

FAIR self assessment for project: MyTardis						
Completed 26/09/2019						
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				Each dataset has a unique web address (URL), and for public datasets, MyTardis can mint a DOI, e.g. https://dx.doi.org/10.4225/52/563BD4A36013F	
	End of project				Each dataset has a unique web address (URL), and for public datasets, MyTardis can mint a DOI, e.g. https://dx.doi.org/10.4225/52/563BD4A36013F	
	Two years time				In two years time, MyTardis will offer a choice of multiple DOI minting services, supporting more than just https://mod.monash.edu	
Q2	Is the identifier included in all metadata records or metadata files describing the data?	No	Yes			
A2	Start of project		Yes. Metadata records are linked to the dataset ID via foreign keys and the dataset's unique URL can be constructed from the dataset ID.			
	End of project		Yes. Metadata records are linked to the dataset ID via foreign keys and the dataset's unique URL can be constructed from the dataset ID.			
	Two years time		Yes. Metadata records are linked to the dataset ID via foreign keys and the dataset's unique URL can be constructed from the dataset ID.			
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			MyTardis provides the capability to record as many metadata fields as desired, but it is not enforced. Each dataset requires a title. Additional metadata fields can be made user-editable or they can be automatically extracted from file(s) captured from the instrument PC. Whilst MyTardis has the capability to provide RIF-CS (OAI-PMH) feeds for public data, MyTardis does not force users to publish their data.		
	End of project			MyTardis provides the capability to record as many metadata fields as desired, but it is not enforced. Each dataset requires a title. Additional metadata fields can be made user-editable or they can be automatically extracted from file(s) captured from the instrument PC. Whilst MyTardis has the capability to provide RIF-CS (OAI-PMH) feeds for public data, MyTardis does not force users to publish their data.		

	Two years time				An improved user interface in MyTardis will make it easier to record metadata, giving researchers an incentive to describe their data in more detail. The metadata records will be indexed by an improved Elasticsearch backend making them easy to find. MyTardis will offer the ability to add Domain-Specific predefined Schema to facilitate and standardize the process of meta-data publishing.	
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 is stored in a relational database cluster, and is accessible to researchers from multiple institutions in a web interface using Australian Access Federation authentication. Public datasets are discoverable through Google Search, and metadata can be made available using OAI-PMH feeds.			
	End of project		Metadata is stored in a relational database cluster, and is accessible to researchers from multiple institutions in a web interface using Australian Access Federation authentication. Public datasets are discoverable through Google Search, and metadata can be made available using OAI-PMH feeds.			
	Two years time					In addition to the existing XML-based OAI-PMH feeds, MyTardis will make metadata available via multiple APIs (REST and GraphQL) offering multiple serialization formats, including XML, JSON and YAML. MyTardis will work on enabling data discovery using Search Engines (e.g. Google Data Search) as an option for published datasets.
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					MyTardis provides mechanisms to store data privately, or to make data public or to make a dataset's metadata public (without making the data public). Newly created datasets are private by default, but researchers can share them with collaborators or publish them.
	End of project					MyTardis provides mechanisms to store data privately, or to make data public or to make a dataset's metadata public (without making the data public). Newly created datasets are private by default, but researchers can share them with collaborators or publish them.
	Two years time					MyTardis provides mechanisms to store data privately, or to make data public or to make a dataset's metadata public (without making the data public). Newly created datasets are private by default, but researchers can share them with collaborators or publish them.

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				MyTardis provides a web interface (and REST API) for downloading data. Individual files can be downloaded, or whole datasets can be downloaded in TAR format or by using an SFTP client.	
	End of project				MyTardis provides a web interface (and REST API) for downloading data. Individual files can be downloaded, or whole datasets can be downloaded in TAR format or by using an SFTP client.	
	Two years time				MyTardis will introduce a new REST API version (v2) based on Django REST Framework in addition to the existing REST API (v1) which will help to future-proof MyTardis, given that the technology behind v1 (Django Tastypie) is now in maintenance mode. API v1 will continue to be supported for as long as needed by clients such as MyData. In addition to the HATEOAS REST APIs, MyTardis will provide a GraphQL interface to allow clients to request only the fields they want from the API.	
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			Yes. The MyTardis database (including metadata) will not be lost in the event that the data product is no longer available from the file storage backend.		
	End of project			Yes. The MyTardis database (including metadata) will not be lost in the event that the data product is no longer available from the file storage backend.		
	Two years time			Yes. The MyTardis database (including metadata) will not be lost in the event that the data product is no longer available from the file storage backend.		
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 is stored in open, documented, widely-used standard formats. The most common file type stored on MyTardis is TIFF. There are many other commonly used file types found in MyTardis deployments, including DICOM and MRC (for CryoEM).		
	End of project			Data is stored in open, documented, widely-used standard formats. The most common file type stored on MyTardis is TIFF. There are many other commonly used file types found in MyTardis deployments, including DICOM and MRC (for CryoEM).		

