



JRC FAIR Data Guidelines

Lowenthal, H., Austin, T., Bonino Da Silva Santos, L.O.,
Chiarelli, C., Cusinato, A., Ferigato, C., Friis-Christensen,
A., Kemper, T., Perrotta, D., Wittwehr, C.

2025

This document is a publication by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The contents of this publication do not necessarily reflect the position or opinion of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contact information

Name: Hagar Lowenthal

Address: Via E. Fermi, 2749, 21027 Ispra (VA), Italy

Email: Hagar-Isabel.LOWENTHAL1@ec.europa.eu

EU Science Hub

<https://joint-research-centre.ec.europa.eu>

JRC140764

EUR 40231

PDF ISBN 978-92-68-24742-6 ISSN 1831-9424 doi:10.2760/5646214 KJ-01-25-098-EN-N

Luxembourg: Publications Office of the European Union, 2025

© European Union, 2025



The reuse policy of the European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

For any use or reproduction of photos or other material that is not owned by the European Union permission must be sought directly from the copyright holders.

How to cite this report: European Commission: Joint Research Centre, Lowenthal, H., Austin, T., Bonino Da Silva Santos, L.O., Chiarelli, C., Cusinato, A., Ferigato, C., Friis-Christensen, A., Kemper, T., Perrotta, D. and Wittwehr, C., *JRC FAIR Data Guidelines*, Publications Office of the European Union, Luxembourg, 2025, <https://data.europa.eu/doi/10.2760/5646214>, JRC140764.

Contents

Abstract	2
Executive summary.....	3
1 Introduction	5
2 The Guidelines.....	7
3 FAIR Levels.....	8
3.1 FAIR start.....	10
3.1.1 Is this enough for you?	11
3.1.2 Examples	11
3.2 FAIR play.....	13
3.2.1 Is this enough for you?	15
3.2.2 Examples	16
3.3 FAIR go.....	17
3.3.1 Is this enough for you?	18
3.3.2 Examples	18
3.4 FAIR share.....	21
3.4.1 Is this enough for you?	22
3.4.2 Examples	23
3.5 FAIR rest of them all.....	25
3.5.1 Is this enough for you?	26
3.5.2 Examples	27
4 Conclusions.....	29
References	30
List of boxes.....	32
List of figures.....	33
Annex 1. FAIR maturity assessment grid.....	34
Annex 2. Further reading	37

Abstract

The JRC Data Policy requires that JRC Data be managed following the FAIR (Findable, Accessible, Interoperable, Reusable) Principles. This paper is a deliverable of the JRC Data Strategy in its call to publish Guidelines to support and enhance the FAIRness of JRC data. The paper provides a framework for evaluating and increasing the level of adherence of JRC datasets to the FAIR Principles, while taking into account specific solutions and strategies available to data publishers. It proposes five progressive levels of FAIR Maturity, indicating which of the FAIR maturity indicators the levels follow and suggesting steps that can improve their FAIR maturity level.

Executive summary

The FAIR (Findable, Accessible, Interoperable, Reusable) Principles were published in 2016 by a group of stakeholders working with scientific data with the aim of facilitating the use of such assets by machine agents. The Joint Research Centre of the European Commission (JRC) acknowledges the importance of publishing data following the FAIR Principles in its Data Policy and Data Strategy. In this context, the Guidelines wish to support JRC data publishers in mapping the intentionally abstract Principles into tangible actions and strategies that can help:

- increase FAIRness of published data;
- set a basis for a simplified evaluation; and
- illustrate how specific EC data policies and JRC solutions support the FAIR principles.

To do so, the Guidelines do not focus on single elements of the FAIR principles or FAIR maturity indicators but rather illustrate how existing EC and JRC solutions readily support adherence to these. For this aim, the paper groups various indicators, following those suggested by the Research Data Alliance FAIR Data Maturity Model, into five progressive maturity levels. Each level specifies what actions are necessary to reach the designated maturity level and lists recommended measures that can be taken to reach higher levels. The paper also includes tip-boxes informing publishers on what can be done to address specific issues such as selecting a data repository, using controlled terms from vocabularies etc. Exemplary JRC datasets that adhere to each defined level are cited and provide a tangible illustration of what each maturity level should look like in practice.

The five FAIR maturity levels progress in correspondence to evaluation of FAIRness of metadata, data and machine-readability. The baseline levels (FAIR start and FAIR play) illustrate how using readily available JRC and EC solutions such as data catalogues ensure adherence to several FAIR Principles. Specifically, the Guidelines demonstrate how indicators assessing Findability and Accessibility of datasets are fulfilled by using prescribed JRC solutions. Furthermore, the baseline levels enjoy full interoperability of metadata thanks to the usage of these solutions. The following three levels (FAIR go, FAIR share and FAIRest of them all) focus on assessing Interoperability of data contents. This is done by examining availability of relevant documentation, use of relevant formats and data services and evaluation of the endorsement of these elements by relevant scientific domain communities. Admittedly, data interoperability is a complex issue and is inevitably specific to the datasets at hand. Nevertheless, the Guidelines suggest strategies and tactics that can increase FAIR maturity level of data contents. Datasets belonging to the highest level represent a scenario where machine agents can not only locate and access the data but also reuse them while maintaining the intended semantics of the original datasets.

A short check-list summarising the elements each level complies with is provided while a full mapping between the RDA indicators and the suggested FAIR levels is available in Annex 1. The annex offers an overview of the elements that are assessed for each level and clearly shows how several FAIR principles, addressed by existing corporate solutions, support FAIRness of JRC data with no need for additional actions by data publishers. Consider, for example, the FAIR Principle “R1.1: (Meta)data are released with a clear and accessible data usage license”, and the three RDA indicators suggested for assessing it (RDA-R1.1-01M: Metadata includes information about the licence under which the data can be reused; RDA-R1.1-02M: Metadata refers to a standard reuse licence; and RDA-R1.1-03M: Metadata refers to a machine-understandable reuse licence). The

Commission Decision on Reuse ([2011/833/EU](#)) provides the strategic framework for assigning the relevant licence to data assets while the data model used by the JRC's data catalogue ensures it is provided in a machine-understandable way by using standard attributes and terms taken from a controlled vocabulary.

Future work could use the Guidelines for developing assessment tools specific to the Commission's metadata specification and be propagated to a wider institutional audience.

1 Introduction

One of the ambitions of the Commission's European Strategy for Data is to deliver high-quality data, compliant with the FAIR (Findable, Accessible, Interoperable and Reusable) Principles¹. The JRC Data Policy explicitly foresees that our data should be FAIR². Furthermore, all receivers of Horizon Europe funding need to abide by DG RTD's Guidelines on Open Data. Specifically, all data produced in this framework must be FAIR by default³.

The FAIR Principles [1] were published in 2016 by a group of various stakeholders in the domain of scientific data publication. In short, the FAIR initiative aims at increasing (re)use of scientific data by making it more Findable, Accessible, Interoperable and Reusable, not just to humans but also, primarily, to machines. The present Guidelines assume that users are already familiar with the Principles without going into great detail. A list of resources for further reading is provided in [Annex 2](#).

The FAIR Principles are a conceptual framework and intentionally do not prescribe specific means or guidelines for implementation of each Principle. This does though present a challenge when attempting to evaluate the FAIRness of digital objects, because the Principles can be applied to any level of the "digital asset" i.e. metadata, data or associated (data) services. To be able to consider the various elements of the JRC data ecosystem, we would therefore first like to define the following elements:

- **Metadata** – The main focus of the FAIR Principles, metadata, is descriptive data. It can be compared to a label on a soda bottle that identifies its contents, ingredients, expiry date, producer etc. However, it is not always easy to distinguish between the various levels of metadata. For example, information about identifiers can be provided on a "**metadata level**" (e.g. PID of the datasets, ORCID identifier of authors) and on a "**data level**" (e.g. tabular data where ISO codes are used to define languages, CAS Registry Numbers to identify chemical substances, or use of ISO 19136 (TC211) Geography Mark-up Language (GML) schemas to model the data).

¹ See: [COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, A European strategy for data, COM/2020/66 final](#). Pillar A (A cross-sectoral governance framework for data access and use) of the data strategy cites the FAIR Principles as means to accomplish "a more harmonised description and overview of datasets, data objects and identifiers to foster data interoperability".

² art. 15 JRC Data Policy DOI 10.2760/637912

³ Cfr. <https://open-research-europe.ec.europa.eu/for-authors/data-guidelines>; Article 17.4 of the Model Grant Agreement https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/agr-contr/general-mga_horizon-euratom_en.pdf

- **Data** – A collection of values that convey information. When the data forms a curated collection (typically by a single agent), it can be referred to as a dataset. When discussing data in these Guidelines it is good to remember its inevitably domain specific nature. For example, encoding sensory geospatial data is by definition very different from encoding genomic sequencing data. The models used for each resource are largely influenced by how the respective communities expect to use it. Another aspect to keep in mind is the fact that data often include embedded metadata (e.g. code lists, data schemas, datatypes etc.), as in the identifier example above.
- **Other Digital Objects** - The FAIR Principles are not limited to (meta)data. Any digital object used to produce research data can and should be FAIR. There is growing body of work emphasising the need to consider all components of the data ecosystem such as code, software, protocols, models, reference data and workflows on the same level that is reserved for data assets. The literature suggested in [9] and [10] chapter 4 is recommended for further reading on the topic.

In recent years, more specific work has been published to support publication of FAIR data assets by providing measures (and tools) for assessment of FAIR maturity levels [3]. Numerous initiatives support researchers who wish to make their data FAIR: the FAIR cookbook [2] is a valid hands-on resource for identifying and understanding many specific issues one must contemplate when wishing to publish FAIR data. The [Research Data Alliance](#) (RDA) actively supports and promotes FAIR Data. The link below [6] lists past and current Work Groups sponsored by the Alliance. It is also possible to search the forum by discipline (<https://www.rd-alliance.org/rda-disciplines>). The [European Open Science Cloud](#) (EOSC) is becoming an important member of the FAIR community. One of the Strategic Objectives (SO4) of the EOSC Association is to make publicly financed research data FAIR by default⁴. Several EOSC Tasks Forces [7] are working towards that aim in close collaboration with the RD Alliance and the GO FAIR Foundation.

