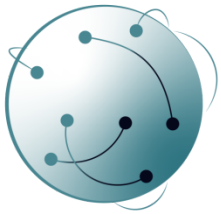
This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.



OPEN EARTH MONITOR

D2.2 Report "Status and prospect for European environmental data"

1st version

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Document control page

Project	Open-Earth-Monitor (OEMC)
Project, full title	A cyberinfrastructure to accelerate uptake of environmental information and help build user communities at European and global levels
Project number	101059548
Project start	June 1st 2022
Deliverable	D2.2
Work Package	WP2 – User-driven system design and FAIR workflows
Document title	Report "Status and prospect for European environmental data" 1st version
Version	V 0.3
Responsible author	Joan Masó Pau (CREAF)
Contributors	
Date of delivery	1st December 2022
Due date of deliverable	30th November 2022
Type	DEC —Websites, patent filings, videos, etc
Language	English
Rights	CC-BY
Status	( ) In progress (x) For review ( ) Approved
Dissemination level	Public

Version history			
Version	**Implemented by**	**Date**	**Description**
V0.5	CREAF	30 Nov 2022	Final version

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Table of contents

Figures

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Tables

Acronyms

API	Application Programming Interface (API means “Web API” in this document)
CAMS	Copernicus Atmosphere Monitoring Service
CLMS	Copernicus Land Monitoring Service
CMEMS	Copernicus Marine Service
CIFS	Common Internet File System
CKAN	Comprehensive Knowledge Archive Network
CSW	Catalogue Service for the Web
DEIMS-SDR	Dynamic Ecological Information Management System - Site and Dataset Registry
DMP	Data Management Principles
DOI	Digital Object Identifier
EEA	European Environmental Agency
EOSC	European Open Science Cloud
ETC	European Topic Center
EV	Essential Variable
GEO	Group for Earth Observation
GEO DAB	GEO Discovery and Access Broker
GEOSS	Global Earth Observation System of Systems
FAIR	Findable, Accessible, Interoperable and Reusable
FTPS	File Transfer Protocol Secured
GBIF	Global Biodiversity Information Facility
HTTPS	Hypertext Transfer Protocol Secure
JRC	Joint Research Center
RPC	Remote Procedure Call
SDI	Spatial Data Infrastructure
STAC	Spatio-Temporal Asset Catalogue
TRUST	Transparency, Responsibility, User focus, Sustainability and Technology

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Executive summary

This report discusses FAIR data principles in detail as they will be applied to the Open Earth Monitor Cyberinfrastructure (OEMC) project, followed by an inexhaustive list of environmental dataset catalogues that may be sourced for future OEMC outputs.

The FAIR principles are well known by the data community at a superficial level, but less so at the sub-principle level. We present here a mapping between the FAIR principles and the GEO Data Management Principles. The FAIR principles were designed to be applicable to data. Here, we apply them to a limited list of environmental data catalogues. From these catalogues a selection of relevant datasets will be selected and a set of metadata fields will be collected following a proposed metadata model.

A second version of this document with the title: D2.7 – Report "Status and prospect for European environmental data" final version will contain a comprehensive version of the sections initiated in this document. D2.7 (final version) should be delivered on month 30 of the OEMC project (November, 2024). D2.2 (this document) does not enumerate environmental datasets but rather catalogues of environmental data. Deliverable D2.7 will report specifically on environmental data layers.

A roadmap consisting of a list of steps, including target due dates, to complete the final version of this report is provided in section 5. The most important steps are:

(1) to establish a set of environmental variables that should be measured (due month 10), and

(2) to examine the catalogues and identify the main environmental data corresponding to the set of Essential Variables (due month 24).

1. Introduction to FAIR

In 2016, the ‘FAIR Guiding Principles for scientific data management and stewardship’ were published in Scientific Data [1]. The authors intended to provide guidelines to improve the Findability, Accessibility, Interoperability, and Reuse of digital assets. The principles emphasise not only the capacity for humans to realise the four words manually but also by computers with no or minimal human intervention. Humans increasingly rely on computational support to process data due to the increase in volume, complexity, and creation speed of data. The FAIR principles were designed to cover all scientific disciplines, including geospatial disciplines which make use of geometrical attributes and refer to environmental topics, as well as non-geospatial (e.g. astronomy or genomics).

The FAIR principles listed below relate to the ten GEO Data Management Principles [4][2]. The GEO Data Management Principles (GEO DMP, formerly known as GEOSS Data Management Principles) build on the GEO Data Sharing Principles, and detail what is required in terms of data management to facilitate data to be shared as Open Data, promptly and at minimum cost. Good data management implies a number of activities which ensure that data are discoverable

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

and accessible, that they may be understood and used, and that they are looked after in the long term. The GEO data management principles were specifically designed for geospatial and environmental data but most of them can be applied to non-geospatial data as well.

The FAIR principles are commonly known by four overarching principles (Findable, Accessible, Interoperable, and Reusable) but in fact consist of a total of 10 principles, some of which are sub-principles. These are the FAIR principles as defined by Go-FAIR.The following subsections will summarise the GEO DMP principles most relevant to the Open Earth Monitor Cyberinfrastructure (OEMC) project.

1.1. Findable

Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.

F1. (Meta)data is assigned a globally unique and persistent identifier. (GEO DMP 10)

F2. Data are described with rich metadata (defined by R1 below) (GEO DMP 4)

F3. Metadata clearly and explicitly include the identifier of the data they describe (GEO DMP 10)

F4. (Meta)data are registered or indexed in a searchable resource (GEO DMP 1)

1.2. Accessible

We need to know how data can be accessed, possibly including authentication and authorisation.

A1. (Meta)data are retrievable by their identifier using a standardised communications protocol (GEO DMP 2)

A1.1 The protocol is open, free, and universally implementable (GEO DMP 2)

A1.2 The protocol allows for an authentication and authorisation procedure, where necessary (GEO DMP 2)

A2. Metadata are accessible, even when the data are no longer available (GEO DMP 4)

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

1.3. Interoperable

The data usually needs to be integrated with other data. In addition, the data needs to interoperate with applications or workflows for analysis, storage, and processing.

