If you are a provider of data

Copernicus Services

Exported on 02/19/2024

Table of Contents

Search

About this page

Scope	This page describes the main steps needed to integrate data in the CDS Catalogue. It **does not describe** what is needed to integrate documentation, the role of the technical officers and other aspects that are very important for a successful integration but that are the scope of other wiki pages.
Intended audience	*Copernicus CDS data providers.*
Outline	*The focus of this page is on what the data provider needs to supply to the CDS team and how to do it.*
Disclaimer	*The information in this page is not guaranteed to describe exactly the actual processes which are subject to change from time to time.* *But, the CDS team intends to keep the information in this page as close as possible of the actual practices.*

1 Provider's role: summary

List of the expected contributions

The provider is expected to work closely with CDS team and the technical officer in order to resolve any issues that come up during the various stages of the publication process and afterwards. Communication is preferred through Jira ticket.
Below there is a list that had been laid out in chronological order with the main contributions.

1. Registers the Dataset: Dataset registration[1] (*Integration process*)
2. Supplies the Information document: Information document template[2] *(Integration process*)
3. Supplies manifest file (See more about manifests below at *Manifest and pseudo-manifest files*. *Integration process.*)
4. Help the CDS team member to reply to the reviewer's and Editorial Board's comments (*Review process*)
5. Provide previous existent DOIs, licences and citations associated with each part of the data (see below: *DOI, citation and licence. Review process.)*
6. After entry published in the CDS Catalogue, the data provider should keep the manifest's filename and path exactly the same for the whole duration of the contract, even when the contents of the manifest is changed. (*Complementary processes: automatic updates*)
7. The data provider is expected to help the CDS team on keeping the entry working as expected when the data provider has the knowledge and the resources to do it (*Complementary processes: Maintenance)*
8. Follows CDS procedures for deprecating data (see below: *Versions, deprecation of entries, replacement of data.Complementary processes: Maintenance*)

1 https://confluence.ecmwf.int/display/COPSRV/Dataset+registration
2 https://confluence.ecmwf.int/display/COPSRV/Information+document+template

2 How to start the integration of your data in the CDS Catalogue

JIRA ticket, Information document and manifest file

Provider's role	Description
JIRA When asked by the CDS management registers the dataset[3] and create a JIRA ticket at https://jira.ecmwf.int/servicedesk/customer/portal/5	All information concerning the creation, modification, merging, updating, deprecation, additions of of data or documentation, DOIs, Citations, etc, is supposed to be managed through the JIRA ticket.
Manifest Have a pseudo-manifest file (or a manifest file) prepared.	This is the central piece of information needed by the CDS. So important that we have a whole section about it at the bottom of this page.
Information document Attach to the JIRA ticket an "Information document" filled in with the information associated to the data that you are delivering for publication in the CDS Catalogue. The template for the information document can be found here: Information document template[4].	The information document is the starting point for the integration process. In order to arrive at an agreed draft entry to submit for review, additional inputs may be required. The document contains fields and tables that should be completed with the information relevant for your data. Guidelines are provide along those fields and tables intending to help you to understand exactly which information is required and in which format.

3 https://confluence.ecmwf.int/display/COPSRV/Dataset+registration
4 https://confluence.ecmwf.int/display/COPSRV/Information+document+template

3 Main processes in which you are expected to participate

Pre-publication process

The aim of this step is to check and agree on the main inputs for the subsequent publication process:

- path and filename conventions,
- the size of the files,
- the number of variables per file,
- where the data will be stored,
- manifest file

The CDS team expects to have access to the information in the list above as soon as possible through a JIRA ticket and at least 2 month before the actual delivery of the data.

For data for which the contents and the container is still modifiable, the CDS team expects to interact with the provider in order to influence the way the data is stored making it more suitable for the needs of the Catalogue and the needs of the Toolbox.

Publication process

The process that goes from the initial trigger of the integration of your data, to the publication of the Catalogue entry in the public Catalogue, is referred as the "publication process".
The publication process has two processes in sequence: the integration process and the review process. Your role in these two processes is summarised below.

