Recommendations for FAIR Evaluation Services, based upon the FAIR Maturity Evaluation Service[12]

The FAIR Maturity Evaluation Service deploys 22 automated tests confirming FAIR properties for the majority of the FAIR Principles. Not all FAIR Principles are covered by the current version of Evaluator. Explicit tests for the following three principles do not exist as of yet (cf. FAIR recommendations intro doc):

- R1.2. (Meta)data to be associated with detailed provenance

- R1.3. (Meta)data to meet domain relevant community standards.

An overview of the 22 tests:

TEST	TEST HEADER	FAIR Principle	TEST DESCRIPTION	Recommendation on TEST-FAILURE	Detailed information relevant to recommended actions in case of TEST-FAILURE
1	UNIQUE IDENTIFIER (METADATA)	F1	Test if the metadata resource uses unique identifier(s) (Test for the existence of URLs pointing to DOIs or FAIRSharing)	Provide the metadata using Unique Identifier(s) – such as DOIs or FAIRSharing	For a digital object (PID, metadata and data) the assumption is that the object is identified using a Globally Unique Identifier (GUID). The GUID should resolve to a 'landing page' that contains two elements; metadata and data, each with its own identifier.

[1] Wilkinson, M.D., Dumontier, M., Sansone, S. *et al.* Evaluating FAIR maturity through a scalable, automated, community-governed framework. *Sci Data* **6,** 174 (2019). https://doi.org/10.1038/s41597-019-0184-5

[2] On line: https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/

2	IDENTIFIER PERSISTENCY (METADATA)	F1	Test for the existence of a persistence policy (Test for the persistence schemas at FAIRSharing or other)	Locate or create a link to an explicit persistence policy, ideally in machine-actionable format	This indicator tests if the unique identifier of the metadata resource is likely to be persistent. Known schema are registered in FAIRSharing (https://fairsharing.org/standards/?q=&selected_facets=type_exact:identifier%20schema). For URLs that don't follow a schema in FAIRSharing we test known URL persistence schemas (purl, oclc, fdlp, purlz, w3id, ark).
3	IDENTIFIER PERSISTENCY (DATA)	F1	Test for the existence of a persistence policy (Test for the persistence schemas at FAIRSharing or other)	Locate or create a link to an explicit persistence policy, ideally in machine-actionable format	This indicator tests if the unique identifier of the data resource is likely to be persistent. Known schema are registered in FAIRSharing (https://fairsharing.org/standards/?q=&selected_facets=type_exact:identifier%20schema). For URLs that don't follow a schema in FAIRSharing we test known URL persistence schemas (purl, oclc, fdlp, purlz, w3id, ark).
4	STRUCTURED METADATA	F2	Test if a machine is able to find structured metadata (RDF or f.i. embedded JSON)	Provide metadata in RDF or embedded JSON – ideally by following existing metadata templates	Examples of (some) metadata templates can be found here https://fairsharing.org/standards/
5	GROUNDED METADATA	F2	Test if a machine is able to find 'grounded' metadata (resolution leads to a definition of the term)	Provide metadata in RDF or embedded JSON – ideally by following existing metadata templates	Examples of (some) metadata templates can be found here https://fairsharing.org/standards/
6	DATA IDENTIFIER EXPLICITLY IN METADATA	F3	Test if the metadata contains the unique identifier to the data	Separate metadata from the data and assure that the metadata explicitly includes the identifier of the related data. Ideally, to augment adherence to 2, the reverse is also true, the metadata can be linked to, from the data.	For a digital object (PID, metadata and data) the assumption is that the data has its own identifier and that this is provided by standardised keys/predicates. Some predicates that are accepted to indicate the data identifier are foaf:primaryTopic, schema:mainEntity, schema:distribution, sio:is-about and iao:is-about (see below for a complete list). For software the following is expected: schema:codeRepository This is a more exhaustive list of predicates accepted for indicating the data identifier: http://www.w3.org/ns/ldp#contains http://xmlns.com/foaf/0.1/primaryTopic http://schema.org/mainEntity