I1. (Meta)data uses a formal, accessible, shared, and broadly applicable language for knowledge representation (GEO DMP 3).

I2. (Meta)data use vocabularies that follow FAIR principles (GEO DMP 3)

I3. (Meta)data include qualified references to other (meta)data (GEO DMP 5)

1.4. Reusable

To achieve reusability, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

R1. (Meta)data are richly described with a plurality of accurate and relevant attributes (GEO DMP-6)

R1.1. (Meta)data are released with a clear and accessible data usage licence (GEO DMP 4)

R1.2. (Meta)data are associated with detailed provenance (GEO DMP 5)

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

1.5. Relation with other principles

The FAIR principles target data, metadata and services but not necessarily to data repositories and data libraries. In contrast, the TRUST principles refer to digital repository trustworthiness and focus on data curation and digital preservation. TRUST is an acronym that emphasises the following five characteristics: Transparency, Responsibility, User focus, Sustainability and Technology [3].

The GEO Data Management Principles include considerations about Data preservation (GEO DMP 7), Data integrity (GEO DMP 8) and Data reprocessing (GEO DMP 9) that are not present in the FAIR principles. GEO DMP-7 and GEO DMP-8 are tasks commonly completed by data libraries and are more connected to the TRUST principles than to the FAIR principles.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

In Figure 1, we can see how the GEO DMPs can be mapped to the FAIR and TRUST principles. Other similar figures have been produced in the past but many ignored the sub-principles behind the FAIR principles presenting only a partial vision.

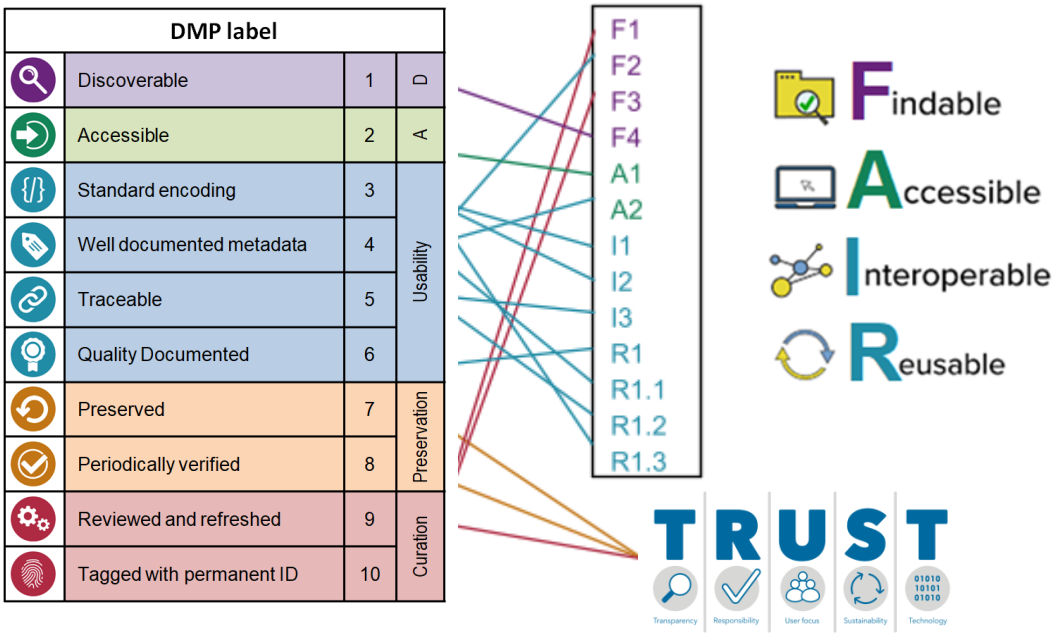


Figure 1: Relation between the DMP and the FAIR and TRUST principles (self-made)

1.6. Implementing FAIR

The report Turning FAIR into Reality [5] presents twenty-seven recommendations to implement the FAIR principles. Fifteen priority recommendations are made. These relate to the key concepts of FAIR Digital Objects and a FAIR ecosystem implemented through interoperability frameworks. The recommendations have been classified into priority recommendations and supporting recommendations.

1.6.1. Priority recommendations

Step 1: Define – concepts for FAIR Digital Objects and the ecosystem

- Rec. 1: Define FAIR for implementation
- Rec. 2: Implement a model for FAIR Digital Objects
- Rec. 3: Develop components of a FAIR ecosystem

Step 2: Implement – culture, technology and skills for FAIR practice

- Rec. 4: Develop interoperability frameworks for FAIR sharing within disciplines and for interdisciplinary Research
- Rec. 5: Ensure Data Management via Data Management Plans
- Rec. 6: Recognise and reward FAIR data and data stewardship
- Rec. 7: Support semantic technologies
- Rec. 8: Facilitate automated processing
- Rec. 9: Develop assessment frameworks to certify FAIR services
- Rec. 10: Professionalise data science and data stewardship roles and train researchers
- Rec. 11: Implement curriculum frameworks and training

Step 3: Embed and sustain – incentives, metrics and investment

- Rec. 12: Develop metrics for FAIR Digital Objects
- Rec. 13: Develop metrics to certify FAIR services
- Rec. 14: Provide strategic and coordinated funding
- Rec. 15: Provide sustainable funding

1.6.2. Supporting Recommendations

- Rec. 16: Apply FAIR broadly
- Rec. 17: Align and harmonise FAIR and Open data policy
- Rec. 18: Cost data management
- Rec. 19: Select and prioritise FAIR Digital Objects
- Rec. 20: Deposit in Trusted Digital Repositories
- Rec. 21: Encourage and incentivise reuse of FAIR outputs
- Rec. 22: Use information held in Data Management Plans
- Rec. 23: Develop FAIR components to meet research needs
- Rec. 24: Incentivise research infrastructures and other services to support FAIR data
- Rec. 25: Implement FAIR metrics to monitor uptake
- Rec. 26: Support data citation and next generation metrics
- Rec. 27: Open EOSC to all providers but ensure services are FAIR

1.6.3. Metrics and FAIR data

The FAIR Metrics group has published a design framework and exemplar metrics. They put forward a template for developing metrics, and the associated GitHub repository provides a core set of quantitative, universally-applicable metrics. Broader international initiatives in this area such as the NIH Data Commons work on FAIR metrics, the COUNTER code and the Code of Practice for research data usage metrics should also be taken into account. A proposed RDA Interest Group aims to develop a FAIR Data Maturity Model and will provide a useful international forum to define core criteria to assess the level of FAIRness.

