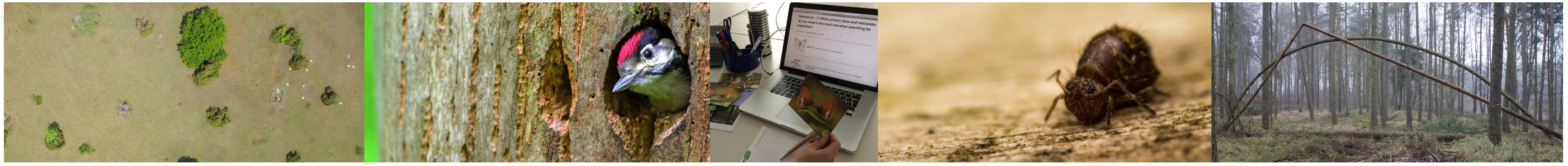


Towards FAIR data and repository within the Biodiversity Exploratories



BIODIVERSITY EXPLORATORIES

Question

The Biodiversity Exploratories Information System (BExIS) acts as centralized data management platform for all datasets providing many additional features to support scientist throughout the whole data life cycle.

Do BExIS and the provided data follow the FAIR data principles?

<https://www.bexis.uni-jena.de/>

Method

Currently no standard or agreed upon measure or metric exists to assess the FAIRness (compliance to principles) of data or repositories. Different initiatives are working on this topic. Few organizations have developed self-assessment tools on their interpretation of the FAIR principles. Thus they do not scale to a standard but are easy to understand and useful for engaging the research community to improve practice [1].

We chose a FAIR Data self-assessment tool developed by the Australian Research Data Commons (ARDC). Answers to questions related to the principles will give a green bar indicator for each section and an overall 'FAIRness' indicator..

Results and Outlook

Parts of the repository and the open data already fulfill some principles. The **Findability** needs small adaptations to enhance discovery by several registries. The **Accessibility** is in a good shape already. Larger improvements are necessary concerning **Interoperability**. This is mainly due the insufficient interlinking of elements to community agreed formats and vocabularies. **Reusability** can be improved by including more provenance information.

FAIR Data Principles

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 open, free, 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:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (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 detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

<https://doi.org/10.1038/sdata.2016.18>

[1] Final report and action plan from the European Commission expert group on FAIR data <https://doi.org/10.2777/1524>

Self-assessment

Findable	
Does the dataset have any identifiers assigned?	Globally Unique, citable and persistent (e.g. DOI, PURL, ARK or Handle)
Is the dataset identifier included in all metadata records/files describing the data?	Yes
How is the data described with metadata?	Comprehensively (see suggestion) using a recognised formal machine-r
What type of repository or registry is the metadata record in?	Domain-specific repository
	
Accessible	
How accessible is the data?	Publicly accessible
Is the data available online without requiring specialised protocols or tools once access has been approved?	Standard web service API (e.g. OGC)
Will the metadata record be available even if the data is no longer available?	Yes
	
Interoperable	
What (file) format(s) is the data available in?	In a structured, open standard, machine-readable format
What best describes the types of vocabularies/ontologies/tagging schemas used to define the data elements?	No standards have been applied in the description of data elements
How is the metadata linked to other data and metadata (to enhance context and clearly indicate relationships)?	There are no links to other metadata
	
Reusable	
Which of the following best describes the license/usage rights attached to the data?	Standard machine-readable license (e.g. Creative Commons)
How much provenance information has been captured to facilitate data reuse?	Partially recorded
	

Total across F.A.I.R.



Via ARDC FAIR Data self-assessment tool. <https://www.andis-nectar-rds.org.au/fair-tool>