



Advances in Collaborative Documentation Support for CMIP6

EGU 2020
4-8 May 2020

C. Pascoe, D. Hassell, M. Stockhause, M. Greenslade, S. Bartholomew

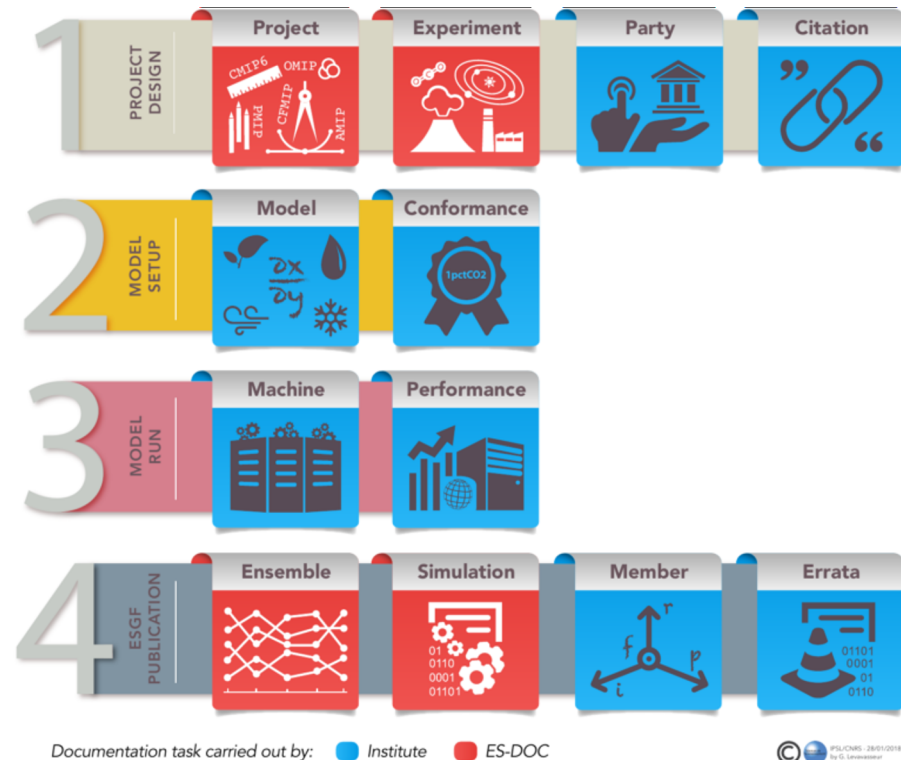
EGU2020-19636: <https://doi.org/10.5194/egusphere-egu2020-19636>

The Earth System Documentation (ES-DOC) project nurtures an ecosystem of tools & services to enable the scientific community to better understand and utilise Earth system model data.

ES-DOC describes the entire workflow that created the data.

ES-DOC documentation includes information about the Project and Experiment design, Model formulation and Ensemble and Simulation descriptions.

ES-DOC documentation also links to the DKRZ data citation service.