⁴ Memorandum of Understanding for the Co-Programmed European Partnership for the European Open Science Cloud (EOSC) https://www.eosc.eu/sites/default/files/EOSC_Memorandum_30_July_2021.pdf p. 3

2 The Guidelines

The call to follow the FAIR principles is evermore recurrent and explicitly made by important institutions and funders who sponsor publication and dissemination of research data. These guidelines wish to provide context-specific guidance considering the availability of JRC tools, platforms, workflows and requirements. Their aim is to facilitate the application of the, primarily conceptual, FAIR principles by providing a grid for a quick evaluation of the current and potential level of FAIRness of JRC data assets. The true aim of the Guidelines, however, is not merely to assess the level of FAIRness but rather to indicate available tools and measures for enhancing it.

Acknowledging that machine-readability and machine-actionability are the primary goals of the FAIR Principles, the Guidelines aim to evaluate machine-to-machine interaction for distinct characteristics of data assets. This assessment is achieved by identifying which aspects can be made FAIR using available JRC data management tools and which further customized efforts may be required to address more specific requirements.

The structuring of the guidelines in five levels must not be confused with a grading scale. Rather, it aims to provide data publishers with a simple categorisation of FAIR levels, including the means to assess their adequacy to their intended use. Making data FAIR requires effort and this is why publishers should assess whether the level they aim to achieve is appropriate for the dataset. This said, even the baseline level should already ensure a primary FAIR maturity level. For each level, an example is provided to make it more tangible. A mapping of each level to the RDA FAIR Data Maturity Model indicators [3] is provided in [Annex 1](#) for a quick overview of the various elements that can be assessed.

The choice of which level to aim for mostly depends on the mandate under which data is curated and the expectation of its re-use. For the latter, it is important to note that, when publishing open data, it may prove difficult to assess who the expected re-users are or what is the expected re-use. A case in point is Artificial Intelligence (A.I.) and neural networks where stacks of data are used to train machine-learning models. Nonetheless, assessment of this aspect, e.g. by interviewing known or potential re-users, is highly recommended and, even if bound to be somewhat inchoate, can guide your choice of the proper level of FAIRness. Naturally, aiming for a particular FAIRness level does not need to be definitive. As context changes, the re-use scenarios may also change and, therefore, the effort to reach other levels may arise.

3 FAIR Levels

The Guidelines define five FAIRness categories and provide a framework for evaluating datasets' compliance to these categories. Each level includes an overview of the observed FAIR principles and guidance for strategies that could elevate datasets to a higher level. The levels include links to exemplary datasets that fulfil the defined requirements. Additionally, [Annex 1](#) maps the assessment of each level to the [RDA Data Maturity Model](#), which can be used to provide a quick overview of the specific aspects that are assessed for each level.

The levels assess the extent to which each of the FAIR Principles is covered with regards to both metadata and data. The first two levels focus mostly on the “F” (findability) and the “A” (accessibility) of FAIR, while the others provide guidance to assessing the “I” (interoperability) and the “R” (reusability) of FAIR by looking more closely at the data itself. Each level concludes with a section called “Is this enough for you?” where additional steps to enhance data FAIRness are proposed.

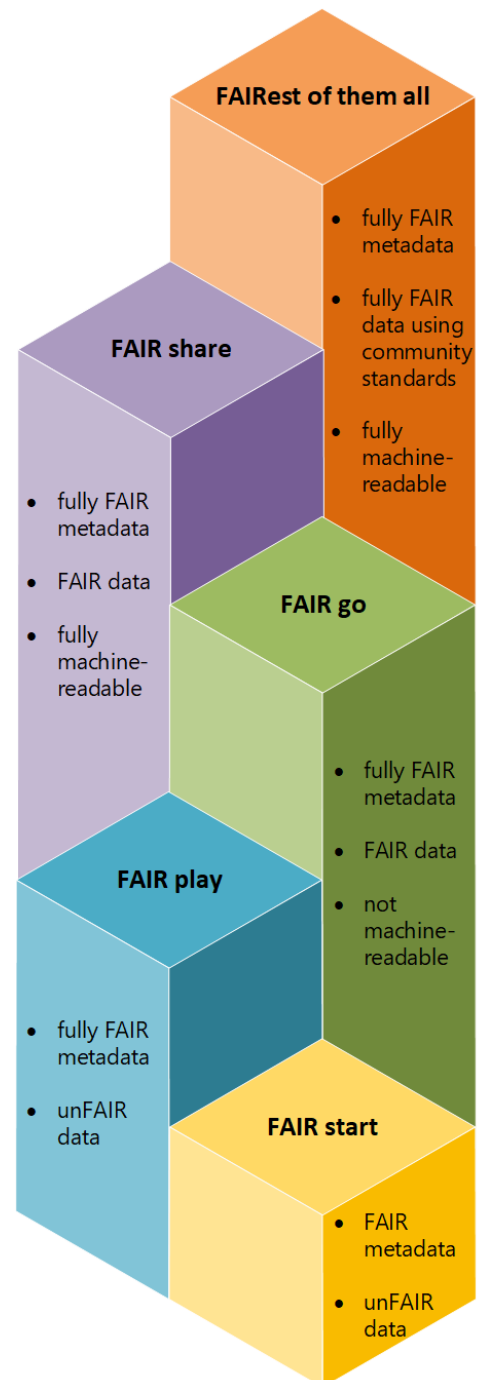
It should be noted that, while the aim is to improve data FAIRness, the levels are context-specific and dynamic. In practice, this means that even if your data complies with a certain level at a given moment, changes in standards, platforms, protocols and other factors may lower data's FAIRness level over time. While many of the technological and societal evolutions will be addressed and updated on an institutional level (such as the metadata standards used by the data catalogue, persistency of unique identifiers etc.), more domain-specific aspects should be monitored by data owners and data stewards to avoid the risk of falling behind and decreasing FAIRness level of data.

The levels have been designed to be progressive, adding features from one to the next. The guidelines also provide strategies and tip-boxes that are relevant for achieving each of the FAIRness levels. The five levels can be summarised as follows:

1. **FAIR start** – the basic level expected of all JRC datasets. Findability and accessibility of datasets are addressed by using specific platforms provided by JRC and by providing essential metadata in keeping with JRC specifications.
2. **FAIR play** – this level further supports the findability and accessibility principles. It also lays the groundwork for interoperability by providing links to resources that offer supplementary information about lineage and methodology used to produce the data in human-readable formats.
3. **FAIR go** – in this level, datasets, in addition to the requisites of the previous ones, make specific reference to the models that standardise how data is encoded. The data contents of the datasets make use of standardised terms and concepts. For this level, the models and terminology may not be available in machine-readable formats, which implies that interoperability and reusability are only possible through human mediation.
4. **FAIR share** – datasets belonging to this level use data models that are machine actionable, in addition to the elements listed for the previous levels. As with the previous level, terms used to denote data elements are standardised, and, in addition, are made accessible to machine agents. These models are not necessarily or fully endorsed by relevant domain communities.

5. **FAIRest of them all** – the top level of FAIRness included in the Guidelines. In this level the data model is not only machine-readable but is endorsed and maintained by the relevant scientific domain community. Terms and concepts used for encoding data elements are likewise machine-accessible and shared by the relevant domain community to maximise interoperability and re-use by machine agents.

Figure 1. Schematic overview of the five FAIR maturity levels



Source: JRC

3.1 FAIR start

Dataset is:

- Published in the data catalogue
 - ☑ Mandatory metadata elements are recorded
- Data is unFAIR:
 - ☑ Exposed using local model
 - ☑ Terms do not refer to a data dictionary

“FAIR start” is the minimum level of FAIRness expected of all JRC datasets. When producing and publishing data in line with JRC data policies, your data is already FAIR to a certain degree. In practical terms, it means that data can be easily found and accessed but may be suboptimal in terms of interoperability and reusability.

In the context of JRC data, information at the “metadata level” is managed following the [EC](#) and [JRC data policies](#) alongside the [EC re-use decision](#). This strategic framework fosters adherence to several FAIR Principles by default. Because all JRC (meta)data must be published in the [JRC Data Catalogue](#), datasets gain a good level of Findability and Accessibility upon submission. The descriptive information is provided following a defined specification⁵ and is made available for automatic harvesting by other data portals, such as the European data portal⁶ and Google Datasets⁷. However, this does not imply that the data itself is FAIR. For example, an accessible PDF resource in the catalogue will be sufficiently FAIR in terms of Findability but will hardly be reusable by machine agents.

In terms of FAIRness, following EC and JRC data policies and using the JRC Data Catalogue to publish your data ensure the following principles:

- Persistent identifiers are used to identify (meta)data, which resolve to a metadata record (for both human and machine access);
- Metadata remains available even if data is not (e.g. data is withdrawn, data repository is no longer maintained);
- Metadata is provided using a community standard ([DCAT vocabulary](#)) that is machine-readable and enables standardised way to access the data;
- Metadata is harvested and indexed by other data portals^{6 7}

⁵ The Data Catalog Vocabulary ([DCAT](#)) is a W3C specification. [DCAT-AP](#) is the application profile fostered by the European Commission to ensure sufficient interoperability between European data portals. The JRC extension of the EU Application Profile (<https://ec-jrc.github.io/dcat-ap-jrc/>) adds additional obligations in line with JRC’s institutional role.

