

Template to framework for FAIR data assessment and data management good practices

<u>INSTITUTION</u>	
<u>EVALUATOR</u>	
<u>DATASET SAMPLE TITLE</u> (significant name)	
<u>DATE</u>	
<u>VERSION Nº</u>	

FAIR	Indicators	Assessment			Comments
FAIR Code	Description	Level 1 (3)	Level 2 (2)	Level 3 (1)	Notes and examples
Administrative information (Findability)					
F1	ID (datasets have a unique identification number)				
F2	Funder (if applicable) + project name + project nº described (data precedence)				
F2, F5, R3	Context: dataset is described in its context? (Briefly summarize the type of study/studies to help others to understand the data)				
F2	Authors/Researchers identified + ORCID				
F2	Datasets Version (first, last)				
F1, F5	Data collection Policies (existing procedures, guidelines, etc.)				
Data Collection: which data is collected and how?					
F1, F2, R3	Data Type (experimental, observational, simulation, derived/compiled data)				
F3, I1	Standard formats use (.xml, .xls, .sql,)				
F4	Data volume (size)				

F1, F2, F5	Software (if applicable)				
F3, I1	Open Software				
F5, R3	Software documentation				
F1, F2, R1, R3	Data description (including any existing data or third parties)				
F1, F2	Standards or Methodologies for data collection described? Or other quality procedures?				
F1, F2, A4	Data location described (structure, naming conventions, folders, servers, repositories)				
Data Access					
F1, A1, A2, A7, R4	Access Conditions Specified?				
A2, A3, A7, R4, R5	Open Data Access?				
A2, A3, A7, A8, R4, R5	Data restrictions Access defined?				
A4	Are mentioned software tools needed to data access?				
F2, A4	Is proprietary software described?				
F4, A5, A6	Storage System defined?				
F1, A4	Does it specify where data and associated metadata and documentation or code are deposited?				
Data Interoperability					

F3, I1	Use of Open Formats				
F2, I2, I3, R1, R2	Use of standard metadata, vocabularies, or protocols for data description				
F2, I4, R3	In case of lack of standards, information about metadata or data description is known				
Data Reusability					
F2, A2, A3, A8, R1, R4, R5	Reusability defined?				
R2	Open License?				
F5, I4, R3	Documentation available (readme, data dictionaries)				
A4, A5, A6, R6	Data Preservation protocols defined (time, place, and responsibility)				
Results					

*28 parameters = 84 points (maximum level of Fairness), 56 (medium level), 28 (minimum = not optimal)

Notes to complete the assessment questionnaire.

Different levels of completeness are defined to evaluate the FAIR Data status of different data collection. These parameters also have correspondence to the data management plan EC guidelines.

Level 1: complete = 3 points (the parameter is described, or the answer is yes)

Level 2: medium = 2 points (parameter is not complete enough – we don't have the complete information)

Level 3: not complete / not exist = 1 point (the parameter it is not defined, or answer is not complete)

In case the description **is not applicable**, please code as: **0**

Basic instructions for parameters scoring

Administrative data (Findability):

- F1. General information about what and how datasets are collected or generated and its identification & organization for its findability.

Data collection (Findability)

- F2. Have a good data description (e.g. [metadata DC](#): title, creator, subject, description, publisher, contributor, date, type, format, identifier ([PIDs](#)), source, language, relation, coverage and rights)
- F3. Use of standard & open formats (e.g. .CSV instead of .xls)
- F4. Data volume (size determine the storage system)
- F5. Documentation: is data well documented for its reproducibility and reuse? What documentation and metadata will accompany the data?

Data Access (Accessibility):

- A1. How will you manage access and security?
- A2. How will you manage ethical issues?
- A3. How will you manage copyright and Intellectual Property Rights (IPR) issues?
- A4. How will the data be stored and backed up during the research?
- A5. Which data should be retained, shared, and/or preserved?
- A6. What is the long-term preservation plan for the dataset?
- A7. How will you share the data?
- A8. Are any restrictions on data sharing required?

Data Interoperability:

- I1. Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organizations, countries, etc. (i.e. adhering to standards for formats) facilitating re-combinations with different datasets from different origins)?
- I2. What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?
- I3. Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?
- I4. In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Data Reusability:

- R1. Are (Meta) data (richly described? (use of community standards)?
- R2. Are (Meta)data released with a clear and accessible data usage license?
- R3. Are the data well documented (readme files, data dictionaries...)?
- R4. When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
- R5. Are the data produced and/or used in the project useable by third parties, after the end of the project? If the re-use of some data is restricted, explain why.

R6. Do you have a data preservation policy? How will the data be stored after the end of the project (data sustainability)? How long can the data be (re)used?