FAIRSFAIR
Fostering Fair Data Practices in Europe

How to improve the FAIRness of your data

Hands-on exercise

Maaike Verburg & Marjan Grootveld (DANS)

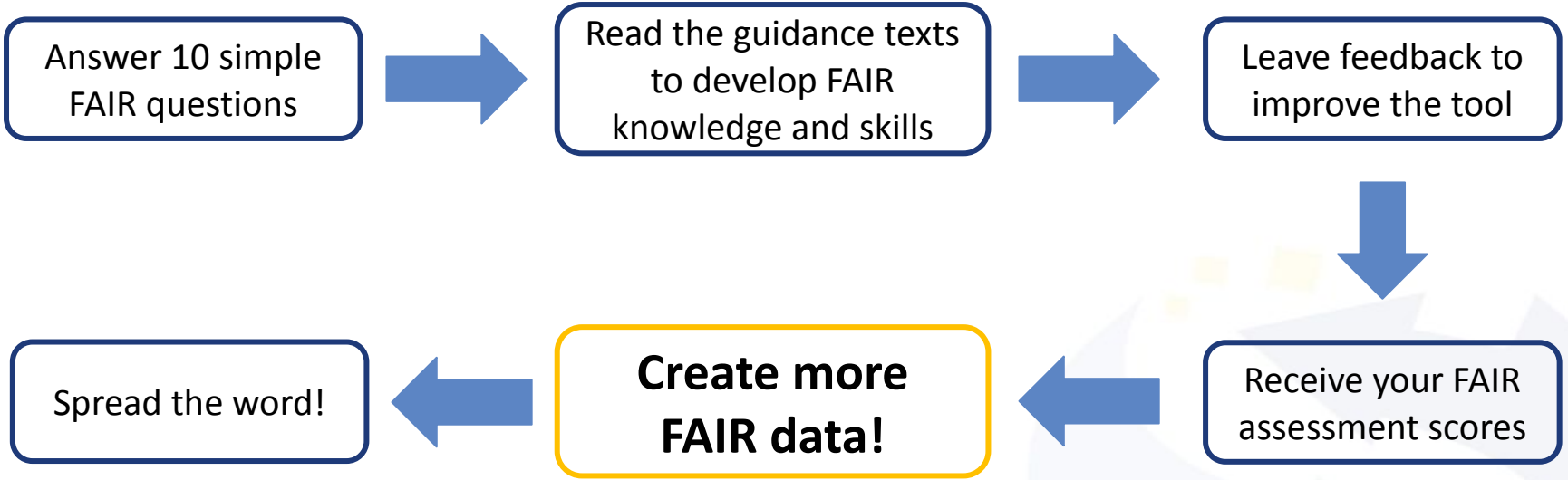
@FAIRsFAIR_EU | @DANS_knaw_nwo | #FAIRAwareTool

FAIRsFAIR Roadshow Lithuania
10 November 2021

FAIRsFAIR "Fostering FAIR Data Practices in Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558

FAIRSFAIR
Fostering Fair Data Practices in Europe

FAIR Aware



Exercise: FAIR-Aware for DMP writing

- Plan for FAIR data - Use FAIR-Aware guidance for DMP writing!

Where in FAIR-Aware can you find help for the following DMP elements?

1. In what **format** will you collect / archive your data?
2. What **metadata** will be provided to help others **discover** the data?
3. Who will be **responsible** for data management (specifically **archiving**)?

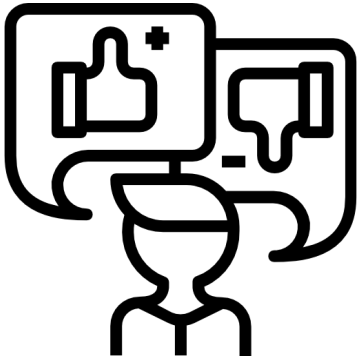
- https://fairaware.dans.knaw.nl/ | Science Europe DMP template (Link)

Feedback

How useful do you find FAIR-Aware for supporting FAIR data?

Would you use it in the future or recommend it?

Put your star rating (1-5) and other feedback in the chat!



FAIRSFAIR

Fostering Fair Data Practices in Europe

Thank you!

Now enjoy a 10 min break

FAIRsFAIR "Fostering FAIR Data Practices in Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558

FAIRSFAIR

Fostering Fair Data Practices in Europe

Answers

FAIRsFAIR "Fostering FAIR Data Practices in Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558

In what format will you collect / archive your data?