<table>
<tr><th>Provider's role</th><th>Process</th><th>Description</th></tr>
<tr><td>Creates JIRA ticket
Supplies the Information document
Supplies manifest file
Replies to CDS team queries</td><td>Integration process (analogous to creating a draft of a paper to be submitted to a scientific journal)</td><td>
<table>
<tr><td>Inputs</td><td>JIRA ticket and Information document</td></tr>
<tr><td>Outputs</td><td>Draft Catalogue entry judged to be good enough to be submitted to review by the CDS team, the technical officer and the data provider</td></tr>
<tr><td>Work</td><td>Based on the manifest file and the Information Document, a CDS team member (or associated) creates one or more possible drafts for the future entry in the Catalogue. When agreed that the draft is good enough to be submitted for publication this process ends.</td></tr>
</table>
</td></tr>
</table>

<table>
<tr>
<td>Help the CDS team member to reply to the reviewer's and Editorial Board's comments.</td>
<td>Review process (analogous to the review process of a paper submitted to a scientific journal)</td>
<td>
<table>
<tr><td>Inputs</td><td>Draft entry</td></tr>
<tr><td>Outputs</td><td>Modified entry reviewed and approved by the Editorial Board published in the public CDS Catalogue</td></tr>
<tr><td>Work</td><td>A CDS team member (or associated) runs an internal review process to guarantee that the entry respects the CDS expectations.</td></tr>
</table>
</td>
</tr>
</table>

Post-publication processes

After publication there are frequently some additions to be made or some issues to be addressed on the entry associated to your data. These are the two main processes where you may be asked to participate:

<table>
<tr><th>Provider's role</th><th>Process</th><th>Description</th></tr>
<tr>
<td>Keep the manifest's filename and path exactly the same. Contents of the manifest is expected to change. But new additions to the contents, other than time extensions, should be discussed with the CDS team. See more about manifests below at Manifest and pseudo-manifest files.</td>
<td>Automatic updates (Updates date and time related widgets in the download form. This allows the automatic release of time extensions of data. Does not work for other widget's updates like new variables. new versions etc.)</td>
<td>
<table>
<tr><td>Inputs</td><td>Entry already published in the Catalogue
Manifest file or equivalent
Update frequency agreed
EC-Flow suite implemented</td></tr>
<tr><td>Outputs</td><td>Entry updated with new dates</td></tr>
<tr><td>Work</td><td>EC-Flow suite will read the manifest file and run CDS scripts able to recreate the download form.</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr>
<td>The data provider is expected to help the CDS team on keeping the entry working as expected. The main observed issues with published datasets are:
download form not providing the expected data
documentation tab not providing the expected documentation
mismatch between data and documentation
When this or other issues are detected by the data provider, the CDS team will be grateful if the provider could notify the CDS team using the JIRA help desk (https://jira.ecmwf.int/servicedesk/customer/portal/5)

Sometimes these issues are detected by users or the CDS team itself, in which case the CDS team will ask the data provider to help to fix the issue only when the data provider has the knowledge or the resources to do it</td>
<td>Maintenance (new programmed versions, new documentation, deprecating data, unexpected issues with the data and the documentation, licences, etc)</td>
<td>
<table>
<tr><td>Inputs</td><td>Published entry
Request for modification of the published entry</td></tr>
<tr><td>Outputs</td><td>Modified entry</td></tr>
<tr><td>Work</td><td>A CDS team member (or associated) modifies the entry as requested. The CDS team evaluates when the required modification needs agreement from the Editorial Board.</td></tr>
</table>
</td>
</tr>
</table>

4 Manifests, deprecation of data, versions, DOI, citation, acknowledgement and licence

Manifest and pseudo-manifest files

Content of the Manifest

The manifest should contain the path and the file name for every file that the CDS catalogue is supposed to provide to the users. **Nothing more nothing else.** No empty lines, no comments.
For instance:

First lines of a manifest for cmip6 data saved in ESGF

```
- path: CMIP/NUIST/NESM3/historical/r1i1p1f1/Amon/evspsbl/gn/v20190705
  ds_id: c3s-
cmip6.CMIP.NUIST.NESM3.historical.r1i1p1f1.Amon.evspsbl.gn.v20190705
  var_id: evspsbl
  array_dims: time lat lon
  array_shape: 1980 96 192
  time: 1850-01-16T12:00:00 2014-12-16T12:00:00
  latitude: -88.57 88.57
  longitude: 0.00 358.12
- path: ScenarioMIP/CNRM-CERFACS/CNRM-CM6-1-HR/ssp245/r1i1p1f2/Amon/pr/gr/
v20191202
  ds_id: c3s-cmip6.ScenarioMIP.CNRM-CERFACS.CNRM-CM6-1-
HR.ssp245.r1i1p1f2.Amon.pr.gr.v20191202
  var_id: pr
  array_dims: time lat lon
  array_shape: 1032 360 720
  time: 2015-01-16T12:00:00 2100-12-16T12:00:00
  latitude: -89.62 89.62
  longitude: 0.00 359.50
- path: CMIP/CNRM-CERFACS/CNRM-CM6-1/historical/r1i1p1f2/Amon/tas/gr/v20180917
  ds_id: c3s-cmip6.CMIP.CNRM-CERFACS.CNRM-
CM6-1.historical.r1i1p1f2.Amon.tas.gr.v20180917
  var_id: tas
  array_dims: time lat lon
  array_shape: 1980 128 256
  time: 1850-01-16T12:00:00 2014-12-16T12:00:00
  level: 2.00 2.00
  latitude: -88.93 88.93
  longitude: 0.00 358.59
```

