



# FAIR Principles

—

## Beyond FAIR

*Analysis of existing approaches  
v0.02*

*Existing approaches*



**F**indable

**A**ccessible

**I**nteroperable

**R**eusable

Non-FAIR  
Principles

Repository

Maintenance

Data quality

...

*Develop common metrics per facet*

# Non-FAIR Principles

Data repository

Curation and maintenance

Open data

Data quality

Others

# LEGEND

1	ANDS-NECTAR-RDS-FAIR data assessment tool	ARDC	<a href="#">[Link]</a>
2	DANS-Fairdat	DANS	<a href="#">[Link]</a>
3	DANS-Fair enough?	DANS	<a href="#">[Link]</a>
4	The CSIRO 5-star Data Rating tool	CSIRO	<a href="#">[Link]</a>
5	FAIR Metrics Questionnaire	The FAIR Metrics Group	<a href="#">[Link]</a>
6	Stewardship Maturity Mix	NOAA's CICS-NC, NOAA's NCDC	<a href="#">[Link]</a>
7	FAIR Evaluator	GO FAIR, LUMC CBGP, IDS, OeRC, IQSS	<a href="#">[Link]</a>
8	Data Stewardship Wizard	ELIXIR NL/CZ	<a href="#">[Link]</a>
9	Checklist for Evaluation of Dataset Fitness for Use	Assessment of Data Fitness for Use WG (WDS/RDA)	<a href="#">[Link]</a>
10	RDA-SHARC Evaluation	SHARC IG (RDA)	<a href="#">[Link]</a>
11	WMO-Wide Stewardship Maturity Matrix for Climate Data	The SMM-CD WG	<a href="#">[Link]</a>
12	Data Use and Services Maturity Matrix	The MM-Serv WG	<a href="#">[Link]</a>

## Principle

### # Facet

#### 1 Question

- Option #1
- Option #2
- Option #3

Potential Overlap

## X1 Data repository

1 What type of repository or registry is the metadata record in ?

- The data is not described in any repository
- Local institutional repository
- Domain-specific repository
- Generalist public repository
- Data is in one place but discoverable through several registries

3 Is the data repository you have chosen trustworthy

- No
- Yes

9 Is dataset located within a CoreTrustSeal-certified repository ?

- No
- Yes

9 Is dataset located within a World Data System or Data Seal of Approval certified repository?

- No
- Yes

9 Repository representative stipulates that structure, harmonization, completeness, and correctness of the dataset comports with typical data curation activities of

- No
- Somewhat
- Yes

10 Does the researcher use data repositories for the storage of data ?

- Never /NA
- If mandatory
- Sometimes
- Always

■ ■ ■

■ ■ ■

## 6 Preservability

Any storage location / Data only

Non-designated repository / Redundancy / Limited archiving metadata

Designated archive / Redundancy / Community-standard archiving metadata / Conforming to limited archiving standards

Conforming to community archiving standards

Archiving process / Performance controlled, measured and audited / Future archiving standard changes planned

## 11 Preservation

Any storage location; Data only; Data not backed up.

Non-designated repository; A backup copy of electronic data is made.

Designated archive; Basic retention policy publicly defined. Routine backups made, including offsite copy.

Previous + Conforming to community archiving standards. Comprehensive retention policy defined and implemented.

Previous + Archiving process performance controlled, measured, and audited; Future archiving standard changes planned.

## X2 Curation and maintenance

- 4** Updated - part of a regular data collection program or series, with clear maintenance arrangements and update schedule
  - one-time dataset
  - part of series - occasional/irregular update
  - part of series - regular scheduled updates
- 4** Curated - commitment to ensuring the data is available long term
  - Once-off dump, no ongoing commitment
  - Best effort, project website
  - Public or institutional repository (e.g. CKAN, GitHub)
  - Certified repository
- 11** Governance
  - Responsibility is not defined; No person is assigned.
  - Responsible entity is identified; Accountability and competency are not well-defined.
  - Responsibility/accountability and compliance mechanisms are defined; Good competency; Processes established conforming to community standards
  - Previous + Competency defined; Conforming to international standards; auditable.
  - Previous + Accountability and responsibility well defined and fully compliant with international standards; transparent; monitored and audited.

## X3 Open data

- 3** FAIR enough, but also Open? Will your data be published as open as possible and as protected as necessary?
  - No
  - Yes
- 8** Will you be working with the philosophy 'as open as possible' for your data?
  - No
  - Yes

## X4 Data quality

### 6 Data Quality Control / Monitoring

None or Sampling unknown or spotty / Analysis unknown or random in time  
 Sampling and analysis are regular in time and space / Limited productspecific metrics defined & implemented  
 Sampling and analysis are frequent and systematic but not automatic / Community metrics defined and partially implemented / Procedure documented and available online  
 Anomaly detection procedure well-documented and fully implemented using community metrics, automatic, tracked and reported / Limited quality monitoring metadata

### 6 Data Quality Assessment

Algorithm/method/model theoretical basis assessed (methods and results online)  
 Research product assessed (methods and results online)  
 Operational product assessed (methods and results online)  
 Quality metadata assessed / Limited quality assessment metadata  
 Assessment performed on a recurring basis Conforming to community quality metadata & standards External ranking

### 11 Quality Assurance & Control

F2

Ad hoc or no data quality assurance (QA) & control (QC) procedure or information unknown.  
 QA/QC procedure are defined, documented, and partially implemented.  
 QA/QC procedure are well-defined according to community best practices, documented and fully applied.  
 Previous + provision of error statistics published or tracked with results made available online and communicated to data providers; Procedure for user feedback, improvement  
 Previous + detailed analysis of errors and gaps at space-time unit level: (Station, grid-points, daily, monthly and or annual time-scale, etc.); QA/QC procedure monitored; Re

### 11 Quality Assessment

Product quality assessment not done or done internally and information not available.  
 Assessed by Principal Investigator (PI) or data producer; Assessment results available online.  
 Previous + Product validation and evaluation done by PI published in peer-reviewed journal.  
 Previous + Independent product validation and evaluation published in peer-reviewed journal.  
 Previous + The complete product provenance is captured and publicly available.



## X5 Others

8 Will data interpretation and modeling require significant compute infrastructure capacity?

No

Yes

8 Will you be doing (automated) knowledge discovery?

No

Yes

11 Uncertainty Analysis

Uncertainty estimates not available.

Uncertainty estimates presented without explanation.

Uncertainty estimates presented with partial explanation.

Full uncertainty budget available with all assumptions; Estimates of accuracy of trend available.

Full uncertainty assessment published in peer reviewed journal.