The next deliverable (D2.7) will be inspired by these metrics to create our own set of metrics and use it for the environmental datasets.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

2. Applying FAIR principles to European Environmental catalogues

The FAIR principles are designed to be applied on individual datasets.

As previously mentioned, FAIR principles are not designed to be applied to entire catalogues. Actually, the TRUST principles could be better suited for the job, as they focus on digital repository trustworthiness, data curation and digital preservation. However, FAIR principles (as well as GEO DMPs) define a set of principles that should be continuously applied by organisations by people dealing with data, and to as many datasets as possible in the organisation. **With that in mind, it makes sense to assume that in most cases most of the data in the catalogue of a data producer will share a common ground of FAIR principles. This common ground is the focus of this first version of the deliverable.** For each presented catalogue, a generic common ground for the FAIR principles will be presented. In fact this approach is not new. The GEO yellow pages record a description of the main catalogues that are being queried by GEO geoportal through the GEO DAB. The yellow pages assign a geolabel to each of the catalogues. The geolabel in each catalogue record shows if each one of the ten DMPs is followed by the catalogue as reported by the original data provider.

3. Summary of Main Catalogues Explored

The following subsections list the most significant catalogues explored.

3.1. JRC data catalogue

The https://data.jrc.ec.europa.eu/ is the central catalogue of the Joint Research centre. It consists of 3195 datasets. They consider their data Findable, Accessible and Interoperable. In this catalogue, you can find an inventory of data produced by the JRC in accordance with the JRC Data Policy. In addition to a portal access to the catalogue through an API is possible with the following API interfaces:

- CKAN action API, RPC-style API that exposes some legacy features to existing API clients.
- ODCAT action API, RPC-style API that we recommend to use by new API clients

An open API can be used to access the portal records as documented here: https://data.jrc.ec.europa.eu/docs/

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

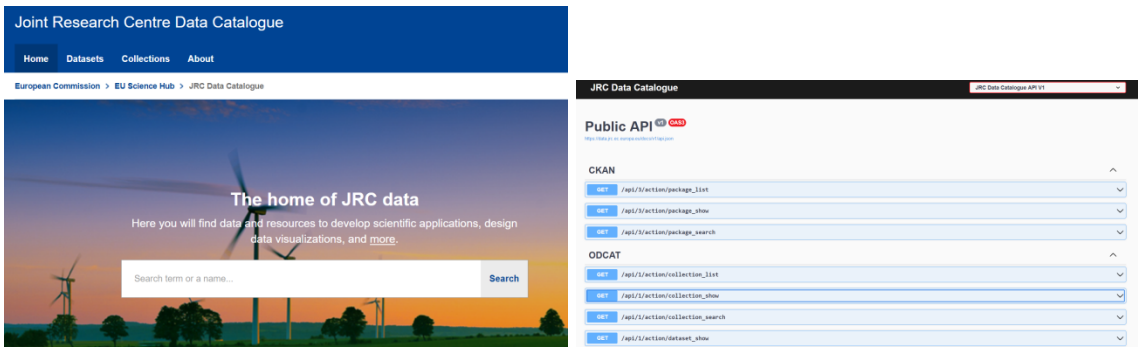


Figure 2: User interface and API of the catalogue

3.2. EEA SDI data catalogue

The https://sdi.eea.europa.eu/catalogue/ is the discovery and access service of the EEA data produced both by EEA staff and from outside (ETC, contractors, general public). It is a reference version of geospatial datasets produced, maintained or published by the EEA.

It provides a persistent archiving of the data sets (not altered over time) where datasets are to be versioned – not to be deleted or edited. Data is classified using the INSPIRE themes. A series of thematic metadata portals (which as subsets of the main SDI catalogue following specific thematic requirements) can be accessed through https://sdi.eea.europa.eu/catalogue/srv/api/sources.

Data is stored in a spatial database: PostgreSQL and PostGIS and in a File-based storage system accessible via CIFS (internally) and "Direct download" (DataShare) as well as HTTPS, WebDavS and FTPS (externally). Metadata is stored in data model called the EEA metadata profile that contains all optional elements offered by ISO 19115 and INSPIRE which are regarded as useful for EEA (https://taskman.eionet.europa.eu/projects/public-docs/wiki/Cataloguemetadata_guidelines) Additional information here: https://taskman.eionet.europa.eu/projects/public-docs/wiki/EEA_SDI

The catalogue is based in GeoNetwork 4.0.8.0 so it should be also available by GeoNetwork API as well as the CSW standard.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

