

FAIR self assessment for project: Percutaneous Coronary Intervention (PCI) Outcomes Registry - Baseline Data

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	No				
	End of project	As this is a pilot project, there is no identifier assigned to the dataset				
	Two years time				The dataset will be held in a central repository with the University of New South Wales (UNSW), and associated with a DOI for future researchers to find and use the dataset	
Q2	Is the identifier included in all metadata records or metadata files describing the data?	No	Yes			
A2	Start of project	No				
	End of project	As this is a pilot project, there is no metadata record or file to describe the data				
	Two years time		Once the central registry is established at UNSW we intend to create a metadata record for the dataset			
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	Not applicable, no metadata record				
	End of project	No metadata record				
	Two years time				Once the central registry is established at UNSW we intend to create a machine readable metadata schema that is accessible to future researchers	
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	Not applicable, no metadata record				
	End of project	No metadata record				
	Two years time					Once the central registry is established at UNSW we intend for the metadata to be located within a UNSW repository that is accessible via a google search
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	The dataset did not exist prior to this project				

Q6	End of project	The dataset is still being created and is not yet accessible				
	Two years time			Researchers will be directed to contact the data custodian to discuss access. Specific criteria must be met to grant access including ethics approval and dataset governance rules.		
	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	The dataset did not exist prior to this project				
	End of project	No open online access due to the sensitive nature of the data contained				
	Two years time	No open online access due to the sensitive nature of the data contained				
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	The dataset did not exist prior to this project				
	End of project	Not applicable, no metadata record				
	Two years time			Yes		
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	The dataset did not exist prior to this project				
	End of project			Data are available in an open, documented, widely-used standard format (CSV)		
	Two years time			Data are available in an open, documented, widely-used standard format (CSV)		
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	The dataset did not exist prior to this project				
	End of project		The data are structured and machine-readable (CSV)			
	Two years time		The data are structured and machine-readable (CSV)			
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	The dataset did not exist prior to this project				

	End of project			Standardised, published definitions have been used for the data elements. A data dictionary with a comprehensive description of each data point is available but not published.		
	Two years time					We intend for the dataset to have persistent resolvable identifiers linked to explanations, understandable by humans and machines
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	No metadata record				
	End of project	No metadata record				
	Two years time		We intend to create metadata record that links to related metadata from similar datasets such as the Victorian Cardiac Outcomes Registry (VCOR)			
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	The dataset did not exist prior to this project				
	End of project	As a pilot project, this does not apply. No license is applied.				
	Two years time		Non-standard license applied, without a license deed URL encoded in a machine-readable format (e.g. RDF/XML) in the metadata record. An attribution will be required in the form a credit line and citation.			
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	The dataset did not exist prior to this project				
	End of project	Not applicable, no metadata record				
	Two years time				We intend for comprehensive provenance information to be recorded for the dataset that is in a machine readable format	