REUSABLE

7. Are you aware that provenance information about the collection and/or generation of data should be included in the metadata? ○ Yes ○ No

8. Are you aware that metadata describing your data should follow the specifications of a community-endorsed standard? ○ Yes ○ No

9. Are you aware that data should be deposited preferably in a file format that is open – to support reuse – and supported by the repository for long-term preservation? ○ Yes ○ No

10. Are you aware that maintaining your dataset FAIR over time requires professional data curation and preservation? ○ Yes ○ No

9. Are you aware that data should be deposited preferably in a file format that is open – to support reuse – and supported by the repository for long-term preservation?

File formats refer to methods for encoding digital information. If there is no open format available, then you may also use a proprietary format. For example, CSV for tabular data, NetCDF for multidimensional data and GeoTIFF for raster imagery.

Data should be made available in a recommended file format that is accepted by the research community to enable data sharing and reuse, and by the repository to enable long-term preservation. Repositories normally have different recommended standard file formats based on data types (for example: UK Data Service) and ISO/TR 22299 Document management - Digital file format recommendations for long-term storage.

Recommended formats are widely used and supported by the most common software and tools ensuring that your data can be read in the future, but they will also help increase the reusability and interoperability. Using recommended formats enables data to be loaded directly into the software and tools used for data analysis. It makes it possible to easily integrate your data with other data using the same format. The use of recommended formats will allow migration of the format to a newer one, in case a preferred format becomes outdated.

Close

What metadata will be provided to help others discover the data?

FINDABLE

1. Are you aware that a dataset should be assigned a globally unique persistent and resolvable identifier when deposited with a data repository? ○ Yes ○ No

2. Are you aware that when you deposit a dataset with a repository, you will need to provide some details (known as discovery metadata) in order to make the data findable, understandable and reusable to others? ○ Yes ○ No

3. Are you aware that the repository providing access to your dataset should make the metadata describing your datasets available in a format readable by machines as well as humans? ○ Yes ○ No

2. Are you aware that when you deposit a dataset with a repository, you will need to provide some details (known as discovery metadata) in order to make the data findable, understandable and reusable to others?

The necessary metadata is descriptive information about the data object (e.g. creator, title, publisher, creation and publication date, summary and keywords describing the data) provided by the researcher when documenting the dataset for deposit in a data repository.

Since the metadata required depends on the type of users and the intended applications, the focus here is on ensuring the minimum descriptive information needed to enable data discovery and citation by potential users - even those from other domains. The metadata should include the unique, persistent and resolvable identifier (PID) for the data so that users can discover and access the data. In cases where the data cannot be shared openly for ethical, legal or commercial reasons, the metadata should make clear how a potential reuser can request legitimate access.

Data content (relates to Reusability in the FAIR principle)
The content of the dataset should be specified in the metadata. It should be an accurate reflection of the actual data deposited. Examples of the properties specifying data content are: resource type (e.g., data or a collection of data), variable(s) measured or observed, method, data format and size. Ideally, semantic vocabularies should be used to describe data content (e.g., variable) to support interdisciplinary reuse. For example, variables should be described by providing a link with a PID to a resource (such as a landing page) containing a structured description of the concept, references and links to other relevant concepts. The access to the linked concepts should be available through the query interface (e.g., using SPARQL or GraphQL).

Links to other research outputs (relates to Interoperability in the FAIR Principles)

Who will be responsible for data management (specifically archiving)?

REUSABLE

7. Are you aware that provenance information about the collection and/or generation of data should be included in the metadata? ○ Yes ○ No

8. Are you aware that metadata describing your data should follow the specifications of a community-endorsed standard? ○ Yes ○ No

9. Are you aware that data should be deposited preferably in a file format that is open – to support reuse – and supported by the repository for long-term preservation? ○ Yes ○ No

10. Are you aware that maintaining your dataset FAIR over time requires professional data curation and preservation? ○ Yes ○ No

10. Are you aware that maintaining your dataset FAIR over time requires professional data curation and preservation?

Data curation is the active and ongoing management of data from the point of creation to ensure that it is fit for contemporary purpose and available for discovery and reuse. Likewise, digital preservation refers to the series of managed activities necessary to ensure continued access to and reusability of digital materials for as long as necessary (i.e FAIR).

Professional data curation and preservation require people, skills and technology. They are central to the mission of several academic institutions, such as universities, libraries, data repositories and archives. Specifically, Trustworthy Digital Repositories (TDRs) play a critical role in both making and preserving data FAIR over time. TDRs provide support and take responsibility for the curation and preservation of data with different levels of FAIRness. Digital repositories demonstrate their trustworthiness, for example, through certification with community-endorsed standards such as the CoreTrustSeal, DIN31644/NESTOR and ISO163638. Trustworthy Data Repositories can be found on registries. For example, CoreTrustSeal certified repositories can be filtered on Re3data.

Want to know more?

Close