First ten lines of a manifest file for a dataset accessible through URL addresses

head ./Integration_of_satellite-earth-radiation-budget/manifest_c3s_312b_lot1_erb_c3s_icdr_latest.txt
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/01/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201701_fv3.1.nc
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/02/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201702_fv3.1.nc
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/03/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201703_fv3.1.nc

http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/04/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201704_fv3.1.nc
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/05/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201705_fv3.1.nc
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/06/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201706_fv3.1.nc
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/07/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201707_fv3.1.nc
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/08/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201708_fv3.1.nc
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/09/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201709_fv3.1.nc
http://gws-access.ceda.ac.uk/public/cds_c3s_cloud/c3s_312b_lot1/data/erb/c3s/icdr/r01/monthly/2017/10/C3S-312bL1-L3C-MONTHLY-ERB-SLSTR_ORAC_Sentinel-3a_201710_fv3.1.nc

First line of a manifest file for a dataset saved in MARS

head reanalysis-uerra-europe-soil-levels/mars.list
class=ur,expver=prod,levtype=sol,origin=eswi,stream=oper,type=an,param=260199/260360,levelist=1/2/3,time=00:00:00/06:00:00/12:00:00/18:00:00,date=1961-01-01/1961-01-02/1961-01-03/1961-01-04/1961-01-05/1961-01-06/1961-01-07/1961-01-08/1961-01-09/1961-01-10/1961-01-11/1961-01-12/1961-01-13/1961-01-14/1961-01-15/1961-01-16/1961-01-17/1961-01-18/1961-01-19/1961-01-20/1961-01-21/1961-01-22/1961-01-23/1961-01-24/1961-01-25/1961-01-26/1961-01-27/1961-01-28/1961-01-29/1961-01-30/1961-01-31
class=ur,expver=prod,levtype=sol,origin=eswi,stream=oper,type=an,param=260199/260360,levelist=1/2/3,time=00:00:00/06:00:00/12:00:00/18:00:00,date=1961-02-01/1961-02-02/1961-02-03/1961-02-04/1961-02-05/1961-02-06/1961-02-07/1961-02-08/1961-02-09/1961-02-10/1961-02-11/1961-02-12/1961-02-13/1961-02-14/1961-02-15/1961-02-16/1961-02-17/1961-02-18/1961-02-19/1961-02-20/1961-02-21/1961-02-22/1961-02-23/1961-02-24/1961-02-25/1961-02-26/1961-02-27/1961-02-28
class=ur,expver=prod,levtype=sol,origin=eswi,stream=oper,type=an,param=260199/260360,levelist=1/2/3,time=00:00:00/06:00:00/12:00:00/18:00:00,date=1961-03-01/1961-03-02/1961-03-03/1961-03-04/1961-03-05/1961-03-06/1961-03-07/1961-03-08/1961-03-09/1961-03-10/1961-03-11/1961-03-12/1961-03-13/1961-03-14/1961-03-15/1961-03-16/1961-03-17/1961-03-18/1961-03-19/1961-03-20/1961-03-21/1961-03-22/1961-03-23/1961-03-24/1961-03-25/1961-03-26/1961-03-27/1961-03-28/1961-03-29/1961-03-30/1961-03-31

Pseudo-manifest

Dataset suppliers to the CDS shall provide a comprehensive description of their data at least two months prior to delivery, using a data registration process established by ECMWF.
For the CDS team this means the delivery of a pseudo-manifest file.

A pseudo-manifest is a manifest file with expected path and filenames for the expected data to be created. Note that the pseudo-manifest should be as close as possible of the final delivery but the CDS team understands that modifications may be needed.
If a pseudo-manifest is provided, then a Catalogue entry can be created and its design agreed and tested. Filenames and paths can be checked to see if they allow a good building of the download form.