					http://schema.org/codeRepository http://www.w3.org/ns/dcat#distribution http://schema.org/distribution http://semanticscience.org/resource/SIO_000332', # is about http://semanticscience.org/resource/is-about', # is about http://purl.obolibrary.org/obo/IAO_0000136', # is about http://purl.obolibrary.org/obo/IAO:0000136', # is about (not a valid URL as it contains a colon) https://www.w3.org/ns/ldp#contains https://xmlns.com/foaf/0.1/primaryTopic https://schema.org/mainEntity https://schema.org/codeRepository https://www.w3.org/ns/dcat#distribution https://schema.org/distribution https://semanticscience.org/resource/SIO_000332 # is about https://semanticscience.org/resource/is-about # is about https://purl.obolibrary.org/obo/IAO_0000136
7	METADATA IDENTIFIER EXPLICITLY IN METADATA	F3	Test if the metadata contains the unique identifier to the metadata itself	Separate metadata from the data and assure that the metadata explicitly includes the identifier of the related metadata.	For a digital object (PID, metadata and data) the assumption is that the data has its own identifier and that this is provided by standardised keys/predicates. These are predicates accepted for indicating the metadata identifier: http://purl.org/dc/terms/identifier http://schema.org/identifier
8	SEARCHABLE IN MAJOR SEARCH ENGINE	F4	Test if a machine is able to discover the resource by search (Using MSFT Bing)	Assure your resource is findable in a registered / indexed searchable resource (for this test : MSFT Bing). Domain communities may also register specialized search engines in FAIRshairing and reference these (for which additional MI tests can be designed).	NO ADDITIONAL COMMENTS
9	USES OPEN FREE PROTOCOL FOR DATA RETRIEVAL	A1.1	Test if data may be retrieved by an open and free protocol by testing	Assure your data resources can be resolved (f.i. InChi keys, DOIs , Handles and URLs)	NO ADDITIONAL COMMENTS

			data GUID for its resolution protocol		
10	USES OPEN FREE PROTOCOL FOR METADATA RETRIEVAL	A1.1	Test if metadata may be retrieved by an open and free protocol by testing metadata GUID for its resolution protocol	Assure your metadata resources can be resolved (f.i. InChi keys, DOIs , Handles and URLs)	NO ADDITIONAL COMMENTS
11	DATA AUTHENTICATION AND AUTHORIZATION	A1.2	Test a discovered data GUID for the availability to implement authentication and authorization in its resolution protocol.	Assure your data is compliant with an explicit data access process. (f.i. InChi keys, DOIs , Handles and URLs). The resolution protocol should support authentication and authorisation.	NO ADDITIONAL COMMENTS
12	METADATA AUTHENTICATION AND AUTHORIZATION	A1.2	Test metadata GUID for the availability to implement authentication and authorization in its resolution protocol.	Assure your metadata is compliant with an explicit data access process. (f.i. InChi keys, DOIs , Handles and URLs). The resolution protocol should support authentication and authorisation.	NO ADDITIONAL COMMENTS
13	METADATA PERSISTENCE	A2	Test if the metadata contains an explicitly identified persistence policy.	Locate or create a link to an explicit persistence policy, ideally in machine-actionable format.	Currently, to pass this test it is required to have the following predicate set in the metadata to declare the relevant persistence policy; http://www.w3.org/2000/10/swap/pim/doc#persistencePolicy (NOTE: "swap" is no longer maintained and this is likely an outdated vocabulary. Rarely used predicate!). Contacted Datacite regarding possibility of establishing a persistence policy declaration for DOIs.)
14	METADATA KNOWLEDGE REPRESENTATION LANGUAGE (WEAK)	I1	Test if the metadata uses a formal language broadly applicable for knowledge representation.	Use a machine-actionable language to structure your metadata or use an existing metadata template.	This test takes a broad view of what defines a 'knowledge representation language'; in this evaluation, anything that can be represented as structured data will be accepted. XML, XHTML, RDF, JSON, Turtle, plain text, triples (check with Mark)
15	METADATA KNOWLEDGE	I1	Test if the metadata uses a formal language broadly	Use a (community defined) ontology to structure your	This test takes a broad view of what defines a 'knowledge representation language'; in this evaluation, a knowledge