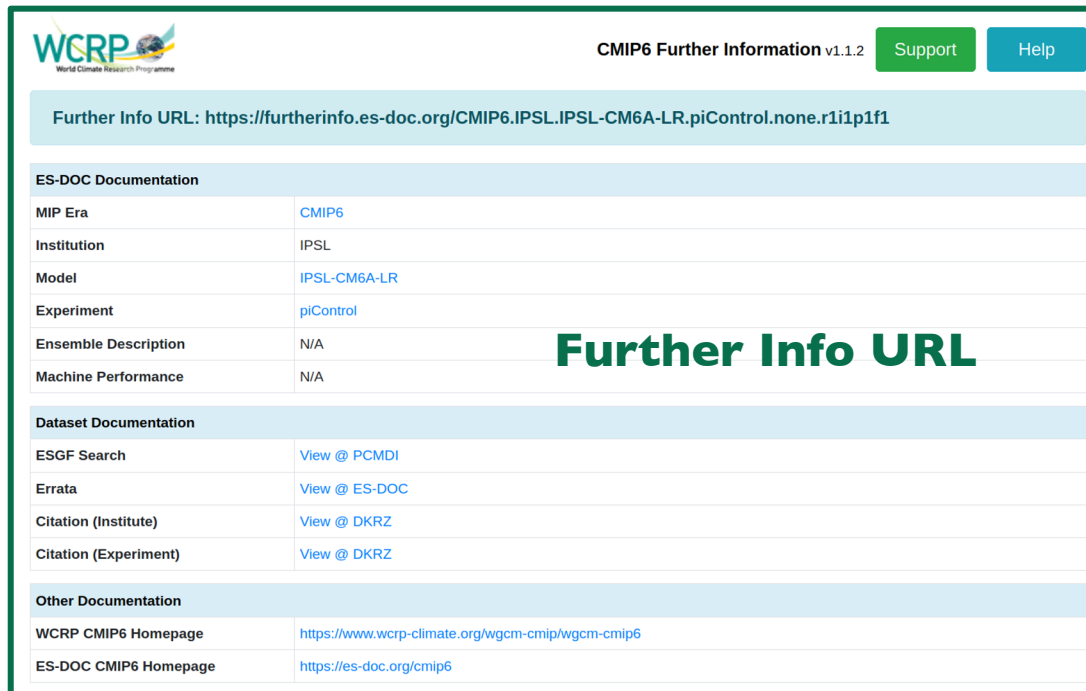
Modelling groups contributing to the Coupled Model Intercomparison Project (Phase 6) (CMIP6) are using ES-DOC to describe their climate models.

ES-DOC has generated complete structured descriptions of the CMIP6 MIPs and experiments.

ES-DOC auto-generates documentation for every CMIP6 simulation in the Earth System Grid Federation (ESGF).

The ES-DOC information can be accessed via the “[further info URL](#)” that is found in every CMIP6 netCDF data file.

The further info URL is a landing page from which all of the CMIP6 documentation relevant to the data may be accessed.



The screenshot shows the 'CMIP6 Further Information v1.1.2' web page. It features a header with the WCRP logo and navigation buttons for 'Support' and 'Help'. A light blue banner displays the 'Further Info URL: <https://furtherinfo.es-doc.org/CMIP6.IPSL.IPSL-CM6A-LR.piControl.none.r1i1p1f1>'. Below this are three sections: 'ES-DOC Documentation', 'Dataset Documentation', and 'Other Documentation', each containing a table of links and information.

ES-DOC Documentation	
MIP Era	CMIP6
Institution	IPSL
Model	IPSL-CM6A-LR
Experiment	piControl
Ensemble Description	N/A
Machine Performance	N/A

Dataset Documentation	
ESGF Search	View @ PCMDI
Errata	View @ ES-DOC
Citation (Institute)	View @ DKRZ
Citation (Experiment)	View @ DKRZ

Other Documentation	
WCRP CMIP6 Homepage	https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6
ES-DOC CMIP6 Homepage	https://es-doc.org/cmip6

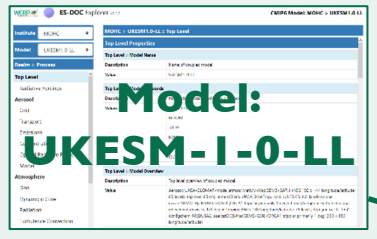
Further Info URL

```
$ ncdump pr_Amon_piControl_IPSL.IPSL-CM6A-LR-LR_r1i1p1f1.nc | grep further_info_url  
:further_info_url = "https://furtherinfo.es-doc.org/CMIP6.IPSL.IPSL-CM6A-LR.piControl.none.r1i1p1f1";
```

Further Info URL

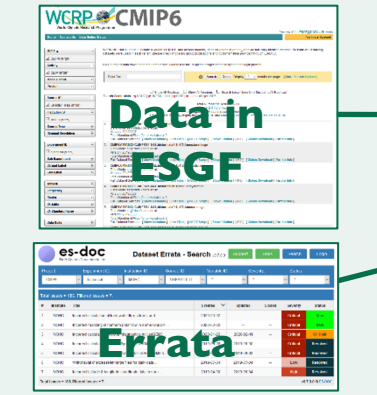
<https://furtherinfo.es-doc.org/CMIP6.MOHC.UKESM1-0-LL.historical.none.r1i1p1f2>

