

**FAIR self assessment for project:
Roadmap to integrate Indigenous genome assemblies into a national in-instance of the international human reference genome resource**

Completed ..08/10/2019

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		Yes			
	End of project					
	Two years time					
Q2	Is the identifier included in all metadata records or metadata files describing the data?	No				
A2	Start of project		yes - embedded in file or code			
	End of project		yes - embedded in file or code			
	Two years time		In the future - potentially 5 years			
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.	Generated automatically
A3	Start of project	The data is not described				
	End of project	The data is not described				
	Two years time	Implementation of road map will enable data description				
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	The data is not described in any registry or repository				
	End of project	The data is not described in any registry or repository				
	Two years time					Implementation of road map will enable description in national and international cohesive repository
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	No metadata record				
	End of project	No metadata record				
	Two years time					Implementation of road map will provide a consistent mechanism for national and international access of the data
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	Access and use is driven by ethics and consent process	Access and use is driven by ethics and consent process - individual arrangement - project specific			

	End of project	No access to data	Access and use is driven by ethics and consent process - individual arrangement - project specific			
	Two years time		Access determined by research participants and use is driven by ethics and consent process - individual arrangement - project specific			
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		
	End of project			Yes		
	Two years time			Yes - current international policy		
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 available in an open, documented, widely-used standard format (i.e. NetCDF, CSV, JSON, XML, etc)		
	End of project			Data are available in an open, documented, widely-used standard format (i.e. NetCDF, CSV, JSON, XML, etc)		
	Two years time			Will adapt to emerging standards as per roadmap		
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			Yes		
	End of project			Yes		
	Two years time			Yes		
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			Recognised standards have been used in the description of data elements, but no published vocabularies with resolvable URIs are used		
	End of project			Recognised standards have been used in the description of data elements, but no published vocabularies with resolvable URIs are used		
	Two years time				5 year aim - Recognised standards have been used in the description of data elements, but no published vocabularies with resolvable URIs are used	
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	There are no links to other metadata or data				

	End of project					Nascent approach to national and international schema
	Two years time					5 years - clarity around and adoption of above
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	Reuse is driven by ethics and consent process				
	End of project					Reuse is driven by ethics and consent process
	Two years time					Reuse is driven by ethics and consent process
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	As per national and international community requirements				
	End of project					As per national and international community requirements
	Two years time					As technology evolves, the national and international community requirements extend