	REPRESENTATION LANGUAGE (STRONG)		applicable for knowledge representation.	metadata or use an existing metadata template	representation language is interpreted as one in which terms are semantically-grounded in ontologies. Any form of RDF will pass this test (including RDF that is automatically extracted by third-party parsers such as Apache Tika)
16	DATA KNOWLEDGE REPRESENTATION LANGUAGE (WEAK)	I1	Test if the data uses a formal language broadly applicable for knowledge representation.	Use a (community defined) ontology to structure your data or use an existing data template	NO ADDITIONAL COMMENTS
17	DATA KNOWLEDGE REPRESENTATION LANGUAGE (STRONG)	I1	Test if the data uses a formal language broadly applicable for knowledge representation.	Use a (community defined) ontology to structure your data or use an existing data template	NO ADDITIONAL COMMENTS
18	METADATA USES FAIR VOCABULARY (WEAK)	I2	Test if the linked data metadata uses terms that resolve.	Assure the metadata identifier resolves and uses FAIR vocabulary (check F1, A1.1. and I1) .	NO ADDITIONAL COMMENTS
19	METADATA USES FAIR VOCABULARY (STRONG)	I2	Test if the linked data metadata uses terms that resolve to linked FAIR data	Assure the metadata identifier resolves and uses FAIR vocabulary (check F1, A1.1. and I1) .	NO ADDITIONAL COMMENTS
20	METADATA CONTAINS QUALIFIED OUTWARD REFERENCES	I3	Test if metadata (as Linked Data) links outward to third-party resource	Assure you dataset can be represented as Linked Data and has machine-actionable references to other metadata.	NO ADDITIONAL COMMENTS
21	METADATA INCLUDES LICENSE (STRONG)	R1.1	Test if metadata contains an explicit pointer to the license	Assure an explicit pointer to the license or use existing schemas that include license terms.	Maturity Indicator to test if the linked data metadata contains an explicit pointer to the license. Tests: xhtml, dvia, dcterms, cc, data.gov.au, and Schema license predicates in linked data, and validates the value of those properties. Standardised predicates/keys used to indicate the license for use of the dataset/software:

					http://www.w3.org/1999/xhtml/vocab#license http://purl.org/ontology/dvia#hasLicense http://purl.org/dc/terms/license http://creativecommons.org/ns#license http://reference.data.gov.au/def/ont/dataset#hasLicense NOTE THAT THE VALUE OF THAT PROPERTY MUST BE A Resource (URL). Strings are not accepted.
22	METADATA INCLUDES LICENSE (WEAK)	R1.1	Test if metadata contains an explicit pointer to the license	Assure an explicit pointer to the license or use existing schemas that include license terms.	Maturity Indicator to test if the metadata contains an explicit pointer to the license. **This 'weak' test will use a case-insensitive regular expression, and scan both key/value style metadata, as well as linked data metadata**. So if your metadata is e.g. JSON, and has a key 'license', that has a resource as its value, you pass the weak test. Tests: xhtml, dvia, dcterms, cc, data.gov.au, and Schema license predicates in linked data, and validates the value of those properties. Standardised predicates/keys used to indicate the license for use of the dataset/software: http://www.w3.org/1999/xhtml/vocab#license http://purl.org/ontology/dvia#hasLicense http://purl.org/dc/terms/license http://creativecommons.org/ns#license http://reference.data.gov.au/def/ont/dataset#hasLicense NOTE THAT THE VALUE OF THAT PROPERTY MUST BE A Resource (URL). Strings are not accepted.