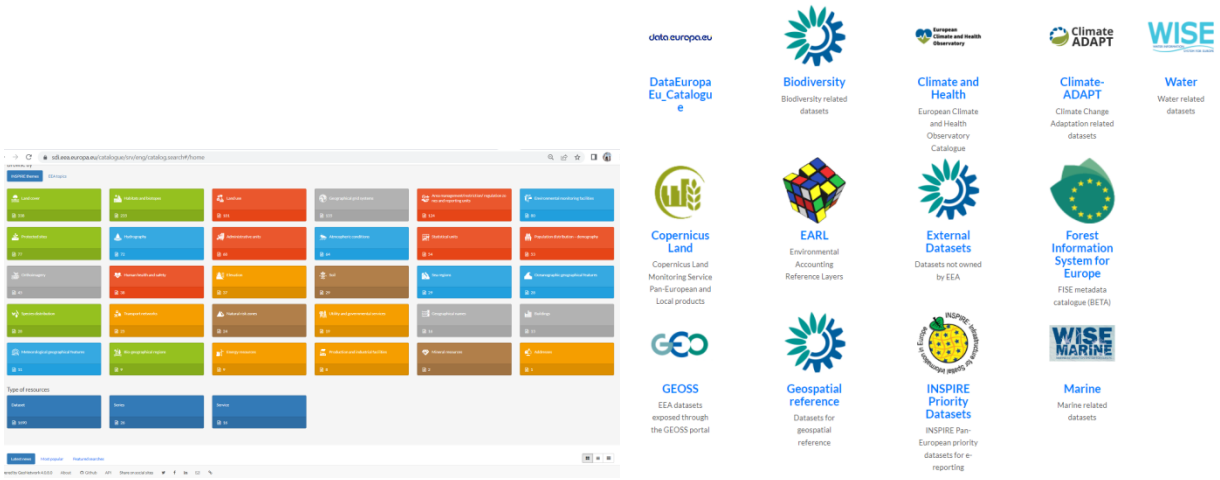


Figure 3: User interface of EEA SDI catalogue and the entry point to the thematic catalogues

3.3. Copernicus Services

3.3.1. Copernicus Atmosphere Monitoring Service

The Copernicus Atmosphere Monitoring Service (CAMS) Catalogue can be found here: https://atmosphere.copernicus.eu/catalogue#/. It offers search, visualisation and download capabilities. Regarding re-use, this is the licence model of the CAMS data: https://apps.ecmwf.int/datasets/licences/copernicus/

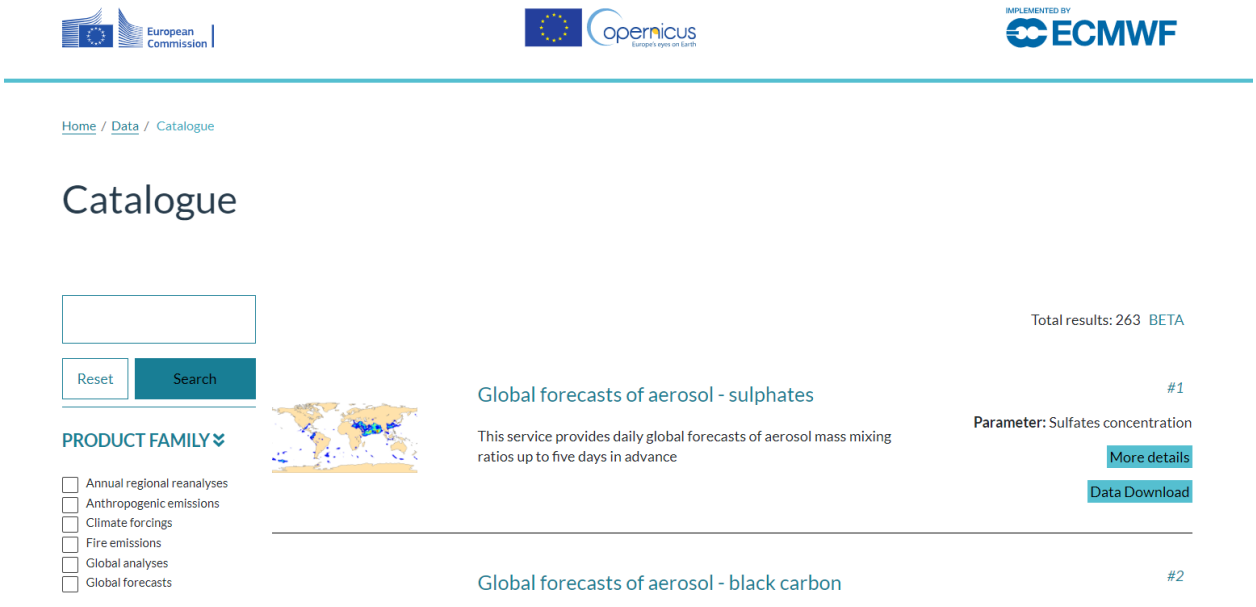


Figure 4: User interface of the CAMS catalogue

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

3.3.2. Copernicus Marine Service

The Copernicus Marine Data Store is a product catalogue that allows to download or to visualise data across nearly 15 variables including hindcast, current and forecast data. It provides 265 products. Findability and Accessibility is covered by it.

Regarding the Reuse of data, the product quality dashboard explores monthly scientific performance and product quality information updates. In it, you will find generic information on the quality of blue ocean parameters, the physical ocean and ocean circulation parameters available in the CMEMS catalogue. The quality of CMEMS products is documented and can be found on the catalogue, along with each product. One major source of deviation is the day-to-day fluctuation in the number of available observations which are used for the elaboration of this product. Hence, recent time series of the daily number of valid observations per parameter are shown to inform on potential quality drops due to this specific source of uncertainty. This information is updated on a daily basis.

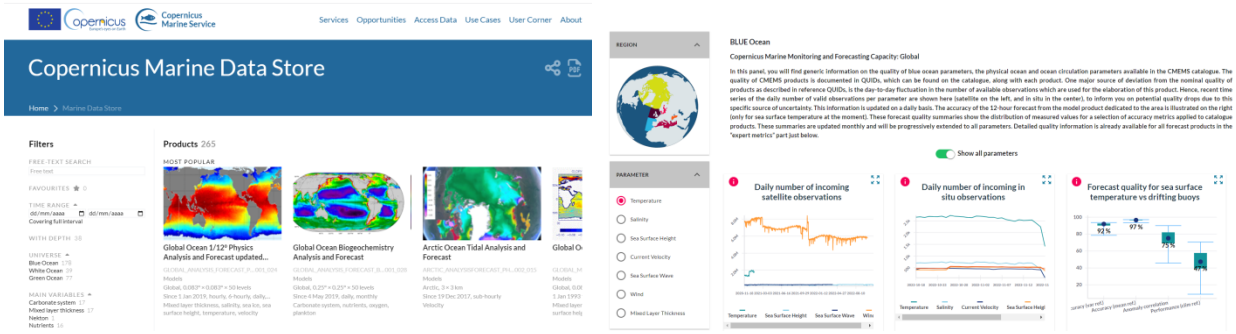


Figure 5: User interface of Copernicus Marine Data Store catalogue and quality dashboard