⁶ <https://data.europa.eu/>

⁷ <https://datasetsearch.research.google.com/>

- The metadata [Application Profile](#) mandates provision of rich metadata and makes extensive use of FAIR vocabularies to describe specified elements (e.g. subjects, licences, locations, languages);
- (Re)use and access conditions are stated in machine-readable manner.

3.1.1 Is this enough for you?

This level is most appropriate for datasets that support delimited projects for which continuity is not foreseen. By default, JRC data must remain available to ensure trustworthiness of the results of JRC projects. While metadata of the dataset is FAIR, data itself is not. This means that any re-use of the underlying data, beyond accessing it, is not straightforward for a machine agent. As we will see in the next levels, to be fit for re-use by machine agents, data needs to adhere to the Interoperability element of FAIR by using community standards and providing data in formats that are recommended for machine readability.

If your data could be re-used by others or is provided according to a known specification, you should consider aligning to a higher level of FAIRness. This level is equally scarce for data that links to other resources, such as publications, data services or other datasets (e.g. source data, datasets that replace/are replaced by your dataset, versions).

Box 1. Storing your data

Particularly when producing data of limited scope, a frequent risk is the discontinuation of infrastructure used for the project. Designated repositories that were used to host data and documentation (e.g. project web pages, visualisations, dashboards) may become obsolete and will result in lost access to these resources in the long or medium term. Even if you choose to host your resources in trustworthy platforms such as GitHub or Zenodo, there is still a tangible risk that changes in the composition of the team responsible for the resources will result in loss of resource management credentials (e.g. editing rights) because, when using an external platform, these are often linked to an individual person.

Unit T.4 can assist you with making sure your data remains available over time by providing data repository services. It is therefore good practice to use JRC dedicated repositories. In addition to increasing the certainty of the resources' availability over time, using JRC repositories will unburden you from the tasks of managing maintenance, migrations, redirections, security etc.

3.1.2 Examples

- **Boyd, Mark; Vaccari, Lorenzino (2020): API best practices references. European Commission, Joint Research Centre (JRC) [Dataset] PID: <http://data.europa.eu/89h/7340ab8a-ef73-459b-a2d9-b64e1a5bb680>**

The dataset provides a list of literary resources on APIs. The data is provided as tabular data (in two formats) and includes ad-hoc classifications, which are explained in the descriptive metadata. Considering the specificity of the scope, it is likely that no community classifications are available to express desired terms. Re-use expectation is limited to

identifying the listed resources, which is supported by making the dataset publicly findable and accessible. It is not feasible that machine agents could use the data beyond the declared scope.

No links are provided to related JRC publications or other resources because the resource is not referenced by any. Nevertheless, being part of a JRC Collection (<https://data.jrc.ec.europa.eu/collection/id-0097>), the dataset is readily linked to additional datasets of the same collection.

- **European Commission, Joint Research Centre (JRC) (2020): Broadband – households having access to the internet. [Dataset] PID:**
http://data.europa.eu/89h/jrc-10113-rio_bband_acc

The dataset provides data on access of European households to broadband internet connections. The data is stored in JRC's corporate data repository to ensure it is accessible and findable, regardless of the duration of the project. No publications are referenced by the dataset. The Collection, of which the dataset is part, gives additional context by stating that the data collected in the project's framework was used to support research and innovation policymaking initiatives.

3.2 FAIR play

Dataset is:

- Published in the data catalogue
 - ☒ Rich metadata is provided
 - ☒ Metadata contains links to other resources
- Data is unFAIR:
 - ☒ Exposed using local model
 - ☒ Terms do not refer to a data dictionary

By “FAIR play”, we refer to datasets published in the Data Catalogue that reference related resources, such as publications citing the data, additional documentation, or other datasets, such as versions or source data from which the dataset is derived. Although many JRC datasets feature such links to other resources, these are not always encoded correctly, nor to their full extent. Adding this information in a structured way to datasets’ metadata, provides meaningful information and facilitates linking to relevant resources and interpreting uncertain values. This, in turn, increases data’s reliability and potential for re-use.

In the context of JRC, publications are often considered the main deliverables of projects. However, even something as obvious as indicating links between data and a publication is not always observed⁸, be it by correctly referencing datasets cited by the publication or referencing publications that cite the data. JRC’s data ecosystem provides means to denote such links, both via PUBSY, the corporate JRC outputs management system, and via the Data Catalogue. In the latter instance, the Dublin Core⁹ term “isReferencedBy” should be used for capturing this information. When citing a publication that is registered in PUBSY, all necessary metadata is retrieved automatically based on the output’s identifier.

The Catalogue’s metadata vocabulary ([DCAT-AP-JRC](#)) provides a handful of attributes to encode qualified relationships between a dataset and other resources. For example, a dataset that is part of a Collection automatically denotes this relationship in a standardised way¹⁰. Additional attributes can provide important information regarding any of the following aspects:

⁸ A useful tool that provides a handful of tips on how to write on data and harness it to support publications was published in 2022 by data.europa.eu. See: <https://data.europa.eu/apps/data-in-publications-guide/>. A dedicated function permitting linking dataset to any PUBSY publication is described here: <https://europea.eu.sharepoint.com/sites/ORG-jrc-portfolio-17/SitePages/Publication-of-data-alongside-reports-and-articles---enhancing-our-data-sharing-culture-with-a-new-dataset-lin.aspx>

⁹ Dublin Core and the Dublin Core Terms are amongst the most widely used set of metadata terms for describing resources. For the latter specification see: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

¹⁰ See: <https://ec-jrc.github.io/dcat-ap-jrc/#collection>

- **related resources** – these include documentation which gives additional information on the dataset, such as publications, methodology, or the project that governs its publication. Other related resources, such as data visualisations or dashboards, can also be linked to the dataset using specific attributes.
- **alternative identifiers** – datasets published in the Data Catalogue are automatically assigned with a primary PID (persistent identifier). It is also possible to indicate the equivalence of external identifiers assigned to the same dataset, such as DOIs, as well as linking it to an identifier used by another repository where the same dataset was deposited. The data model further supports using identifiers of entities associated with the dataset such as authors, contributors and funders (e.g. by stating ORCID identifiers of personal authors, or ROR identifiers of research organisations).
- **provenance (lineage) of the dataset** -
 - Textual information – this can be indicated either as a URI of a document or literal text providing information on any change in datasets' lineage to support evaluation of its integrity and interpretation.
 - Common provenance relations between datasets or datasets and other catalogued resources can be expressed using specific attributes to denote **source data** used for the creation of the described dataset; **versions** of the dataset including links to resources that deprecate or are replaced by it, **membership** relationships to indicate resources that are part of or included in other resources.
 - Creation and update dates are recorded on various levels of the dataset (e.g. when was the metadata record created/modified; when was the dataset created/modified; when was the described data created/modified).
 - The metadata model supports expressing more complex qualified relationships by stating specific roles that other resources, namely agents, fulfilled in the curation of the dataset. This information is important for supporting interpretation and reproducibility of data results and will be the subject of the [“FAIR share” level](#).

Put together, denoting these relations increases trustworthiness and reliability of datasets. They facilitate quick discovery of related resources. They guide (re)users in how to interpret and assess data contents and lay the basis for reproducibility of data results.

In terms of FAIRness, publishing your data in the Data Catalogue and including additional links to related resources adds compliance with these principles (i.e. in addition to those listed for the previous level):

- Cross-referencing resources enhances interoperability
- Provenance information facilitates re-use in a manner that takes into account all relevant data

Box 2. Linking to data visualisations

Data visualisations are a powerful tool for communicating data in an accessible way. They assist in providing insights and quickly identifying trends in complex data in a manner that is not reserved for expert users. EC Store contains several data viz software tools (under the category: Business applications: Data analysis). When planning to visualise your data, the [Data Visualisation Guide](#) can help you assess many relevant issues and choose the best approach for your needs.

When using data visualisations, it is important to remember that these are complementary resources and do not replace the data itself. It is therefore important to provide access to the data used for creating the visualisations in a clear and accessible manner.

3.2.1 Is this enough for you?

This level is most appropriate for datasets that include relations to other resources but where the data is not readily machine-readable or actionable. Any dataset that can be linked to publications, source data, versions of the dataset and other related resources should be included in this level. Datasets included in this level have rich and meaningful metadata, but the data itself cannot be re-used by machine agents without human intermediation. In other words, the data contents do not fulfil the principle of Interoperability. This can be the case when data is exposed in a non-standardised manner and where the scope of the project did not include resources for exposing the data via standardised protocols and formats that are not optimised for actionability (e.g. APIs) and readability (e.g. JSON, XML, RDF).

If your data follows a common or well-defined standard, you should consider aligning to a higher level of FAIRness. The level is likewise scarce if your data makes use of (FAIR) vocabularies to define data elements or if the data itself directly contains links to other data.

Box 3. Referencing URLs

When linking to resources external to the JRC Data Catalogue, it is important to verify their persistency. Otherwise, provided links may result in a 404-resource-not-found error. It is common to use external URLs to point to landing pages or additional documentation. If you have any doubt regarding the persistency of these URLs, remember that all EU Institutions web pages are periodically archived by the [Publication Office's WebArchive service](#). These URLs are guaranteed to remain stable and so you may prefer to link to an archived version of the web page if not sure which redirection policy is in place for the specific domain. It is possible to request the service to crawl a specific site you are managing, for example, when a domain change is being programmed or if no archived version exists. However, the solution is designed to capture static HTML pages. Dynamic pages that contain forms, scripts or other types of server-side services will not be archived in a way that preserves the site's original functionalities.

3.2.2 Examples

- **Pagano, Andrea; Di Girolamo, Francesca (2023): European Union Banking Sector Statistics (2007). European Commission, Joint Research Centre [Dataset] PID: <http://data.europa.eu/89h/jrc-eubss-eubss-2007>**

The dataset assesses the impact of EU banking sector legislation on reduction in public financing costs. The linked publication explains how the dataset was produced and cites input data sources. Other resources linked to the dataset provide a file including additional information on which source data was used and a bar chart visualising the main findings of the assessment. ORCID identifiers of the authors of the dataset are provided to reduce ambiguity and support linking of data. The dataset is part of a larger collection of datasets on the same topic.

