FAIR self assessment for project: Bringing Long-Tail Microscopy and Characterisation Data into the Light

Completed 8/10/2019

	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 is organised by ProjectID. A Project is a collection of files (Datafiles) grouped by (Dataset). Projects, Datasets and Datafiles each have a local identifier (MyTardis link) specific to the repository			
	End of project		Data is organised by ProjectID. A Project is a collection of files (Datafiles) grouped by (Dataset). Projects, Datasets and Datafiles each have a local identifier (MyTardis link) specific to the repository			
	Two years time		Data is organised by ProjectID. A Project is a collection of files (Datafiles) grouped by (Dataset). Projects, Datasets and Datafiles each have a local identifier (MyTardis link) specific to the repository	Yes for published Datasets	DOI for user-published Datasets via MyTardis publication wizard (http://www.mytardis.org/features)	
Q2	Is the identifier included in all metadata records or metadata files describing the data?	No	Yes			
A2	Start of project		Record in repository database associated with Dataset link			
	End of project		Record in repository database associated with Dataset link			
	Two years time		Record in repository database associated with Dataset link			
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		Project - string description; Dataset - NIF certification; Datafile - Bioformats			
	End of project		Project - string description; Dataset - NIF certification; Datafile - Bioformats			
	Two years time				NIF/MA agreed metadata schemas for Project, Dataset and Datafile	
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			*Metadata record in MyTardis repository; example https://trudat.cmca.uwa.edu.au/dataset/103*		
	End of project			*Metadata record in MyTardis repository; example https://trudat.cmca.uwa.edu.au/dataset/103*		
	Two years time					*Metadata record in MyTardis repository and RDA*
	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					Data is only accessible to users who have authenticated via AAF and have been assigned access to the Project containing the datasets.
	End of project					Data is only accessible to users who have authenticated via AAF and have been assigned access to the Project containing the datasets.
	Two years time					Data is accessible to either: (i) users who have authenticated via AAF and have been assigned access to the Project containing the datasets, or (ii) public access for published datasets with specified license
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			Files downloadable via web interface		
	End of project			Files downloadable via web interface		
	Two years time			Files downloadable via web interface		
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			Data and metadata are archived to UWA's Institution Research Data Store		
	End of project			Data and metadata are archived to UWA's Institution Research Data Store		
	Two years time			Data and metadata are archived to UWA's Institution Research Data Store		
	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		MRI data available in both proprietary and one or more open formats (e.g. DICOM, MINC, NIFTI)			
	End of project		MRI data available in both proprietary and one or more open formats (e.g. DICOM, MINC, NIFTI)			
	Two years time		Yes for all formats, perhaps enabled via a fileconversion service			
Q9	Are the data machine readable?	The data are unstructured	The data are structured and machinereadable (i.e. csv, JSON, XML, RDF, database files, etc)			

A9	Start of project	Downloaded data comprises the datafiles only but no metadata				
	End of project	Downloaded data comprises the datafiles only but no metadata				
	Two years time		Downloaded data is structured and machinereadable; Data Crate / BagIt			
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		Objects are labelled as Project/Experiment, Dataset or Datafile; Internal schemas for Dataset and Datafile			
	End of project		Objects are labelled as Project/Experiment, Dataset or Datafile; Internal schemas for Dataset and Datafile			
	Two years time			For Datasets and Datafiles		
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		Each Dataset has a link to an instrument record in RDA			
	End of project		Each Dataset has a link to an instrument record in RDA			
	Two years time		Each Dataset has a link to an instrument record in RDA			
	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 machinereadable format (e.g. RDF/XML) in the metadata record	Non-standard license applied, WITH the license deed URL encoded in a machinereadable 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	No license is applied				
	End of project	No license is applied				
	Two years time					For user-published datasets

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		MRI data provides some provenance information with respect to the subject/study, scan parameters, and processing performed (derived data)			
	End of project		MRI data provides some provenance information with respect to the subject/study, scan parameters, and processing performed (derived data)			
	Two years time				For all instrument data	