3.3.3. Copernicus Land Monitoring Service

In this case the service is divided into subservices

3.3.3.1. Copernicus Global Land Service

For findability there is a catalogue service that has a user interface that can be found here: https://land.copernicus.vgt.vito.be/geonetwork/srv/cat/catalog.search#/home or from machines at this endpoint.
https://land.copernicus.vgt.vito.be/geonetwork/srv/eng/csw?request=GetCapabilities&service=CSW
A WMTS and WMS view service is available here:
https://viewer.globalland.vgt.vito.be/mapcache/wmts?service=wmts&request=GetCapabilities
https://viewer.globalland.vgt.vito.be/geoserver/ows?service=WMS&request=GetCapabilities
and an access service for download here:
https://viewer.globalland.vgt.vito.be/geoserver/ows?service=WCS&request=GetCapabilities

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

For the reuse of this service, a Jupyter Notebook is provided to create and share documents that contain live code, equations, visualisations and narrative text. As such, they are ideal for step-by-step development of smaller-scale applications (e.g. prototypes, snippets), in common programming languages (Python, R, etc.), for developing one-off illustrations and for educational or training purposes. Notebook access is currently limited to the products from the vegetation and energy themes.

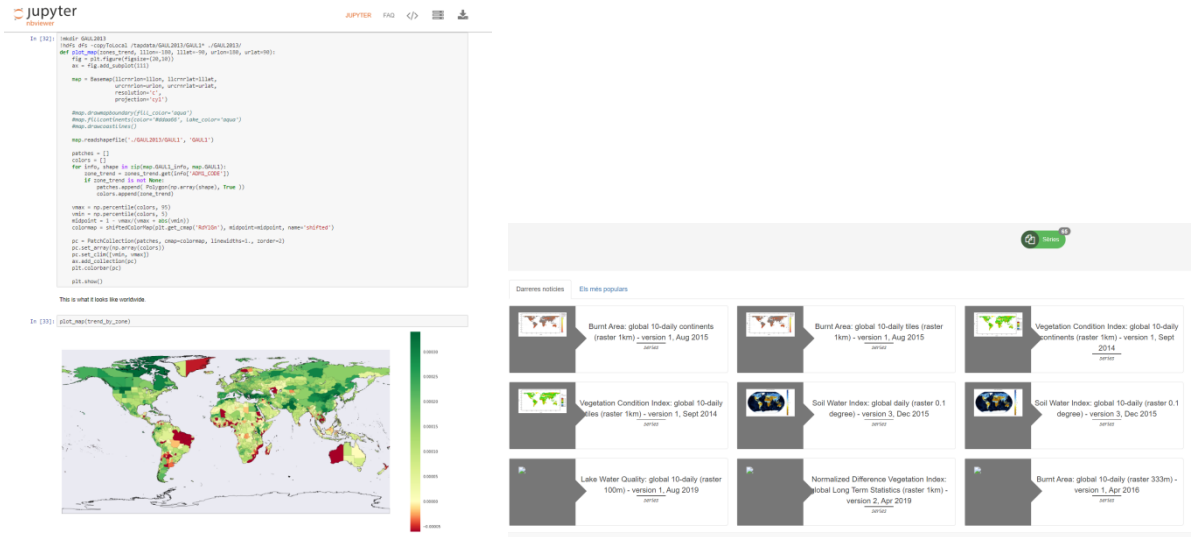
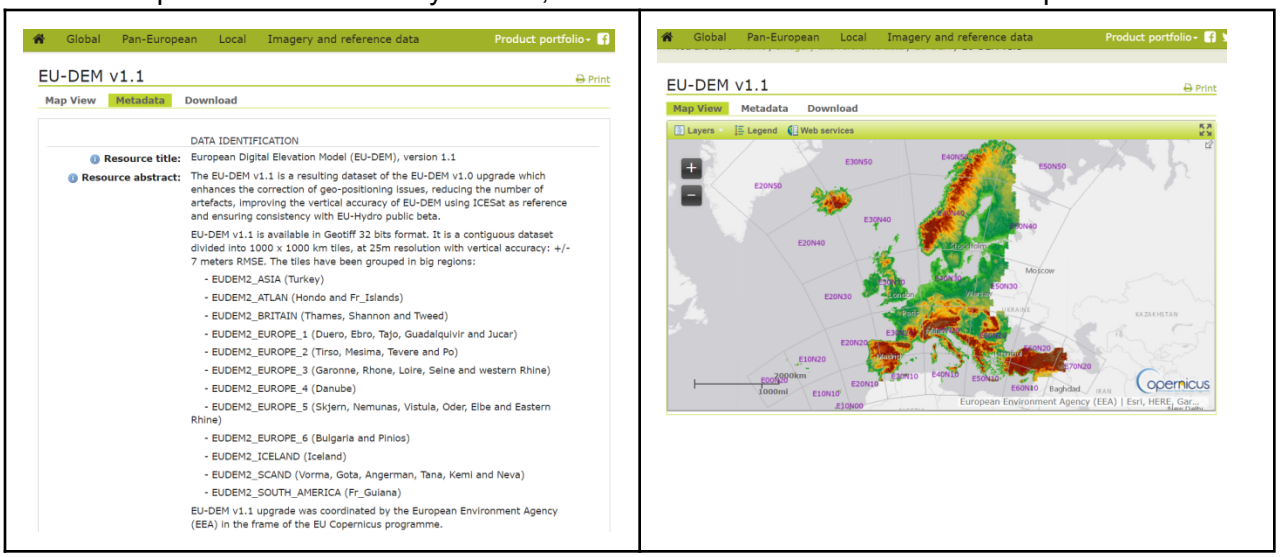


Figure 6: Jupyter notebook and user interface of Copernicus Global Land Service catalogue

3.3.3.2. Others

The rest of the Copernicus subservices for pan-European, local and reference data, do not provide a search interface and users need to travel to different pages to find what they need. For each product there is a way to view, read the metadata and download the product.



