

Generic FAIR notes (also to help make data FAIR)

Generic codes against our own metadata plans (suppliers or API usually do not provide FAIR data).
 Inspired by codes from **GoFAIR**. There may be many valid approaches to this.

FAIR CODE	Very brief explanation	Standards suggested (our approach, by design)
FM1-F1A	Global identifiers used.	<p>e.g.</p> <p>DOI, ror.org, ORCID and well known IRI schemes based on good practices e.g., concepts and their associated UUID or Hex hash digest.</p> <p>IRI</p> <p>Internationalised URI. RFC 3987 (identifiers)</p> <p>(Usually a resolvable URL)</p> <p>DOI</p> <p>(Crossref) Digital Object Identifier. Popular with scientific assets like published papers and data.</p> <p>ORCID</p> <p>Open Researcher and Contributor ID</p> <p>Ror.org</p> <p>Permanent ID registry for organisation and institutions.</p> <p>Hex hash digest (among others)</p> <p>MD5</p> <p>Hash algorithm. RFC 1321</p> <p>SHA</p>

		<p>Family of hash algorithms RFC 4634.</p> <p>W3C Asset Description Metadata Schema (ADMS)</p> <p>We can use ADMS for identifiers that are text literals like e.g. ISBN</p>
FM1-F1B	<p>The identifier must be permanent.</p> <p>(Persistent Identifier)</p>	Same as above.
FM-F2	Rich metadata.	In our case, semantic linked data. Extendible in many ways and with multiple serialisations.
FM-F3	Data includes identifiers.	The metadata contains the identifiers of other objects, including the data.
FM-F4	Data indexed in a searchable resource.	Public data can be searched somewhere public.
FMA1	Retrievable by identifier using a standard protocol	Yes, HTTP
FM-A1.1	Open, free etc communications protocol.	Yes, HTTP
FM-A1.2	Protocol allows for authentication and authorisation.	Yes
FM-A2	Metadata serialisation is accessible even when data files are no longer available.	Yes
FM-I1	It uses language for knowledge representation.	<p><u>Enumeration of standards.</u></p> <p>This list can be expanded as needs arise.</p>

		<p>W3C OWL Web Ontology Language</p> <p>Topics categories (coarse)/ Themes (fine) (i.e. not free keywords)</p> <p>Scientific categories: EuroSciVoc</p> <p>Multiple serialisations of RDF:</p> <p>W3C RDF Schema Resource Description Framework. (e.g., <i>W3C JSON-LD</i> <i>JSON for Linked Data. RDF document (serialisation). 100% backward compatible with JSON.</i>)</p> <p>COAR for <i>access level</i> (e.g.) Confederation of Open Access Repositories.</p> <p>Metadata for CSV files:</p> <p>W3C CSVW Vocabulary for tabular data on the web. (Albeit other representations are a better choice)</p> <p>Datasets:</p> <p>W3C DCAT Data Catalog Vocabulary based on OWL and RDF. Includes elements from DCTERMS, FOAF and PROV-O. Non-normative vocabularies and ontologies also suggested.</p>
--	--	---

		<p>Further provenance and lineage:</p> <p>W3C PROV-V</p> <p>Provenance Vocabulary.</p> <p>Duration and timestamping:</p> <p>ISO 639-1</p> <p>Two characters language ID.</p> <p>ISO 639-2</p> <p>Three characters language ID.</p> <p>ISO 8601</p> <p>Datetime standard (RFC 3339).</p> <p>Scientific parameters and I/O:</p> <p>W3C DQV</p> <p>Data Quality Vocabulary. (e.g., for scientific I/O, parametrisation)</p> <p>ISO 17369:2013 aka SDMX 2.1</p> <p>Statistical Data and Metadata Exchange (e.g. for scientific attributes)</p> <p>QUDT (for SI units)</p> <p>Scientific Description:</p> <p>EMMO</p>
FM-I2	Vocabularies used are themselves FAIR.	Yes. International standards.