Name of the manifest and updates of the contents of the manifest

The manifest should be named "manifest_<Contract tag>_<ECV_name_tag or SIS_name_tag>_<optional_tag>_yyyymmdd.txt" where yyyymmdd is the date where this manifest was created.
It is expected that the providers replace the strings <...> in the manifest filename with the actual names for the dataset they are providing.

When a new manifest file is added to the the providers site, that manifest should also be copied to "manifest_<Contract tag>_<ECV_name|SIS_name>_<optional_tag>_**latest**.txt".
Remove the date and leave just the string "latest".

This convention is central for the CDS computers to find and access the correct manifest.

Where to save the manifest file:

The manifest should be in a directory named http://web[5] address/c3s_manifest/ accessible through wget and http(s)f or the CDS computers to automatically download/check it.
Old manifest files may be removed from the providers site. The idea is to store 2 or 3 previous manifest files to track back any issues.
At least one manifest file should be always present and that providers site: the latest manifest file.

Why the CDS values so much the manifest file:

For the Catalogue a dataset is a manifest file. Not the description in the contracts, not overviews.
Nothing else is so central and important than the manifest: it tells the CDS computers what should be present in the public Catalogue.

How the download form is directly related with filenames and paths in the manifest?

The widgets in the download pages of the CDS Catalogue are the way by which the user builds the name and the path of the file that corresponds to the data the user wants to download.
In other words, there is a direct link between the filename and path convention and what we can offer for the user to click in the download form.
To see it better let's consider the CIMP5 datasets which has addressed like:

:/output1/NOAA-GFDL/GFDL-CM3/historical/day/atmos/day/r3i1p1/ua/v20120227/ua_day_GFDL-CM3_historical_r3i1p1_19800101-19841231.nc[6]

the path and filename convention is:

/output1/(?P<__organisation>[\-\w]+)/(?P<model_1>[\-\w]+)/(?P<experiment_1>[a-zA-Z0-9]+)/(?P<frequency>[a-z0-9]+)/(?P<__realm>\w+)/(?P<__cmor_table_1>[a-zA-Z0-9]+)/(?P<ensemble_member_1>[a-z0-9]+)/(?P<variable_1>[a-zA-Z]+)/(?P<__version>[v0-9]+)/(?P<variable_2>[a-zA-Z]+)_(?P<__cmor_table_2>[a-zA-Z0-9]+)_(?P<model_2>[a-zA-Z0-9\-]+)_(?P<experiment_2>[a-zA-Z0-9]+)_(?P<ensemble_member_2>[a-z0-9]+)_(?P<period>\d{8}-\d{8}).nc.

When one looks at https://cds.climate.copernicus.eu/cdsapp#!/dataset/projections-cmip5-daily-pressure-levels?tab=form it is easy to notice that the widgets on that download page are the ones defined in the convention.
By clicking on the boxes the user is in fact providing values to each part of the convention and building the name of the file that will be downloaded. Each part of the convention will eventually lead to a widget in the Catalogue for the dataset.

CDS preferences:

The CDS prefer long names than short names. We prefer understandable than smart. For instance we prefer L3-U.nc[7] than L3U.nc[8] since in the first case it is clear that 2 things are at play.

5 http://web/
6 http://ua_day_gfdl-cm3_historical_r3i1p1_19800101-19841231.nc/
7 http://l3-u.nc/
8 http://l3u.nc/

Filenames should follow conventions. More than one convention is OK. Different main variables should be preferably in different files.

Examples of what would be desirable for the filename conventions:

If possible, in the filenames the underscore "_" should be used to split between place holders and hyphen "-" to say that different words belong in fact to the same placeholder.
For instance:
sfcWind_climatology_prevailing/01/sfcWind_climatology_prevailing_01_v0.0.nc[9]
would be better as:
sfcWind-prevailing_climatology/01/sfcWind-prevailing_climatology_01_v0.0.nc
[10]By using this grouping and splitting *this will help to design more well organised filenames that will be easier to use by the CDS scripts.*

Warning:
The same thing should be named in the same way whenever it is referenced.
For instance, for a computer "version_0.0" is different from "v0.0". If we mean the same thing then the string should be exactly the same, no differences in capitalisation, or more letters or less letters. One can choose either v0.0 or version_0.0 or another string one finds convenient but should then keep it the same everywhere when the same thing is meant.