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

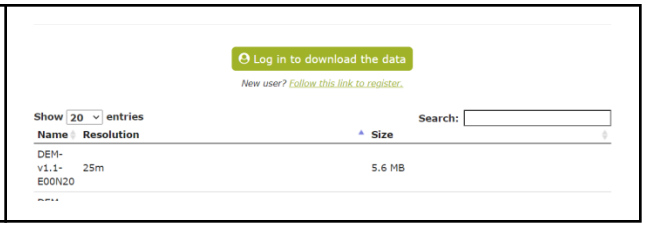


Figure 7: An example of metadata, visualisation and download pages of the pan-european CLMS

3.4. DEIMS-SDR

DEIMS-SDR (Dynamic Ecological Information Management System - Site and dataset registry) is an information management system powered by eLTER. It allows you to discover long-term ecosystem research sites around the globe, along with the data gathered at those sites and the people and networks associated with them. DEIMS-SDR describes a wide range of sites, providing a wealth of information, including each site's location, ecosystems, facilities, parameters measured and research themes. It is also possible to access a growing number of datasets and data products associated with the sites. DEIMS-SDR belongs to the service components of the emerging eLTER Research Infrastructure. While being used as eLTER site registry it also offers services to European peers and the global user communities.

There are 4 main resources in DEIMS: Sites, Datasets, Sensors and Activities.
This is essentially a discovery facility. When an asset is found (e.g. a site), access to the data on this site is not directly provided. Only the side itself allows access to the data.

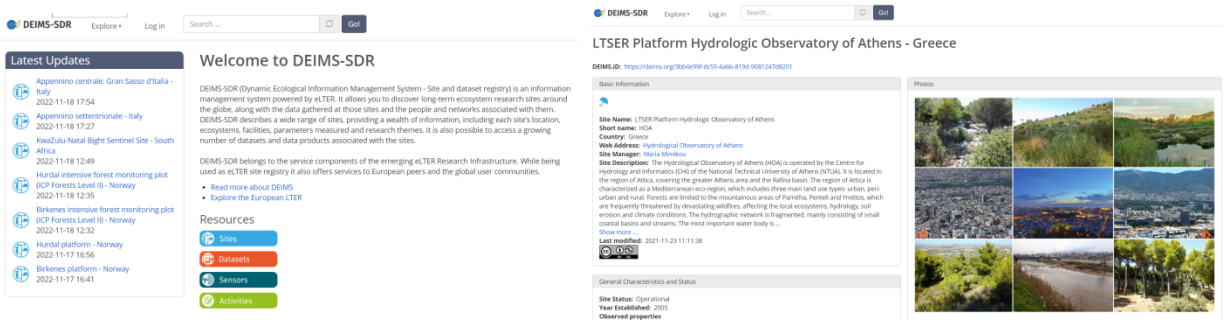


Figure 8: Home page and example of a site page in DEIMS.

3.5. GBIF

GBIF—the Global Biodiversity Information Facility—is an international network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth. The GBIF network working through the participant nodes, provides data-holding institutions around the world with common standards, best practices and open-source tools enabling them to share information about where and when

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

species have been recorded. This knowledge derives from many different kinds of sources, including everything from museum specimens collected in the 18th and 19th century to DNA barcodes and smartphone photos recorded in recent days and weeks.
The portal and API provides easy access to occurrences (observations) and species (taxonomy) datasets (occurrences datasets from organisations integrated in GBIF).

The GBIF API is a RESTful JSON based API. The base URL for v1 you should use is: https://api.gbif.org/v1/

The API is split into logical sections to ease understanding:

Registry: Provides means to create, edit, update and search for information about the datasets, organisations (e.g. data publishers), networks and the means to access them (technical endpoints). The registered content controls what is crawled and indexed in the GBIF data portal, but as a shared API may also be used for other initiatives.

Species: Provides services to discover and access information about species and higher taxa, and utility services for interpreting names and looking up the identifiers and complete scientific names used for species in the GBIF portal.

Occurrence: Provides access to occurrence information crawled and indexed by GBIF and search services to do real time paged search and asynchronous download services to do large batch downloads.

Maps: Provides simple services to show the maps of GBIF mobilised content on other sites. The site allows for finding and accessing data in an interoperable semantic framework (GBIF taxonomy) for reuse. They even provide an API to discover scientific papers reusing GBIF data.

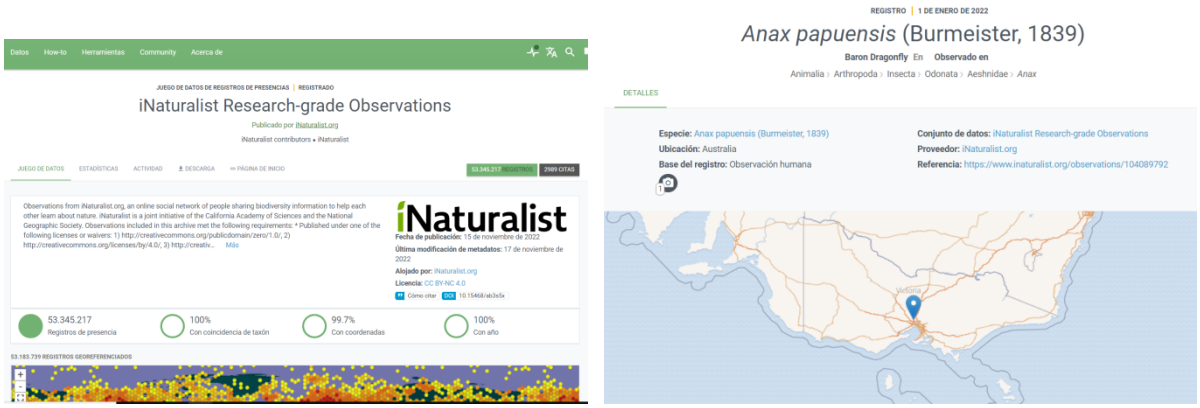


Figure 9: iNaturalist integrated in GBIF (https://www.gbif.org/es/dataset/50c9509d-22c7-4a22-a47d-8c48425ef4a7) and one iNaturalist observation integrated as a GBIF occurrence (https://www.gbif.org/es/occurrence/3455477481)

3.6. Envidat

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

EnviDat is the environmental data portal and repository developed by the Swiss Federal Research Institute WSL. They make environmental monitoring and research data accessible. EnviDat supports the user-friendly registration, documentation, storage, publication, search and retrieval of datasets from the environmental domain in Switzerland.