- **Garaffa, Rafael; Ordonez, Jose; Vandyck, Toon; Weitzel, Matthias (2024): Baseline GECO 2022. European Commission, Joint Research Centre [Dataset] doi: 10.2905/DF6CFD52-EE0C-4647-A2B3-5FA56B8B5AB0 PID: <http://data.europa.eu/89h/df6cfd52-ee0c-4647-a2b3-5fa56b8b5ab0>**

The dataset provides Multi-Regional Input-Output (I-O) tables which are used for creating the Baseline of the Global Energy and Climate Outlook (GECO). In it, baseline projections of energy demand and greenhouse gas emissions are provided based on Global Trade, Assistance, and Production database. In addition to the persistent identifier assigned to the dataset via the JRC Data Catalogue, a DOI is declared as an equivalent identifier of the dataset, supporting its persistent identification and dissemination of its metadata via the DOI citation channels. The dataset is included in the GECO collection, facilitating the discovery of related datasets. A landing page and a publication are also linked to the dataset, defining qualified relationships the dataset has with additional resources that provides complementary information useful for gaining knowledge of how the dataset was produced and how it is used in compiling the assessment. As in the previous example, ORCID identifiers are provided to increase data quality of the attributes related to the authors of the dataset.

3.3 FAIR go

Dataset is:

- Published in the data catalogue
 - ☑ Rich metadata is provided
 - ☑ Metadata contains links to other resources
- Data is partially FAIR:
 - ☑ Exposed using non machine-readable community data model
 - ☑ Terms refer to a data dictionary (either machine-readable or not)

The previous sections analysed maturity levels of the Findability and Accessibility principles. These are mostly sustained by providing useful metadata, which was the focus of the first two levels. To assess Interoperability and Reusability of data, it is necessary to go beyond the descriptive elements and look more closely at the data content itself.

Interoperability is a multi-factor consideration, and we can identify different degrees, or levels, in which different systems can interoperate. Therefore, the "solution" to Interoperability should address different issues, from systems being able to communicate among themselves via technological networking infrastructure to being able to interpret the exchanged messages consistently.

Assuming that networking is addressed, for instance, by using the widespread Web technologies, the next crucial aspect of interoperability is that interoperable systems can process the encoding format in which the messages are exchanged. For instance, if one system is offering data in XML, but the other system can only process the JSON format, their interoperation is not going to succeed. Once they are able to process the encoding format, they need to interpret the structure in which the data are organised up to the point where they are able to interpret the meaning of the exchanged data. Given this scale of interoperability, and the fact that not all data can have the same encoding format, due to their inherent characteristics (e.g. structured data differs from images, video, etc.), and, even when they share the same encoding format, their structures and meaning (semantics) may need to be different. Therefore, from the FAIR Principles' perspective, it is important to be clear and explicit on how data are encoded, structured and what they represent, i.e. the meaning. Naturally, we can also improve interoperability and re-use if, for similar types of characteristics of data, we agree on adopting the same domain-relevant community standards.

At this level, we expect data to be encoded and structured using relevant community standards and these choices to be explicitly declared as part of their metadata. If no relevant standards exist, the data models and vocabularies used need to be explicitly mentioned and defined. At this level, the models representing data structure and semantics are not required to be machine-readable. For instance, data dictionaries and other types of textual descriptions are acceptable.

Box 4. Distinction between data and metadata

A common definition of metadata is data about data. Undoubtedly, metadata is data. However, we can also have metadata about other types of entities such as software, services, people, organisations, etc. When considering data as the type of entity being described, we can say that, while data constitutes the actual content or observations within a dataset, metadata provides essential context, structure, and descriptive information about the data itself, facilitating its organization, discovery, and interpretation.

To illustrate the distinction, consider a library catalogue as an analogy. The books on the shelves represent the data—these are the tangible, substantive content that users are interested in accessing. Meanwhile, the library catalogue entries provide metadata—they offer descriptive information about each book, such as the title, author, publication date, and subject categories, enabling users to locate and evaluate the books more efficiently. However, it should be noted that books also contain metadata as an integral part of themselves. Consider the book cover, frontispiece or copyright page where information such as title, authors, editor, edition and copyrights are provided. This example shows that a precise distinction between data and metadata is not always clear. In fact, to improve discoverability and reusability of data it may be useful to replicate some elements from the data (the books) in the metadata (the catalogue). For instance, the OFFICAIR dataset contains data representing the concentration of 33 pollutants measured in the indoor air of 37 modern office buildings in eight European Countries. In this case, adding the list of pollutants that are present in the dataset to the metadata is useful for potential users in considering whether this dataset is relevant for them.

The question that immediately emerges is: which elements of or about my data should I have in my metadata? The FAIR Principles can help us answer this question by considering which properties can be used to facilitate the findability, accessibility, interoperability and reusability of the data. For example, properties such as title, description and keywords can be used for findability while, license, used community-relevant standards and detailed provenance will facilitate its reusability.

3.3.1 Is this enough for you?

Alongside rich metadata, datasets included in this level provide a basic level of data interoperability. This implies that data is exposed following the same pattern, or model, across instances. The models and terminologies used by these datasets are explicit and well defined. These are often defined in human-readable terms that cannot be readily interpreted by a machine agent. Conversely, if your dataset follows a common data model that is machine-accessible, you should consider aligning to the next two levels of FAIR maturity. Similarly, if your data uses controlled FAIR vocabularies to denote terms, it should be aligned to higher levels of FAIR maturity.

3.3.2 Examples

- a) **Munn, Sharon; Landesmann, Brigitte; Dumont, Coralie (2016): Covalent Protein binding leading to Skin Sensitisation. AOPWiki [dataset], <https://aopwiki.org/aops/40>.**

The Adverse Outcome Pathway Wiki (AOP-Wiki) serves as the primary repository of qualitative information about adverse outcome pathways, that are intended to provide evidence for demonstrating and assessing causality between measurable toxicological mechanisms and

human or environmental adverse effects. The AOPs made available through the AOP-Wiki are results of collaborative efforts with contributions from various stakeholders, including research, regulators and industry professionals. Each AOP has a unique URI, which makes it straightforward to locate and reference specific AOPs. Additionally, the AOPs are described with rich metadata including information about authors, point of contact, modification history and description. The data is freely and openly available without any login credentials or other barriers. Besides the human-oriented HTML representation, the data is also provided in XML and JSON formats, facilitating their processing by computational agents. Concepts and terms used to describe AOPs and related Key Events and Key Event Relationships are part of controlled lists of terms leading to enhanced interoperability.

The model for creating or consuming AOP wikis entities is described in a dedicated manual (the Developers' handbook: <https://aopwiki.org/handbooks/4>). As the model evolves, previous versions of the model are archived and remain accessible to ensure compatibility. Although the model is not available in machine-readable form, its structure supports translation to machine-readable formats (namely RDF serialisation cfr. <http://doi.org/10.1089/aivt.2017.0017>). A recent deliverable of the Partnership for Assessment of the Risks from Chemicals (PARC)¹¹ provides a specific example of how relevant AOP-wikis were transported to a Labelled Property Graph to leverage querying functionalities.

- **Frank, Stefan; Havlik, Petr; Stehfest, Elke; van Meijl, Hans; Witzke, Peter; Pérez Domínguez, Ignacio; van Dijk, Michiel; Doelman, Jonathan; Fellmann, Thomas; Koopman, Jason F.L.; Tabeau, Andrzej; Valin, Hugo (2018): AgMIP – Agricultural non-CO2 emission reduction potential in the context of the 1.5 °C target. European Commission, Joint Research Centre (JRC) [Dataset] PID: <http://data.europa.eu/89h/5a06cad1-6c12-4d17-b008-4b58956ec3d8>**

The dataset provides an assessment of the agricultural non-CO2 emission reduction potential, while defining relative variables essential for the assessment. Four impact assessment models and predefined scenarios were used to project impact outcomes. The impact outcomes are made explicit and structured for machine-readability. The dataset is accessible as open data in tabular format, available for bulk download as a CSV file or through an interactive interface. Registered users can also access data and metadata via a REST API. Each column header of the tabular data, representing a distinct “dimension,” is documented in an accompanying text file. Additionally, value lists used within the data are downloadable as separate CSV files via a dedicated web form. Additional context and insights for the correct use and interpretation of the data is made available via links to the associated publication. A general description of the open data principles of the DataM collection and of its API, is available at: <https://doi.org/10.2760/278135> (chapter 7 and Annex III). Although the model used for disseminating the data cannot be considered a “community standard,” also considering the specific scope of the data, the data provided facilitates machine-driven re-use with human

¹¹ Saurav Kumar, Deepika Deepika, Karin Slater, Vikas Kumar, AOPWIKI-EXPLORER: An interactive graph-based query engine leveraging large language models, Computational Toxicology, Volume 30, 2024, 100308, ISSN 2468-1113, <https://doi.org/10.1016/j.comtox.2024.100308>.

intervention. This facilitation is achieved through explicit linking to specific data elements and the provision of downloadable resources for key terminologies in the dataset.

The monitoring of access and re-use of datasets in the DataM collection (see: https://datam.jrc.ec.europa.eu/datam/mashup/DATAM_USAGE_STATISTICS/) is a noteworthy approach that supports the Re-use element of the FAIR principles.