Versions, deprecation of entries, replacement of data

Amount of data to deprecate	**Provider's role**	**CDS team**
Large amount of data	• Provide old and new data in the same updated manifest file • Keep old and new data • Remove deprecated data and corresponding lines from the manifest at the end of the deprecation period	• Deprecate the whole entry and create a new one. The deprecated entry will not be searchable in the CDS, but API request will continue to work. (This prevents new users to find and download deprecated data, allowing at the same time scientific traceability and reproducibility), Example: Deprecated SST[11]. New entry with corrected data: Corrected SST[12] • Remove the deprecated data after 1 to 3 year deprecation period

9 http://global-shipping.copernicus-climate.eu/shipping_metocean_variables_monthly_climatology/v0.0/sfcWind_climatology_prevailing/01/sfcWind_climatology_prevailing_01_v0.0.nc_
10 http://global-shipping.copernicus-climate.eu/shipping-metocean-variables_monthly-climatology/v0.0/sfcWind-prevailing_climatology/01/sfcWind-prevailing_climatology_01_v0.0.nc
11 https://cds-test.climate.copernicus.eu/cdsapp#!/dataset/satellite-sst-esa-cci?tab=overview
12 https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-sea-surface-temperature?tab=overview

Small amount of data	• Create new files with a different version tag for the corrected data • Include those files in the manifest • Manifest should contain both old and new versions • Remove deprecated data and corresponding entries in the manifest at the end of the deprecation period	• Deprecate the version of the data corresponding to the wrong data • Modify overview to explain the deprecation or use a new widget called "Known issues" under the Documenation tab • Modify the download form making clear the deprecated version of the data. (When the CDS will have the tools to do it: the deprecated data will only be accessible through the API). • Remove the deprecated data after 1 to 3 year deprecation period

DOI, citation, licence and acknowledgement

DOI

Type of data in the CDS Catalogue entry	**Provider's role**	**CDS team**
Data **without** DOI issued before the publication in the CDS Catalogue	• No active role	• Provides a DOI to the Catalogue entry (which can be see as a DOI for the data themselves) Example: https://cds.climate.copernicus.eu/cdsapp#!/dataset/cems-glofas-reforecast?tab=overview

Data **with** DOIs issued before publication in the CDS Catalogue	• Provides a mapping between the data and the previous DOIs	• Create a DOI box allowing for multiple DOIs.
		• DOI's box will show
		  • all the DOIs supplied by the data provider (with a clear association to which data they refer to)
		  • the DOI of the Catalogue entry itself
		Example: https://cds-test.climate.copernicus.eu/cdsapp#!/dataset/satellite-total-column-water-vapour?tab=overview https://cds-dev.copernicus-climate.eu/cdsapp#!/dataset/satellite-cloud-properties?tab=overview

Mixing of data **with and without** DOIs issued before the publication in the CDS Catalogue	• Provides a mapping between the data and the DOIs	• Create a DOI box allowing for multiple DOIs.
		• DOI's box will show
		  • all the DOIs supplied by the data provider (with a clear association to which data they refer to)
		  • the DOI of the Catalogue entry itself
		  • data with no DOI attribute will be associated with the string: "no specific DOI"
		Example: https://cds-dev.copernicus-climate.eu/cdsapp#!/dataset/satellite-surface-radiation-budget?tab=overview

Citations

Citations are like file formats, there are a few available, no one better than all the others in all situations.
The "Citation" link in the Catalogue entry **does not say how** people should cite the data, that depends on the journal, site and publisher where the data will be cited.
The "Citation" link in the Catalogue entry **is the Catalogue citing** the contents that it is exposing.
In this way it also shows how to cite the data, but that is just an example of how to cite the data and contents from where **people can extract all information** required to cite the data using other formats in other places.

Type of data in the CDS Catalogue entry	**Provider's role**	**CDS team**
Data **without** citation issued before the publication in the CDS Catalogue	• Interact with the CDS team on this. Most probably you will be asked for the names of the authors of the data.	• Interact with the provider and create a Citation following the Catalogue citation format

Data **with** citation issued before publication in the CDS Catalogue	• Provides those citations to the CDS team	• Create a Citation box allowing for multiple citations.
		• Citation's box will show
		  • all the Citations supplied by the data provider (with a clear association to which data they refer to)
		  • the Citation of the Catalogue entry itself
		
		Example:
		https://cds-test.climate.copernicus.eu/cdsapp#!/dataset/satellite-total-column-water-vapour?tab=overview
		https://cds-dev.copernicus-climate.eu/cdsapp#!/dataset/satellite-cloud-properties?tab=overview