EnviDat provides support on managing and publishing your research data with DOIs. EnviDat is generally open for all types of WSL research data in a range of data formats. The data sets must be: Curated, Quality-controlled, Documented (reusable) and Publication-ready.

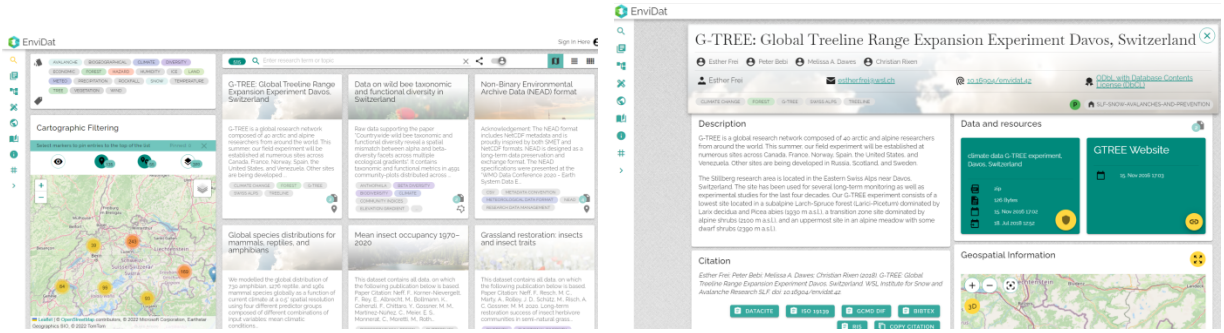
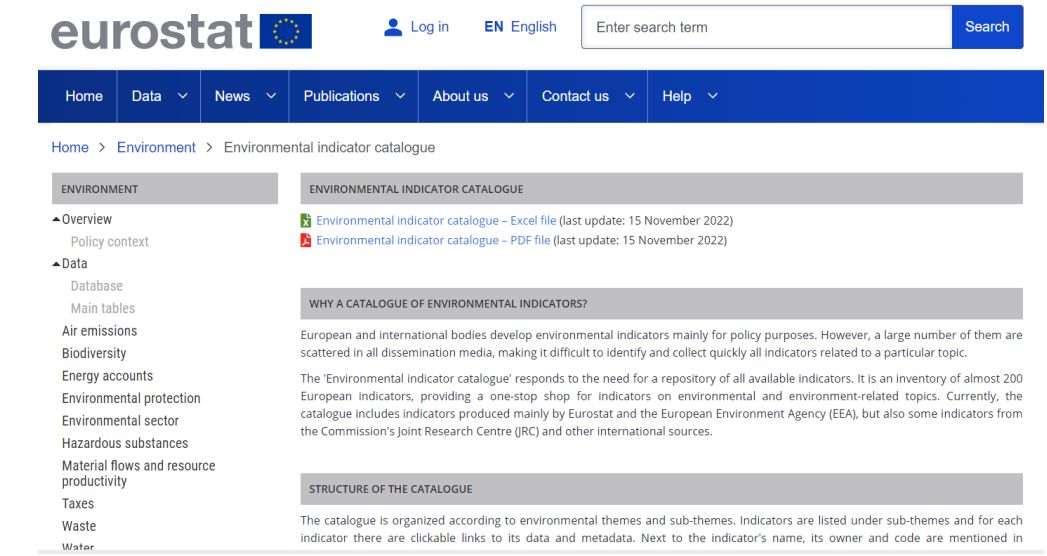


Figure 10: EnviDat portal and the view of a dataset reference accessible as a ZIP file

3.7. Environmental indicator catalogue

European and international bodies develop environmental indicators mainly for policy purposes. However, a large number of them are scattered in all dissemination media, making it difficult to identify and quickly collect all indicators related to a particular topic. The 'Environmental indicator catalogue' responds to the need for a repository of all available indicators. It is an inventory of almost 200 European indicators, providing a one-stop shop for indicators on environmental and environment-related topics. Currently, the catalogue includes indicators produced mainly by Eurostat and the European Environment Agency (EEA), but also some indicators from the Commission's Joint Research Centre (JRC) and other international sources. This is a different catalogue that does not deal with environmental data but environmental indicators. It is shared as an excel file that enumerates the indicators and presents links for where to find the indicator as well as the metadata about it. The catalogue is organised according to environmental themes and sub-themes and the indicators are listed under sub-themes. Themes, sub-themes and indicators are all listed in alphabetical order. Next to an indicator's name, its owner and code are mentioned in parentheses. An indicator metadata file provides information on the definition, properties, dissemination and methodology for developing the indicator. The catalogue can be downloaded from:
https://ec.europa.eu/eurostat/web/environment/environmental-indicator-catalogue

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.