- **Barrero, Josefa (2020): OFFICAIR - Indoor air pollution in modern office buildings. European Commission, Joint Research Centre (JRC) [Dataset] PID: <http://data.europa.eu/89h/eaf0191d-456c-4c2c-8b30-1ecae9b2f999>**

The OFFICAIR dataset was produced in the FP7 framework (concluded in 2014) as part of a project that had the objective of establishing a framework to provide new knowledge in terms of databases, modelling tools and assessment methods towards an integrated approach in assessing the health risk from indoor air pollution. The project aimed at identifying possibilities for reducing negative health effects from exposure to indoor air pollutants in modern office spaces¹². The dataset's metadata provides links to the relevant publication¹³ providing additional information with regards to the dataset. Notably, the OFFICAIR dataset exemplifies an early endeavour to increase data interoperability by means of applying a common data model to various measurements data in a manner that is compatible with related datasets, namely the AIRMEX (European Indoor Air Monitoring and Exposure Assessment Project) and BUMAC (Building material and Consumer product) databases. In addition to the cited data model, the dataset uses controlled terms to denote chemical elements that are monitored. Currently, the OFFICAIR dataset is available via the IPCHEM portal¹⁴. The portal was launched later in the year of the project's conclusion date and includes a larger collection of measurements data. For this aim, IPCHEM further defined a schema for the harmonised data submitted to the portal. The common data schema used to encode sampling data in a unified manner, is available as a document for IPCHEM portal data providers. Work is ongoing to render the data models accessible online by machine agents. A JSON schema – which is expected to be defined by the data providers for each dataset – complements the prior model with machine-readable data contained in the original sampling data.

¹² Bartzis, J.G. & Reina, Vittorio & Goelen, Eddy & Mandin, Corinne & PederWolkoff, & Terry, Andrew & Carrer, Paolo & De Oliveira Fernandes, Eduardo. (2013). OFFICAIR Final Workshop – Results Summary. [10.13140/RG.2.2.15781.70881](https://pubsy.jrc.ec.europa.eu/int/workflow/pages/output-detail/87783).

¹³ Public documentation available via landing page at: <https://ipchem.jrc.ec.europa.eu/#showmetadata/OFFICAIR>. Further documentation – including a description of the data model – is not available for public consultation but is recorded in PUBSY at: <https://pubsy.jrc.ec.europa.eu/int/workflow/pages/output-detail/87783>

¹⁴ <https://ipchem.jrc.ec.europa.eu>

3.4 FAIR share

Dataset is:

- Published in the data catalogue
 - ☑ Rich metadata is provided
 - ☑ Metadata contains links to other resources
- Data is mostly FAIR:
 - ☑ Exposed using machine-readable local data model
 - ☑ Terms refer to machine-readable data dictionaries

The “FAIR share” level denotes datasets that have rich metadata describing how the data was produced (provenance), and how data contents are exposed using machine-readable data models. Furthermore, data of these datasets use concepts that can be interpreted in a consistent manner across systems and domains and may also include reference to other data or metadata. Using machine-readable models and standardised terminology not only enables interoperability, as described in the previous level, but also allows machine agents to interpret the data correctly with no need for human mediation.

This level builds on the previous one by requiring not only that datasets have data models, but also that these models are machine-readable, i.e. they should be represented in a format that provides a structure to describe the data model in a manner that software systems can parse and interpret. Examples of such formats include JSON Schema¹⁵, XML Schema (XSD)¹⁶, SHACL¹⁷, ShEx¹⁸, Entity-Relationship notation and machine-readable UML representations.

Datasets included in this level pay special attention to documenting the lineage of the data. For example, source data and processes used to transform it into the published dataset are made explicit. These processes may include codes, algorithms, workflows or, generally speaking, any activity taken by any agent during the dataset’s lifecycle, which resulted in its transformation.

Datasets in this category represent the highest FAIR level for resources that cannot be fully exposed according to existing community standards. It is often the case, when data is collected to represent innovative observations, that no existing model can capture all data. For this reason, publishers may need to extend existing models or even create new ones. When the adoption of these models is clearly stated, and the models can be accessed by machine agents, then interoperability is supported. To move to the next level, these models will need to be endorsed by a larger community.

¹⁵ <https://json-schema.org/specification.html>

¹⁶ <https://www.w3.org/TR/xmlschema11-1/>

¹⁷ <https://www.w3.org/TR/shacl/>

¹⁸ <http://shex.io/shex-primer/>

Box 5. Interoperability: FAIR vocabularies

Controlled FAIR vocabularies play a pivotal role in enhancing interoperability within various domains. These structured sets of terms and concepts, tailored to specific communities or disciplines, serve as a lingua franca that facilitates seamless communication and data exchange among diverse stakeholders. By establishing common terminology and definitions, these vocabularies minimise ambiguity and misunderstanding, ensuring that information is interpreted consistently across different systems and applications.

To be FAIR, vocabularies need to follow the FAIR Principles. To be controlled, vocabularies need to adhere to a governance policy. When these prerequisites are met, vocabularies sustain the development of interoperable systems and tools by providing a common framework for data integration and exchange. More complex vocabularies like thesauri, taxonomies and ontologies facilitate use of sophisticated data analysis and visualization techniques and can be leveraged to unlock deeper insights and discoveries that can transcend the original context for which they were created.

To give one simple example of the benefits using FAIR vocabularies convey, consider using the URI "<http://publications.europa.eu/resource/authority/country/CSK>" to denote the former Czechoslovak Republic. This URI resolves to a [SKOS](#) vocabulary Concept that allows machine-agents to understand its meaning, regardless of the language used. Furthermore, the deprecated status of the term, as well as the period during which the country name was in use and the terms which replace it are clearly stated. Equivalence of the URI to terms from other concept schemes (e.g. ISO 3166, UNSD et al.) is also included. In other words, using a codified concept rather than a natural language term intrinsically supports interoperability across domains and instances.

To see if relevant standards and vocabularies are available for use for your data, the [FAIRSharing portal](#) provides lists of data and metadata standards, as well as repositories and policies. The portal has a user-friendly search interface, and a knowledge-graph can be displayed for each described resource. The [EU Vocabularies catalogue](#) is the recommended access point to vocabularies that are specifically relevant to the work of EU Institutions and beyond.

If no suitable vocabulary exists to express the concepts your data denotes, available tools and frameworks in the context of JRC and the European Institutions can support you in the creation and management of controlled FAIR vocabularies. JRC's [R3gistry](#) software and the publications office's [VocBench](#) and [ShowVoc](#) platform can be used to create, explore, and update vocabularies. The publications office can also act as publisher and curator of vocabularies authored by DGs and Agencies of the Commission or other EU Institutions.

3.4.1 Is this enough for you?

Datasets at this level include standardised metadata that denotes their provenance and reference to other resources. Furthermore, the data described by the metadata is made available via standard protocols and follows defined models, which are in turn accessible for machine-agents. Using the analogy of the books and library catalogue in the [data and metadata tip-box](#), this means that not only are the books of this category registered in the library catalogue, but also that the metadata included within the data contents (i.e. the book cover and frontispiece in the analogy) as well as the data elements themselves (chapters, illustrations, references etc.) are explicitly marked and defined. Moreover, data elements are exposed using standardised terminology whenever possible.

In essence, this level is reserved for datasets with a very high level of FAIRness, but maintaining that level requires sustained efforts. For example, a data management plan would need to accompany these datasets to determine what safeguards need to be put in place to ensure models and codes used by the datasets are kept pertinent and accessible. If your dataset uses models and vocabularies that may be relevant to a larger community of users, taking actions to promote these to community-standards can improve the FAIR maturity level of the dataset. However, moderating communities and standards may require additional efforts and can prove difficult to sustain if not already part of your mandate. The [JRC hub for standardisation](#) can guide you in this journey. The [EC Standardisation Knowledge Base](#) contains additional information and contacts

3.4.2 Examples

- b) European Commission, Joint Research Centre (JRC) (2024): Flood in Saarland region, Germany (2024-01-02). [Dataset] PID: <http://data.europa.eu/89h/84265e62-3a2a-5b50-8d78-327d26011122>**

This dataset is part of the Copernicus Emergency Mapping Service for Rapid Mapping. It encompasses data collected during the activation of the flood alert for the Prims river. A summary of the event and relevant information is available via the landing page. The data is exposed following geospatial community specifications. A complementary model is used to expose data concerning the rapid mapping service. The complementary Rapid Mapping schema is accessible via:

https://rapidmapping.emergency.copernicus.eu/static/cems_rapidmapping_openapi_specs.yaml.

Data contents collected for this alert are available in KMZ and ESRI Shapefile formats. These formats are widely used standards for geospatial data and include vector files with geospatial metadata. A technical report presenting information in human readable format (<https://doi.org/10.2760/29876>) is available via the landing page. Although the dataset does not employ “traditional” vocabulary terms, it utilises geospatial coordinates following community-standard models. This approach ensures that terms (in this instance, spatial coordinates) are consistently interpreted by machine agents to support interoperability within and across domains. All products resulting from Rapid Mapping activations are made publicly available and accessible free of charge.

- c) European Commission, Directorate-General for Financial Stability, Financial Services and Capital Markets Union (2016): Consolidated list of persons, groups and entities subject to EU financial sanctions. [Dataset] PID: <http://data.europa.eu/88u/dataset/consolidated-list-of-persons-groups-and-entities-subject-to-eu-financial-sanctions>**

The Consolidated financial sanctions file provides a list of entities and individuals that are subject to financial sanctions mandated by the European Union. The dataset is published in a machine-readable format, to facilitate the dataset’s re-use and integration in relevant apps and platforms (for example with other sanctions datasets, as done by the platform). It is one of the most popular datasets on the European data portal, data.europa.eu, which is used by various stakeholders such as companies, law-enforcement authorities, and journalists. The file is available in three different formats, namely PDF, CSV, and XML. The latter format is provided alongside an XML schema (accessible via: <https://webgate.ec.europa.eu/fsd/fsf#!/files> - EU Login

required) to support parsing and interpretation of the data by machine agents. The schema, to some extent, may be considered a community standard, inasmuch as it is used by all EU member states. However, it presents some differences with regards to similar data models used by other countries and institutions (cfr. the [UN sanctions list](#)¹⁹). The XML schema sustains a correct alignment of the EU model to external ones.