Mixing of data **with and without** citations issued before the publication in the CDS Catalogue	• Provides those citations to the CDS team	• Create a Citation box allowing for multiple citations.
		• Citation's box will show
		  • all the Citations supplied by the data provider (with a clear association to which data they refer to)
		  • the Citation of the Catalogue entry itself
		Example:
		https://cds-dev.copernicus-climate.eu/cdsapp#!/dataset/satellite-surface-radiation-budget?tab=overview

Licence

Provider's role	**CDS team**
Provide all licences related to the data and a mapping between the licences and the parts of the data they are related to	For datasets with multiple licences use a "Origin" button in the download form making related to the name of the licence Example: https://cds-dev.copernicus-climate.eu/cdsapp#!/dataset/satellite-surface-radiation-budget?tab=form

Acknowledgement

Provider's role	**CDS team**

No active role but may want to have a look at: How to acknowledge and cite a Climate Data Store (CDS) catalogue entry and the data published as part of it[13] https://cds-test.climate.copernicus.eu/cdsapp#!/dataset/satellite-total-column-water-vapour?tab=overview	For datasets with multiple licences use a "Origin" button in the download form making related to the name of the licence Example: https://cds-dev.copernicus-climate.eu/cdsapp#!/dataset/satellite-surface-radiation-budget?tab=form

13 https://confluence.ecmwf.int/display/CKB/How+to+acknowledge+and+cite+a+Climate+Data+Store+%28CDS%29+catalogue+entry+and+the+data+published+as+part+of+it

5 Publishing under FAIR principles[14]

14 https://confluence.ecmwf.int/display/PS/Publishing+under+FAIR+principles

Publishing under FAIR principles

Production Section

Exported on 02/21/2024

Table of Contents

Search

1 Core interpretation of the FAIR principles

The core principles of the FAIR guidelines have not changed since they were first published in 2016 (*), and have since been widely adopted by the scientific community as a way to improve the quality and usability of research data. However, the principles are intended to be flexible and adaptable, and different organizations and communities may have different interpretations and implementations of the principles. It's also possible that the principles may be updated or refined over time as the field of data science and technology evolves.

FAIR principles are a set of guidelines for making data more Findable, Accessible, Interoperable, and Reusable:

1. Is the data Findable?
   Can the data be easily discovered by those who need it, using relevant keywords and metadata?
2. Is the data Accessible?
   Can the data be accessed, read, and understood by a machine or a human? Is it available in a widely used, open format?
3. Is the data Interoperable?
   Can the data be easily integrated with other data sources, using common standards and formats?
4. Is the data Reusable?
   Can the data be used and reused for multiple purposes, without significant effort or additional licensing restrictions?

If the data meets all of these criteria, it can be considered "FAIR." It's important to note, however, that the FAIR principles are guidelines rather than strict rules, and different organizations and communities may have different interpretations and implementations of the principles.

(*) 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

Source : Australian National Data Service (ANDS)
Entire FAIR resources graphic is licensed under a Creative Commons Attribution 4.0 International License

2 ECMWF data: Extending the FAIR principles to all our data

At the ECMWF, we are committed to making our data as useful and accessible as possible. That's why we aim to publish all our data in accordance with the FAIR principles.

Our data is carefully curated and described using relevant metadata, that provides detailed information about the variables and parameters included in the data. For each variable, we provide a clear definition, specify the units, and include any relevant notes or caveats that users should be aware of, ensuring that the data can be used accurately and reliably.

We use DOIs (Digital Object Identifiers) to provide persistent, stable links to our data, allowing users to easily find and access the data they need. We also use open, standardized formats for our data and provide API (Application Programming Interface) access, allowing users to easily integrate our data with other systems and applications.

And we provide clear licensing information, enabling users to freely reuse and repurpose the data for their own purposes.

By following the FAIR principles, we are helping to make our data more valuable and useful for a wide range of users, from meteorologists and researchers to policymakers and the general public. We are proud to be part of the growing community of organizations that are working to make data more FAIR (*).

(*) OGC FAIR Climate Services: ECMWF is co-chair for the OGC Climate Resilience Domain Working Group[1]

- Decommissioning plan of ECMWF public datasets service[2]

1 https://www.ogc.org/blog/4460
2 https://confluence.ecmwf.int/display/PS/Decommissioning+plan+of+ECMWF+public+datasets+service