FAIR self assessment for project: AusTraits

Completed 1/8/2019, updated 20/12/2019

	Questions for each FAIR component ↓	Answer options: Increasingly FAIR -->				
FINDABLE						
Q1	Does the dataset have any identifiers assigned?	No identifier	Local identifier	Web address (URL)	Globally unique, citable and persistent identifier (e.g. DOI, PURL, or Handle)	
A1	Start of project	Data in private GitHub repository				
	End of project				DOI for each release of AusTraits	
	Two years time			Web addresses for both github repository and possibly something like austraits.org	as above	
Q2	Is the identifier included in all metadata records or metadata files describing the data?	No	Yes			
A2	Start of project	No identifier				
	End of project				Yes	
	Two years time				as above	
Q3	Is the data described by a metadata record?	The data is not described	Brief title and description	Brief title and description, and multiple other fields filled out, albeit briefly.	Comprehensively (a min metadata template will be provided) using a formal machine-readable metadata schema.	
A3	Start of project	No metadata record				
	End of project				Based on standard of zenodo.org	
	Two years time				as above	
Q4	What type of repository or registry is the metadata record in?	The data is not described in any registry or repository	Local institutional repository	Domain-specific repository	Generalist public repository	Data is in one place but discoverable through several places (i.e. other registries, RDA, Google Data Search)
A4	Start of project	No metadata record				
	End of project				Based on standard of zenodo.org	
	Two years time					Upon publication of our paper, a copy of metadata record will be harvested into Research Data Australia
ACCESSIBLE						
Q5	How accessible is the data? Note: The access method(s) must be explicitly stated in the metadata record, e.g. if any authentication is needed, or there are any restrictions to access.	No metadata record	Access to metadata only	Unspecified access conditions e.g. "contact the data custodian to discuss access"	Embargoed access after a specified date; or A deidentified version of the data is publicly accessible	Fully accessible public, or to persons who meet and follow explicitly stated conditions and processes, e.g. ethics approval for sensitive data
A5	Start of project					
	End of project				Access is to data is emabrgoed until our paper descrobong the data is published	
	Two years time					Like the GitHub releases feature, archived in Zenodo
Q6	Is the data available online without requiring specialised protocols or tools once access has been approved?	No access to data	By individual arrangement	File download from online location	Non-standard web service (e.g. OpenAPI/Swagger/informal API)	Standard web service API (e.g. OGC)
A6	Start of project					
	End of project					Based on API of zenodo.org
	Two years time					as above
Q7	Does the repository/registry agree to maintain the persistence of the metadata record, even if the data product is no longer available?	No (or not applicable, if no metadata record exists)	Unsure	Yes		
A7	Start of project	No record, also unsure of Uni repository policy				
	End of project			Zenodo established for long-term archiving		
	Two years time			as above		
INTEROPERABLE						
Q8	Are the data available in (an) open (file) format(s)?	Data are mostly available only in a proprietary format	Data are available in an open format	Data are available in an open, documented, widely-used standard format (i.e. NetCDF, CSV, JSON, XML, etc)		
A8	Start of project			Data are in csv files, documentation in YAML		
	End of project			as above		
	Two years time			as above		
Q9	Are the data is machine readable?	The data are unstructured	The data are structured and machine-readable (i.e. csv, JSON, XML, RDF, database files, etc)			
A9	Start of project		Data are in csv files			
	End of project		as above			
	Two years time		as above			
Q10	What best describes the types of vocabularies/ontologies/tagging schemas used to define the data elements?	Data elements are not described (i.e. fields or objects are labelled with codes or not at all)	Data elements are described (so that a human user can correctly interpret the data), but no standards have been used in the description	Recognised standards have been used in the description of data elements, but no published vocabularies with resolvable URIs are used	Published vocabularies using resolvable identifiers linking to explanations are used, so that the data can be read and understood by machines as well as humans.	Published vocabularies using **persistent** resolvable identifiers linking to explanations are used, so that the data can be read and understood by machines as well as humans.
A10	Start of project		Columns are labelled with names such as temperature, and units are given	Some standardisation is being done on the desciption		
	End of project		as above	as above		
	Two years time			We hope that in the future a standard will be published against which we will link		
Q11	How is the relationship to other data and resources (e.g. related datasets, services, publications, etc) described in the metadata, to provide context around the data?	There are no links to other metadata or data	The metadata record includes URI links to related metadata, data and definitions	Qualified links to other resources are recorded in a machine readable format, e.g. a linked data format such as RDF		
A11	Start of project		Clear documenation provided with some links to related publications (provenance of data), but no metadata record, so no qualified links			
	End of project		as above			
	Two years time			We hope that in the future a standard will be published against which we will link		
REUSABLE						
Q12	Which of the following best describes the license (usage rights) attached to the data?	No license is applied	Non-standard license applied, without a license deed URL encoded in a machine-readable format (e.g. RDF/XML) in the metadata record	Non-standard license applied, WITH the license deed URL encoded in a machine-readable format (e.g. RDF/XML) in the metadata record	Standard license applied (e.g. Creative Commons), without a license deed URL encoded in a machine-readable format (e.g. RDF/XML) in the metadata record	Standard license applied (e.g. Creative Commons), WITH the license deed URL encoded in a machine-readable format (e.g. RDF/XML) in the metadata record
A12	Start of project					
	End of project					Release with CC-BY
	Two years time					Release with CC-BY
Q13	How much provenance information has been captured to facilitate data reuse? i.e. project objectives, data generation/collection (including from external sources) and processing workflows.	No provenance information is recorded	Partially recorded	Comprehensively recorded in a text format (i.e. TXT or PDF)	Comprehensively recorded in a machine readable format (i.e. in metadata record's schema or PROV, or in RDF, JSON, NetCDF, XML, etc)	
A13	Start of project			Extensive documentation in YAML & HTM		
	End of project			Source data and code will be available in archived containers enabling dataset to be reproduced from raw data		
	Two years time			as above		