To maximise exploration of the data, an interactive sanctions map is available primarily for human users. An additional app is available for exploring financial sanctions and travel bans via alternate dimensions (see: [Dashboard | EU sanctions tracker \(europa.eu\)](#)).

¹⁹ scsanctions.un.org/resources/xml/en/consolidated.xml. The schema name space is: <https://www.un.org/sc/resources/sc-sanctions.xsd>

3.5 FAIRest of them all

Dataset is:

- Published in the data catalogue
 - ☑ Rich metadata is provided
 - ☑ Metadata contains links to other resources
- Data is FAIR:
 - ☑ Exposed using machine-readable community data model
 - ☑ Terms refer to machine-readable data dictionaries

By “FAIRest of them all” we indicate the highest level of FAIRness to which JRC data may aspire. Datasets at this level are identified by persistent identifiers and are comprehensively described in accordance with agreed-upon standards. The data itself is formatted in adherence to declared standards that are endorsed by the domain community and are machine-readable. Terms and concepts contained in the data are exposed using FAIR vocabularies and have qualified relations to other terms or objects. Ideally, data contents of datasets at this level can be consumed by machine agents via data services. In other words, datasets belonging to this top level can be readily re-used by humans or machines operating in the relevant domain and, to a large extent, can be readily re-used by agents of other domains thanks to the use of standardised protocols, models, and semantics, many of them expressed in machine-actionable formats.

However, as standards, technology and re-use expectations are dynamic factors, having reached this level at some point does not guarantee it will be maintained over time. To give one example from the geospatial domain, ISO 19115/CSW (and the INSPIRE specification extension) may be replaced or extended by the STAC specification or GeoDCAT-AP for geospatial data catalogues. With regards to encoding of geospatial data, standards like GML and GeoJSON may also be updated or become deprecated over time, which will result in reduced interoperability.

In summary, the “FAIRest of them all” category represents a commitment to the highest standards of data findability, accessibility, interoperability, and reusability. While reaching this level is a significant milestone, it also underscores the need for continual vigilance and adaptability in response to evolving standards, technologies, and best practices within specific domains. By embracing dynamic factors and remaining responsive to changes in the data ecosystem, we can strive to maintain the highest level of FAIRness and ensure the enduring value and utility of our datasets for both current and future stakeholders.

Box 6. Linking data to other data

The examples cited in the second level, “FAIR play”, describe how metadata is linked to other metadata or data resources, but referencing other resources can also be done on the level of the data itself, for example by linking patient’s information to a given test sample. Ideally, the data also states the relation between the instances, e.g. that a test sample was *collected from* a specific patient. The solutions for encoding relationships at this level of granularity are not discussed in these Guidelines because of how context specific they are. However, any formalisation of these data elements and their relationships intrinsically means that a data model is used. The data model can define, for example, types of entities to be used for each subject and possible relationships to other entities.

In the domain of Linked Data, this is often formalised using ([OWL](#)) ontologies, which can support complex expressivity. In the biomedical domain, the use of OWL ontologies supports researchers in capturing complex and heterogeneous data of different granularity levels, while maintaining semantic and system interoperability.

The choice of any modelling approach is inevitably context specific. Whether the chosen model is expressed as a human readable specification, an XML or JSON schema or an OWL ontology and SHACL shapes, curators of data should strive to analyse and define the nature of data elements and the links between them in an explicit manner. This in turn will support coherent re-use of the produced data within and across domains.

3.5.1 Is this enough for you?

This is the top FAIR maturity level envisaged by the Guidelines. Datasets belonging to this level are easily discoverable. Their data and metadata are accessible for human and machine agents in keeping with standard protocols. Data and metadata are readily interoperable and can be re-used by designated communities of users. Although not explicitly addressed by the FAIR Principles, datasets at this maturity level generally yield reproducible results with a high level of certainty. It can be presumed with high degree of confidence that these datasets are reliable resources that are fit-for-use and fit-for-purpose (the evaluation of data quality is, however, not explicitly covered by the present Guidelines, nor by the FAIR Principles²⁰).

Achieving and maintaining this level requires substantial effort. To be able to maintain it, publishers of data should actively monitor and engage with stakeholders involved in maintaining the standards and specifications the data adheres to. To justify the effort and to predict future developments, it is useful to monitor re-use of the data and understand who are the data re-users. While an active exchange with re-users is a valuable way of gaining meaningful insight, several general approaches can assist you in doing so. For example, monitoring the number of unique views and downloads of the data and providing contact details or the possibility to include comments in relevant forums where data is published (the public JRC data catalogue includes several such features) may be available off-the-shelf and can provide a basic overview of re-use trends

²⁰ Publications Office of the European Union, *Data.europa.eu data quality guidelines*, Publications Office, 2021, <https://data.europa.eu/doi/10.2830/79367> suggests a grouping of Data Quality dimensions to relevant FAIR principles (see Figure 2).

3.5.2 Examples

- d) **Ruiz, A. (2018): Nanoindentation (single cycle) test data for Gr. 91 material at 23 °C and maximum indenter force of 100 mN, version 1.0, European Commission JRC, [Dataset], <https://doi.org/10.5290/37>**

The [ODIN Portal](#) is an early adopter of DOIs for publication of the data it hosts, namely data relating to the qualification of advanced materials for nuclear energy systems. The example citation, which is intentionally very similar to that for a traditional scientific publication, provides access to a collection of datasets, each of which is also identified by a DOI. Whereas any individual dataset may be restricted or open, the bibliographic metadata (authors, title, abstract, etc.) are always accessible, hence allowing datasets to be discovered. Where datasets are of restricted access, the ODIN Portal data citation feature supports on-demand access requests to be submitted, thereby allowing the data owner and any interested third party to discuss possible data sharing terms. In all circumstances where citable datasets are re-used, the data owner can be confident of being acknowledged in derivative works exactly because the datasets can be cited in the references section in the same way as traditional scientific publications.

DOIs are amongst the most widely used metadata schemas for identification of research data assets. The specific schema used here follows the DataCite specification (<https://schema.datacite.org>) and is machine-actionable via the DataCite API (<https://api.datacite.org>). An appointed registration authority is responsible for assigning the DOIs and in the case of the ODIN Portal, the registration service is hosted by the Publications Office of the European Union. Use of DataCite addresses the Findable and Accessible components of FAIR by promoting discoverability and specifying the legal conditions for access and re-use of the data. As for all datasets published in the JRC data catalogue, interoperability of the metadata is supported.

The data provided in the [doi:10.5290/37](https://doi.org/10.5290/37) example are created in accordance with the [ISO 14577-1:2015](#) testing standard. In turn, the data are encoded according to a corresponding data specification published as a CEN Workshop deliverable ([CWA 17552:2020](#)), thereby giving attention to the Interoperable component of FAIR. Various projects have been undertaken at the initiative of the JRC that demonstrate testing standards can provide the basis for corresponding data specifications, the derivation of which depends entirely on an examination of the vocabulary of the standard. Thereafter, reference implementations (e.g. XSD, JSON, [HDF5](#), etc.) can be implemented in accordance with the data specification. For the case in question, an XSD reference implementation is available (at <http://uri.cen.eu/cen/cwa/17552/1/xsd/iso-14577.xsd>) for encoding the data. The schema defines specific data-types to be used for each encoded value. This ensures data is used and interpreted correctly by machine agents. It can reasonably be anticipated that the methodology for deriving a data specification from a procedural standard or protocol, as suggested in the [FAIR Share](#) level sub-section, can be applied to any domain, not just engineering materials.

While data quality is not explicitly assessed by the FAIR Principles, it is nonetheless a fundamental aspect of the Re-use dimension in the sense that if data is Findable, Accessible, and Interoperable but is of low quality, then it is less likely to be re-used by relevant stakeholders. To maximise Re-use, the ODIN Portal is designed to encourage adherence to various quality assurance procedures, including adherence to standards and protocols for data

creation; adoption of data protocols that identify the data of primary interest; and data peer review.

- **Pesaresi, Martino; Politis, Panagiotis (2023): GHS-BUILT-H R2023A - GHS building height, derived from AW3D30, SRTM30, and Sentinel2 composite (2018). European Commission, Joint Research Centre (JRC) [Dataset] doi: 10.2905/85005901-3A49-48DD-9D19-6261354F56FE PID: <http://data.europa.eu/89h/85005901-3a49-48dd-9d19-6261354f56fe>**

The Global Human Settlement Layer collection includes some of the most popular datasets produced by JRC. The dataset referenced in this example gives access to georeferenced raster image files depicting distribution of building heights (as average net and gross values) across the globe. The declared scope of the dataset is to increase emergency preparedness, but its re-use extends the declared scope by far. To give one example, the dataset is used to produce population density maps by combining it with census data and distributing the values to building volumes or floors. In turn, the resulting datasets can be used to calculate building stock or for modelling of climate impact scenarios.

To maximise the dataset's discoverability, metadata is available following the W3C Data Catalogues specification (DCAT), as well as STAC - a specification for geospatial data catalogues (see: <https://jeodpp.jrc.ec.europa.eu/eu/data/stac-browser/#/collections/Landcover.JRC.GHSL.BUILT-H-R2023A.V1-0>). The data contents likewise adhere to well established community standards. The data for this citation dataset, which includes the average net building height data-package, is available as a data-package with a geoTiff file, and a geoTiff overview image (OVR). Data is also available as tiles. Each tile includes, alongside the GeoTIFF and OVR files, a CLR file that defines the semantics of the colour palette used to encode the raster images. All packages include a copy of the technical documentation describing the contents of the publication. Special attention is reserved for evaluating the quality of the data and known gaps or considerations are made explicit (for instance in the SDATA data-package). Data is downloadable using alternative coordinates systems via: <https://human-settlement.emergency.copernicus.eu/download.php?ds=builtH> where additional parameters can be used to filter the data. To understand the full scope of the GHSL dataset, it is useful to consider the complete GHSL collection where data is harmonised according to several standards in a way that enhances interoperability beyond the immediate scope of each dataset. These include, for instance, harmonised sensor data for source data obtained by different sensors using different resolutions, and harmonised statistical data per spatial areas.

