cross-community language	•	Useful
R1.3	RDA-R1.3-01M	Metadata complies with a community standard	000	Essential
R1.3	RDA-R1.3-01D	Data complies with a community standard	•••	Essential
R1.3	RDA-R1.3-02M	Metadata is expressed in compliance with a machine- understandable community standard	•••	Essential
R1.3	RDA-R1.3-02D	Data is expressed in compliance with a machine-understandable community standard	••	Important