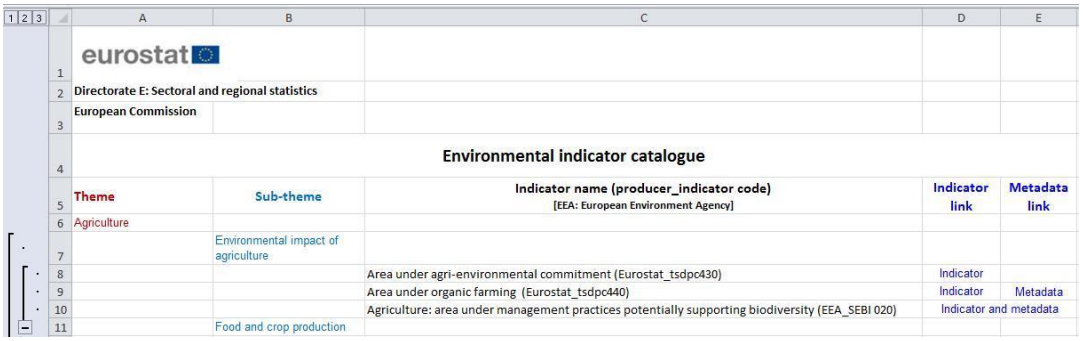


Figure 11 Eurostat Environmental Indicator Catalogue

4. Metadata model to describe environmental datasets

To catalogue environmental datasets, there is a need to define a minimum set of metadata necessary to collect and query the catalogue and to allow for comparison of similar products based on data quality parameters. This is a first draft of the minimum set of metadata fields we should collect for each dataset. These proposed metadata fields will be further discussed and validated among the WP2 partners for the final draft of this report.

Table 1: Minimum set of metadata to characterise environmental data

Field	Explanation	Example
ID		1, 2, 3 …
Geographic scope		EU, national, regional or local

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Geographic extent	Name of the locality, region, nation of geographic scope	e.g. Catalunya, Schleswig-Holstein, Lower Austria
Dataset name English **Dataset name original**		EUNIS Biotopes Austria 2018 *EUNIS Biotoptypen Österreichs 2018*
Short description / Resumé	A short description of of the dataset to get a rough impression; if not available in English to be translated or created by user	Dataset representing EUNIS classes at Level 2 and partly Level 3
Data provider	Name of the data provider and point of contact	Umweltbundesamt / Enviroment Ageny Austria (EAA)
Reference year(s) available	Years of data available	2018
Geometry	Raster/Vector (Point/Line/Polygon)	Raster
Format	Shape, Geopackage, TIFF	TIFF
CRS/EPSG	Coordinate System	ETRS89/LAEA, EPSG 3035
Spatial Resolution/Scale		10 m
Thematic coverage/Essential variable		e.g. forest ecosystems; wetlands; agricultural ecosystems…
Temporal extent	Date or interval of dates of the data capture.	2018
Time series available		No
Time series steps	Interval between data releases	unknown - not to be expected
Legal constraints		Freely accessible Full and open Members state/provider data policy Payment required
Licence	Licence applied to the product	Creative Commons Namensnennung 4.0 International
Attribution & use restrictions	Short sentence summarising both provider description and licence	Umweltbundesamt GmbH, CC-BY 4.0

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Data URL	URL of the data	https://www.data.gv.at/katalog/dataset/1d1a8d2d-cfe8-46ef-aeda-136c88726b5d
Metadata URL	URL of the data	https://www.data.gv.at/katalog/metadata/1d1a8d2d-cfe8-46ef-aeda-136c88726b5d
Reference/Technical documentation	if available	
Additional Comments		

It may also be useful to have 4 additional fields to include considerations about the support to the four main FAIR principles:

Table 2: Additional fields to characterise environmental data “FAIRness”

Findable	E.g. Persistent Identifiers	There is no guaranty of persistency
Accessible	E.g. Availability of downloading service	Downloading service requires authentication
Interoperable	E.g. Use of internationally recognised vocabularies	National habitat map classifications difficult to map to EUNIS codes
Reusable	E.g. Licences and data quality issues	Licence prevent derivative products

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Next steps

There will be a second version of this document with the title: D2.7 – Report "Status and prospect for European environmental data" final version which will contain a comprehensive version of the sections initiated in this document. D2.7 will be delivered in month 30 of the OEMC Project, or November 2024. Considering we are on month 6, this gives us 24 months to complete it. D2.2 (this document) does not enumerate environmental datasets, but catalogues of environmental data. Deliverable D2.7 will report specifically on environmental data.

5.1. Roadmap to Deliverable 2.7

The foreseen phases for the elaboration of the deliverable 2.7 are (months numbers are from the beginning of the project, month 1 = June 2022):

1. Review and extend the list of catalogues to examine (month 8);
2. Establish a set of environmental variables that should be measured. The Essential Variable framework should be the starting point for this (month 10);
3. Get the final agreement on a common set of metadata to characterise the datasets in terms of data model, spatial and temporal extent and spatial, temporal and thematic resolution as well as some other indicator for data quality (month 11);
4. Examine the catalogues and identify the main environmental data that corresponds to the set of EV (month 24);
5. Catalogue the assets considering the metadata collected fields (month 24);
6. Determine obvious data gaps in the required essential variables (month 26);
7. Consider building a STAC catalogue with the metadata that allows access to the actual data (month 28);
8. Consider a gap analysis that compares the existing assets with the project needs and requirements (month 30).

Related tasks and outputs

This deliverable collects the work completed in the first part of Task 2.2: Analysis of FAIR data status and workflows in WP2 – User-driven system design and FAIR workflows of the project Open Earth Monitor.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059548.

Cited references

1. Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18 and https://www.nature.com/articles/sdata201618
2. Revised GEO Data Sharing and Data Management Principles. 23rd GEOSS Programme Board Meeting – 21-22 June 2022 PB-23.12 https://earthobservations.org/documents/open_eo_data/PB-23-12_Revised%20GEO%20Data%20Sharing%20and%20Data%20Management%20Principles.pdf
3. *Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. Sci Data 7, 144 (2020). https://doi.org/10.1038/s41597-020-0486-7 and https://www.nature.com/articles/s41597-020-0486-7
4. GEOSS Data Management Principles Prepared by the GEO Data Management Principles Task Force Approved by Data Management Principles Task Force on April 28, 2015 https://www.earthobservations.org/documents/dswg/201504_data_management_principles_long_final.pdf
5. Turning FAIR into reality, European Commission, ISBN 978-92-79-96546-3, DOI: 10.2777/1524 https://ec.europa.eu/info/sites/default/files/turning_fair_into_reality_1.pdf