	Two years time			Data is stored in open, documented, widely-used standard formats. The most common file type stored on MyTardis is TIFF. There are many other commonly used file types found in MyTardis deployments, including DICOM and MRC (for CryoEM).		
Q9	Are the data 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		Files stored on MyTardis are in a machine readable format, with TIFF being the most common file format. The Bio-Formats tools are used to extract metadata and thumbnail images from the image files recorded with MyTardis.			
	End of project		Files stored on MyTardis are in a machine readable format, with TIFF being the most common file format. The Bio-Formats tools are used to extract metadata and thumbnail images from the image files recorded with MyTardis.			
	Two years time		Files stored on MyTardis are in a machine readable format, with TIFF being the most common file format. The Bio-Formats tools are used to extract metadata and thumbnail images from the image files recorded with MyTardis.			
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				MyTardis deployments can optionally enable an OAI-PMH feed which contains machine-readable (and human-readable) metadata, as available in https://store.synchrotron.org.au/apps/oaipmh/?verb=ListRecords&metadataPrefix=rif	
	End of project				MyTardis deployments can optionally enable an OAI-PMH feed which contains machine-readable (and human-readable) metadata, as available in https://store.synchrotron.org.au/apps/oaipmh/?verb=ListRecords&metadataPrefix=rif	
	Two years time				MyTardis deployments can optionally enable an OAI-PMH feed which contains machine-readable (and human-readable) metadata, as available in https://store.synchrotron.org.au/apps/oaipmh/?verb=ListRecords&metadataPrefix=rif	
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		MyTardis's metadata system makes it possible to add links to published articles and other references relating to a public dataset, e.g. a PDB record for a protein structure. Some example links can be found here: https://store.synchrotron.org.au/experiment/view/1037/#metadata			

	End of project		MyTardis's metadata system makes it possible to add links to published articles and other references relating to a public dataset, e.g. a PDB record for a protein structure. Some example links can be found here: https://store.synchrotron.org.au/experiment/view/1037/#metadata			
	Two years time		MyTardis's metadata system makes it possible to add links to published articles and other references relating to a public dataset, e.g. a PDB record for a protein structure. Some example links can be found here: https://store.synchrotron.org.au/experiment/view/1037/#metadata MyTardis will add the ability to link related/derived datasets (e.g. linking the different formats of the same datafile), link related publications, link to Lab notebooks, and link to Quality control datasets.			
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					Standard licenses (Creative Commons) are used for public datasets. MyTardis deployments can optionally enable an OAI-PMH feed which contains the license deed URL encoded in a machine-readable format, as included in https://store.synchrotron.org.au/apps/oaipmh/?verb=ListRecords&metadataPrefix=rif
	End of project					Standard licenses (Creative Commons) are used for public datasets. MyTardis deployments can optionally enable an OAI-PMH feed which contains the license deed URL encoded in a machine-readable format, as included in https://store.synchrotron.org.au/apps/oaipmh/?verb=ListRecords&metadataPrefix=rif
	Two years time					Standard licenses (Creative Commons) are used for public datasets. MyTardis deployments can optionally enable an OAI-PMH feed which contains the license deed URL encoded in a machine-readable format, as included in https://store.synchrotron.org.au/apps/oaipmh/?verb=ListRecords&metadataPrefix=rif
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		MyTardis offers a flexible system for recording metadata which can include provenance information. MyTardis does not force researchers to enter this information.			
	End of project		MyTardis offers a flexible system for recording metadata which can include provenance information. MyTardis does not force researchers to enter this information.			

	Two years time		An improved user interface in MyTardis will make it easier to record metadata including provenance information, giving researchers an incentive to describe their data in more detail. Linking with Lab Notebooks and publications will facilitate the recording of provenance information.		
--	----------------	--	---	--	--