4 Conclusions

Publishing FAIR data is not a static operation. Digital assets that are FAIR at a given point in time can change their status in line with technological and societal transformations. For this reason, the call to observe the FAIR Principles in the JRC, has always aimed at nurturing a cultural shift, which was, and still is, supported by relevant strategic and specific outputs and solutions. In drafting this paper, it was chosen to adopt the same strategy, not limiting the paper to pure evaluation of FAIR maturity levels but rather supporting and guiding data publishers in how to interpret and implement various indicators of FAIRness.

The Guidelines suggest a mapping of available best practices and literature commonly used to evaluate FAIR maturity to five FAIRness levels. They further indicate which available JRC solutions support each level. Elements that are managed as institutional solutions – such as assigning and maintaining persistent identifiers, managing metadata following a determined specification and managing access and interoperability of the metadata – comply with several FAIR maturity indicators and are the focus of the first two maturity levels. These solutions mostly address the principles related to Findability and Accessibility, as well as Interoperability of metadata and the conditions for Reuse.

Other elements, primarily those related to the Interoperability of data assets, are inherently domain specific and managed in a decentralised manner. For this reason, the Guidelines do not provide an all-encompassing stratagem. Instead, they list issues and advocate strategies that can help assessing and enhancing data's interoperability and subsequent reuse. The more these elements are usable by machine-agents, and the more they make use of domain relevant standards and vocabularies, the higher their maturity level is.

The work acknowledges the fact that not all datasets need to strive reaching the highest FAIR maturity level. However, data produced in the context of high impact or durable initiatives should strive to reach the top two levels by offering data that is reusable by machine agents. Such datasets, when a high level of interoperability is supported, foster reuse and dissemination on a vast scale and are more apt for reuse in (multi)disciplinary areas beyond those in which the data was originally produced.

References

- Barker, Michelle, Neil P. Chue Hong, Daniel S. Katz, Anna-Lena Lamprecht, Carlos Martinez-Ortiz, Fotis Psomopoulos, Jennifer Harrow, et al., "Introducing the FAIR Principles for Research Software", *Scientific Data*, Vol. 9, No. 1, October 14, 2022, p. 622. <https://doi.org/10.1038/s41597-022-01710-X>.
- Batista, Dominique, Alejandra Gonzalez-Beltran, Susanna-Assunta Sansone, and Philippe Rocca-Serra, "Machine Actionable Metadata Models", *Scientific Data*, Vol. 9, No. 1, September 30, 2022, p. 592. <https://doi.org/10.1038/s41597-022-01707-6>.
- Bechhofer, Sean, Iain Buchan, David De Roure, Paolo Missier, John Ainsworth, Jiten Bhagat, Philip Couch, et al., "Why Linked Data Is Not Enough for Scientists", *Future Generation Computer Systems*, Vol. 29, No. 2, February 2013, pp. 599–611. <https://doi.org/10.1016/j.future.2011.08.004>.
- Cox, Simon J. D., Alejandra N. Gonzalez-Beltran, Barbara Magagna, and Maria-Cristina Marinescu, "Ten Simple Rules for Making a Vocabulary FAIR", Edited by Scott Markel, *PLOS Computational Biology*, Vol. 17, No. 6, June 16, 2021, p. e1009041. <https://doi.org/10.1371/journal.pcbi.1009041>.
- d'Aquin, Mathieu, Fabian Kirstein, Daniela Oliveira, Sonja Schimmler, and Sebastian Urbanek, "FAIREST: A Framework for Assessing Research Repositories", *Data Intelligence*, Vol. 5, No. 1, March 8, 2023, pp. 202–241. https://doi.org/10.1162/dint_a_00159.
- Devaraju, Anusuriya, Robert Huber, Mustapha Mokrane, Patricia Herterich, Linas Cepinskas, Jerry de Vries, Herve L'Hours, Joy Davidson, and Angus White, "FAIRsFAIR Data Object Assessment Metrics", April 14, 2022. <https://doi.org/10.5281/zenodo.3775793>.
- European Commission. Directorate General for Research and Innovation, *Turning FAIR into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data*, Publications Office, LU, 2018. <https://data.europa.eu/doi/10.2777/1524>.
- European Commission. Directorate General for Research and Innovation and EOSC Executive Board., *Recommendations on FAIR Metrics for EOSC*, Publications Office, LU, 2021. <https://data.europa.eu/doi/10.2777/70791>.
- Gruenpeter, Morane, Daniel S. Katz, Anna-Lena Lamprecht, Tom Honeyman, Daniel Garijo, Alexander Struck, Anna Niehues, et al., *Defining Research Software: A Controversial Discussion*, Zenodo, September 13, 2021. <https://doi.org/10.5281/zenodo.5504016>.
- Jacobsen, Annika, Ricardo De Miranda Azevedo, Nick Juty, Dominique Batista, Simon Coles, Ronald Cornet, Mélanie Courtot, et al., "FAIR Principles: Interpretations and Implementation Considerations", *Data Intelligence*, Vol. 2, No. 1–2, January 2020, pp. 10–29. https://doi.org/10.1162/dint_r_00024.
- Mons, Barend, Cameron Neylon, Jan Velterop, Michel Dumontier, Luiz Olavo Bonino Da Silva Santos, and Mark D. Wilkinson, "Cloudy, Increasingly FAIR; Revisiting the FAIR Data Guiding Principles for the European Open Science Cloud", *Information Services & Use*, Vol. 37, No. 1, March 7, 2017, pp. 49–56. <https://doi.org/10.3233/ISU-170824>.
- Mons, Barend, "FAIR Science for Social Machines: Let's Share Metadata Knowlets in the Internet of FAIR Data and Services", *Data Intelligence*, Vol. 1, No. 1, March 2019, pp. 22–42. https://doi.org/10.1162/dint_a_00002.
- Research Data Alliance FAIR Data Maturity Model Working Group, "FAIR Data Maturity Model: Specification and Guidelines", 2020. <https://doi.org/10.15497/rda00050>.

Task Group on URIs in MARC, Program for Cooperative Cataloging. "Formulating and Obtaining URIs: A Guide to Commonly Used Vocabularies and Reference Sources." Library of Congress, 2020.

https://www.loc.gov/aba/pcc/bibframe/TaskGroups/formulate_obtain_URI_guide.pdf.

The FAIRsharing Community, Susanna-Assunta Sansone, Peter McQuilton, Philippe Rocca-Serra, Alejandra Gonzalez-Beltran, Massimiliano Izzo, Allyson L. Lister, and Milo Thurston, "FAIRsharing as a Community Approach to Standards, Repositories and Policies", *Nature Biotechnology*, Vol. 37, No. 4, April 2019, pp. 358–367. <https://doi.org/10.1038/s41587-019-0080-8>.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al., "The FAIR Guiding Principles for Scientific Data Management and Stewardship", *Scientific Data*, Vol. 3, No. 1, March 15, 2016, p. 160018.

<https://doi.org/10.1038/sdata.2016.18>.

Wilkinson, Mark D., Susanna-Assunta Sansone, Erik Schultes, Peter Doorn, Luiz Olavo Bonino Da Silva Santos, and Michel Dumontier, "A Design Framework and Exemplar Metrics for FAIRness", *Scientific Data*, Vol. 5, No. 1, June 26, 2018, p. 180118. <https://doi.org/10.1038/sdata.2018.118>.

Wilkinson, Mark D., Michel Dumontier, Susanna-Assunta Sansone, Luiz Olavo Bonino Da Silva Santos, Mario Prieto, Dominique Batista, Peter McQuilton, et al., "Evaluating FAIR Maturity through a Scalable, Automated, Community-Governed Framework", *Scientific Data*, Vol. 6, No. 1, September 20, 2019, p. 174. <https://doi.org/10.1038/s41597-019-0184-5>.

Wilkinson, Mark D, Susanna-Assunta Sansone, Grootveld Marjan, Josefine Nordling, Richard Dennis, and David Hecker, "FAIR Assessment Tools: Towards an 'Apples to Apples' Comparisons", December 20, 2022. <https://doi.org/10.5281/zenodo.7463420>.

Wittner, Rudolf, Cecilia Mascia, Matej Gallo, Francesca Frexia, Heimo Müller, Markus Plass, Jörg Geiger, and Petr Holub, "Lightweight Distributed Provenance Model for Complex Real-World Environments", *Scientific Data*, Vol. 9, No. 1, August 17, 2022, p. 503.

<https://doi.org/10.1038/s41597-022-01537-6>.

List of boxes

Box 1. Storing your data	11
Box 2. Linking to data visualisations.....	15
Box 3. Referencing URLs.....	15
Box 4. Distinction between data and metadata.....	18
Box 5. Interoperability: FAIR vocabularies	22
Box 6. Linking data to other data.....	26

List of figures

Figure 1. Schematic overview of the five FAIR maturity levels.....	9
---	---

Annex 1. FAIR maturity assessment grid

The [FAIR Data Maturity Model](#) - Specification and Guidelines was published by the RDA working group on FAIR data maturity model in 2020. The model provides a set of core assessment criteria of data's FAIR maturity level that can be used by different evaluators, regardless of the choice of specific tools or approaches. The Model enjoys large acceptance and is currently in maintenance mode after being recognised and endorsed by the RDA.

The model breaks down the rather general FAIR Principles into forty-one indicators that add context and specificity to the original FAIR Principle from which they are derived. Generally speaking, each Principle is mapped to distinct indicators for the data and the metadata parts and, when applicable, indicators assessing machine-readability and community adoption are added.