FM-I4	It uses qualified references to other metadata.	Yes. International standards.
FM-R1	Richly described with accurate and relevant attributes.	Yes, it is extendible and progressively enriched.
FM-R1.1	It must have open licences identifiers. Other licences identifiers can be added.	ISO/IEC 5962:2021 aka SPDX 2.2.1 Software Package Data Exchange. (e.g., Open licences identifiers)
FM-R1.2	Detailed provenance must be possible	DCAT (datasets) and W3C PROV-O.
FM-R1.3	Meets community standards	Metadata and data (files) are processable by users. Open file standards encouraged. Data views on request (if applicable).

Brief F-UJI tool FAIR interpretation analysis

F-UJI list of codes

e.g., The F-UJI assessment tool [interpretation of FAIR with corresponding codes](#). We contrasted the FAIR codes from their solution against others, their implementation and our FAIR approach (intended to be FAIR by design and leveraging a triplestore database as the main foundation).

F-UJI tool FAIR metrics	Our comments against our proposal for supplied data.	Brief implementation analysis
FsF-F1-01D Data is assigned a globally unique identifier	Same. (Global PID). Note: includes GUID and hashes.	<ul style="list-style-type: none"> • Requires working URL. • Features may rely on external services. • List of valid schemes.
FsF-F1-02D Data is assigned a persistent identifier	Same.	
FsF-F2-01M Metadata includes descriptive core elements to support data findability.	e.g. a creators list with corresponding identifiers.	<ul style="list-style-type: none"> • Requires working URL to fetch the object. • It expects hardcoded elements as metadata.
FsF-F3-01M	By design. In DCAT datasets that is in the	<ul style="list-style-type: none"> • Requires proper IDs for data files linked in the metadata.

Metadata includes the identifier of the data it describes.	distribution entry (data files identifiers).	
FsF-F4-01M Metadata is offered in such a way that it can be retrieved by machines.	We use a serialisation standard. (RDF in our case).	<ul style="list-style-type: none"> • Fetching metadata by automation possible.
FsF-A1-01M Metadata contains access level and access conditions of the data.	e.g. COAR	<ul style="list-style-type: none"> • Conditional access to data and metadata.
FsF-A1-02M Metadata is accessible through a standardized communication protocol	HTTP	<ul style="list-style-type: none"> • Both metadata and data are accessible by their respective communications protocols.
FsF-A1-03D Data is accessible through a standardized communication protocol	HTTP	
FsF-A2-01M Metadata remains available, even if the data is no longer available.	Separation of metadata and data (files) access.	<ul style="list-style-type: none"> • Independent access for data and metadata.
FsF-I1-01M Metadata is represented using a formal knowledge representation language.	Standards: Vocabularies and Ontologies	<ul style="list-style-type: none"> • List of valid serialisation standards.
FsF-I1-02M Metadata uses semantic resources.	Standards: Vocabularies and Ontologies	<ul style="list-style-type: none"> • List of expected semantic namespaces.
FsF-I3-01M Metadata includes links between the data and its related entities.	Relation between the metadata and the data. Provenance. By design in DCAT.	<ul style="list-style-type: none"> • Data Included attribution/provenance related to metadata (DCAT/PROV-O etc)
FsF-R1-01MD Metadata specifies the content of the data.	e.g., data files metadata (content, MediaType, size)	<ul style="list-style-type: none"> • Data files have their respective metadata description (e.g., distribution view)

<p>FsF-R1.1-01M Metadata includes license information under which data can be reused.</p>	<p>Licensing metadata included.</p>	<ul style="list-style-type: none"> • Licensing info.
<p>FsF-R1.2-01M Metadata includes provenance information about data creation or generation.</p>	<p>Attribution and provenance included and processable.</p>	<ul style="list-style-type: none"> • attribution/provenance (e.g., creator) (DCAT/PROV-O etc)
<p>FsF-R1.3-01M Metadata follows a standard recommended by the target research community of the data.</p>	<p>F-UJI uses standards in this catalogue list (any alternative list may be valid)</p>	<ul style="list-style-type: none"> • Metadata standard must be in this catalogue list
<p>FsF-R1.3-02D Data is available in a file format recommended by the target research community.</p>	<p>Essentially a list of open standards.</p>	<ul style="list-style-type: none"> • Data must be in the list of standards.