

FAIR Metric FM-I2

Mark D. Wilkinson, Susanna-Assunta Sansone,
Erik Schultes, Peter Doorn,
Luiz Olavo Bonino da Silva Santos, Michel Dumontier

November 21, 2017

<i>FIELD</i>	<i>DESCRIPTION</i>
Metric Identifier	FM-I2
Metric Name	Use FAIR Vocabularies
To which principle does it apply?	I2 - (meta)data use vocabularies that follow FAIR principles
What is being measured?	The metadata values and qualified relations should themselves be FAIR, for example, terms from open, community-accepted vocabularies published in an appropriate knowledge-exchange format.
Why should we measure it?	It is not possible to unambiguously interpret metadata represented as simple keywords or other non-qualified symbols. For interoperability, it must be possible to identify data that can be integrated like-with-like. This requires that the data, and the provenance descriptors of the data, should (where reasonable) use vocabularies and terminologies that are, themselves, FAIR.
What must be provided?	UUIDs representing the vocabularies used for (meta)data
How do we measure it?	Resolve UUIDs, check FAIRness of the returned document(s)
What is a valid result?	Successful resolution; document is amenable to machine-parsing and identification of terms within it.
For which digital resource(s) is this relevant?	All
Examples of their application across types of digital resource	None

Comments	<p> michel: there must be a syntax and associated semantics for that language. This is sufficient mark: there needs to be some identity or denotation in the language; ('vanilla') xml and json are not FAIR, so should fail this test </p> <p> *** can you (i) identify elements and (ii) make statements about them, and iii) is there a formally defined interpretation for that -> HTML fails; PDF fails shared -> that there are many users of the language . acknowledged within your community -> hard to prove. . could we use google to query for your filetype (can't discriminate between different models) -> has a media type </p> <p> -> This SHOULD be stated as a IANA code [IANA-MT] </p> <p> standardization of at least this listing process is a good measure of "sharedness" </p> <p> broadly applicable . that the language is extensible to a domain of interest . you can define your own elements in accordance with the semantics of the language </p> <p> gff3 is not in the IANA list -> what steps would the community need to execute to be listed here? cases like GFF, PDB are not broadly applicable biopax -> is defined vnd.biopax.rdf+xml and built on rdf -> allows users to create new elements and relate them jpg -> widely used, registered, but primarily for image content pdf -> registered, enables users to create their own dictionary. </p>
----------	---