The following table provides a mapping that indicates if each distinct FAIR level defined in the Guidelines satisfies a particular RDA indicator. When the indicator is assessed as covered by the level, a "Y" is used to show this. Alternatively, "N" indicates that the level does not fulfil the requirement set by the indicator, while "P" means that the indicator is partially fulfilled.

Indicator	JRC FAIR level compliance				
	FAIR Start	FAIR Play	FAIR Go	FAIR Share	FAIRest of them all
Metadata is identified by a persistent identifier	Y	Y	Y	Y	Y
Data is identified by a persistent identifier	Y	Y	Y	Y	Y
Metadata is identified by a globally unique identifier	Y	Y	Y	Y	Y
Data is identified by a globally unique identifier	Y	Y	Y	Y	Y
Rich metadata is provided to allow discovery	Y	Y	Y	Y	Y
Metadata includes the identifier for the data	Y	Y	Y	Y	Y
Metadata is offered in such a way that it can be harvested and indexed	Y	Y	Y	Y	Y
Metadata contains information to enable the user to get access to the data	Y	Y	Y	Y	Y
Metadata can be accessed manually (i.e. with human intervention)	Y	Y	Y	Y	Y
Data can be accessed manually (i.e. with human intervention)	Y	Y	Y	Y	Y
Metadata identifier resolves to a metadata record	Y	Y	Y	Y	Y
Data identifier resolves to a digital object	Y	Y	Y	Y	Y

Indicator	JRC FAIR level compliance				
	FAIR Start	FAIR Play	FAIR Go	FAIR Share	FAIRest of them all
Metadata is accessed through standardised protocol	Y	Y	Y	Y	Y
Data is accessible through standardised protocol	Y	Y	Y	Y	Y
Data can be accessed automatically (i.e. by a computer program)	Y	Y	Y	Y	Y
Metadata is accessible through a free access protocol	Y	Y	Y	Y	Y
Data is accessible through a free access protocol	Y	Y	Y	Y	Y
Data is accessible through an access protocol that supports authentication and authorisation	Y	Y	Y	Y	Y
Metadata is guaranteed to remain available after data is no longer available	Y	Y	Y	Y	Y
Metadata uses knowledge representation expressed in standardised format	Y	Y	Y	Y	Y
Data uses knowledge representation expressed in standardised format	N	N	Y	Y	Y
Metadata uses machine-understandable knowledge representation	Y	Y	Y	Y	Y
Data uses machine-understandable knowledge representation	N	N	N	Y	Y
Metadata uses FAIR-compliant vocabularies	Y	Y	Y	Y	Y
Data uses FAIR-compliant vocabularies	N	N	P	Y	Y
Metadata includes references to other metadata	Y	Y	Y	Y	Y
Data includes references to other data	N	N	P	Y	Y
Metadata includes references to other data	N	Y	Y	Y	Y
Data includes qualified references to other data	N	N	P	Y	Y
Metadata includes qualified references to other metadata	Y	Y	Y	Y	Y
Metadata include qualified references to other data	N	Y	Y	Y	Y

Indicator	JRC FAIR level compliance				
	FAIR Start	FAIR Play	FAIR Go	FAIR Share	FAIRest of them all
Plurality of accurate and relevant attributes are provided to allow re-use	P	Y	Y	Y	Y
Metadata includes information about the licence under which the data can be re-used	Y	Y	Y	Y	Y
Metadata refers to a standard re-use licence	Y	Y	Y	Y	Y
Metadata refers to a machine-understandable re-use licence	Y	Y	Y	Y	Y
Metadata includes provenance information according to community-specific standards	N	Y	Y	Y	Y
Metadata includes provenance information according to a cross-community language	N	Y	Y	Y	Y
Metadata complies with a community standard	Y	Y	Y	Y	Y
Data complies with a community standard	N	N	P	P	Y
Metadata is expressed in compliance with a machine-understandable community standard	Y	Y	Y	Y	Y
Data is expressed in compliance with a machine-understandable community standard	N	N	N	Y	Y

Annex 2. Further reading

In recent years, work has been published in support of publication of FAIR data assets by providing measures (and tools) for assessment of FAIR maturity levels [3]. Numerous initiatives support researchers who wish to make their data FAIR: the FAIR cookbook [2] is a valid hands-on resource for identifying and understanding many specific issues one must contemplate when wishing to publish FAIR data. The [Research Data Alliance](#) (RDA) actively supports and promotes FAIR Data. The link below [6] lists past and current Work Groups sponsored by the Alliance. It is also possible to search the forum by discipline (<https://www.rd-alliance.org/rda-disciplines>). The [European Open Science Cloud](#) (EOSC) is becoming an important member of the FAIR community. One of the Strategic Objectives (SO4) of the EOSC Association is to make publicly financed research data FAIR by default²¹. Several EOSC Task Forces [7] are working towards that aim in close collaboration with the RD Alliance and the GO FAIR Foundation.

- [1] Wilkinson, M.D., M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. Bonino da Silva Santos, et al., 'The FAIR Guiding Principles for Scientific Data Management and Stewardship', *Scientific Data*, Vol. 3, No. 1, March 15, 2016, p. 160018. <https://doi.org/10.1038/sdata.2016.18>.

The FAIR Principles are the fruit of the 2014 Workshop "Jointly designing a Data FAIRport". The Principles can also be accessed via: <https://www.go-fair.org/fair-principles/>

- [2] Rocca-Serra, Philippe, Alasdair J G Gray, Alejandra Delfin Rossaro, Andrea Splendiani, Andrea Zaliani, Andreas Pippow, and Anne Cambon-Thomsen, "The FAIR Cookbook", 2022. <https://github.com/FAIRplus/the-fair-cookbook/>.

The FAIR cookbook is mostly curated by practitioners in the life science community but is a good starting point for any discipline. The resource provides many "recipes" from utilising permanent identifiers for denoting digital objects to using FAIR vocabularies.

- [3] Research Data Alliance FAIR Data Maturity Model Working Group, "FAIR Data Maturity Model: Specification and Guidelines", 2020. <https://doi.org/10.15497/RDA00050>.

RDA Working Group on [FAIR data maturity model](#) published the specification to assist evaluators of FAIRness to interpret the Principles in a homogenous manner. The specification provides a "[check-list](#)" which summarises the indicators.

- [4] Wilkinson, Mark D, Susanna-Assunta Sansone, Grootveld Marjan, Josefine Nordling, Richard Dennis, and David Hecker, "FAIR Assessment Tools: Towards an 'Apples to Apples' Comparisons", December 20, 2022. <https://doi.org/10.5281/ZENODO.7463421>.

The publication curated by the EOSC task force on FAIR Metrics examines some of the issues leading to very heterogeneous assessment results.

- [5] FAIR Evaluation tools:

- a) "The FAIR Maturity Evaluation Service", FAIRSharing. <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#/>.

²¹ Memorandum of Understanding for the Co-Programmed European Partnership for the European Open Science Cloud (EOSC) https://www.eosc.eu/sites/default/files/EOSC_Memorandum_30_July_2021.pdf p. 3

The tool is offered by the [FAIRSharing](#) community as explained in: the FAIRsharing Community, Susanna-Assunta Sansone, Peter McQuilton, Philippe Rocca-Serra, Alejandra Gonzalez-Beltran, Massimiliano Izzo, Allyson L. Lister, and Milo Thurston, “FAIRsharing as a Community Approach to Standards, Repositories and Policies”, *Nature Biotechnology*, Vol. 37, No. 4, April 2019, pp. 358–367. <https://doi.org/10.1038/s41587-019-0080-8>.

- b) “F-UJI Automated FAIR Data Assessment Tool”, FAIRsFAIR, September 22, 2020. <https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>.

The tool curated by the [FAIRsFAIR](#) project.

- [6] “Group Directory - RDA”, <https://www.rd-alliance.org/>, n.d. <https://www.rd-alliance.org/group-directory/>.

- [7] “EOSC-A Task Forces”, *EOSC Association*, n.d. <https://eosc.eu/eosc-task-forces/>.

Task Force teams of relevance to FAIR data include those grouped under Metadata and data Quality (FAIR metrics and data quality, Semantic interoperability and PID policy and implementation) and technical challenges (cfr. Long-term data preservation and technical interoperability of data and services).

- [8] “Data on the Web Best Practices”, January 31, 2017. <https://www.w3.org/TR/dwbp/>.

Cited here is the W3C recommendation for data identifiers on the Web (part of the Data on the Web Best Practices recommendation). A full list of W3C Standards is available at: <https://www.w3.org/TR/?status%5B0%5D=standard>.

- [9] Bonino Da Silva Santos, Luiz Olavo, Tiago Prince Sales, Claudenir M. Fonseca, and Giancarlo Guizzardi, “Towards a Conceptual Model for the FAIR Digital Object Framework”, in Nathalie Aussenac-Gilles, Torsten Hahmann, Antony Galton, and Maria M. Hedblom (eds.), *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2023. <https://doi.org/10.3233/FAIA231131>.

The FAIR Digital Object Framework Documentation (<https://fairdigitalobjectframework.org/>) explains the context and motivation for managing digital objects in line with the FAIR principles. It defines a baseline for facilitating interoperability while being technology-agnostic.

- [10] European Commission. Directorate General for Research and Innovation., *Turning FAIR into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data.*, Publications Office, LU, 2018. <https://data.europa.eu/doi/10.2777/1524>.

The final report and action plan on FAIR data “Turning FAIR into reality” published in 2018 contains the roadmap for developing the EOSC platform while considering each aspect of any data ecosystem in a holistic manner.

Getting in touch with the EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

Finding information about the EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

Science for policy

The Joint Research Centre (JRC) provides independent, evidence-based knowledge and science, supporting EU policies to positively impact society



EU Science Hub

[Joint-research-centre.ec.europa.eu](https://joint-research-centre.ec.europa.eu)



Publications Office
of the European Union