Model Info



Model:
UKESM-1-0-LL

Data Access and Errata



Data in ESGF

Errata



CMIP6 Further Information v1.1.2

Support

Further Info URL: <https://furtherinfo.es-doc.org/CMIP6.MOHC.UKESM1-0-LL.historical.none.r1i1p1f2>

ES-DOC Documentation

MIP Era	CMIP6
Institution	MOHC
Model	UKESM1.0-LL
Experiment	historical
Ensemble Description	N/A
Machine Performance	N/A

Dataset Documentation

ESGF Search	View @ PCMDI
Errata	View @ ES-DOC
Citation (Institute)	View @ DKRZ
Citation (Experiment)	View @ DKRZ

Other Documentation

WCRP CMIP6 Homepage	https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6
ES-DOC CMIP6 Homepage	https://es-doc.org/cmip6


CMIP6 Info



CMIP6

Experiment: historical

DKRZ Data Citation



DOI landing page

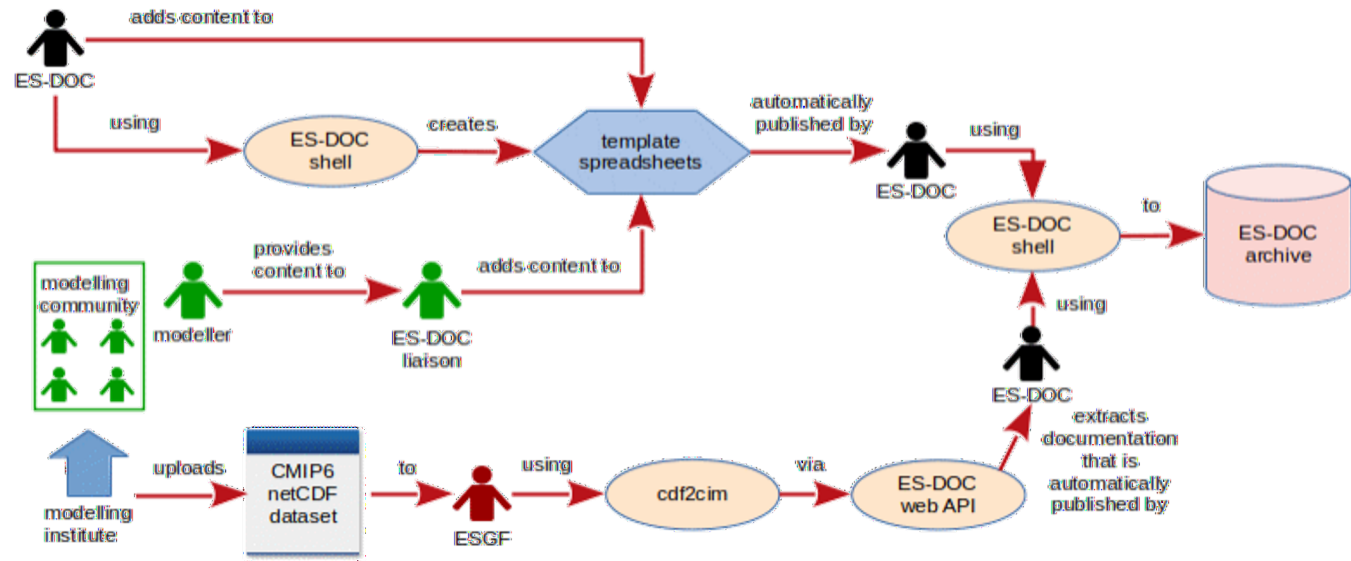
Structured scientific metadata has an important role in science communication, however it's creation and collation exacts a cost in time, energy and attention.

We look to find a balance between the ease of information collection and the complexity of our information handling structures.

Modellers are willing to engage in such endeavours where the effort is perceived as being relevant and helpful to their science.

To this end it is important to have tools in place that can be used by scientists to exploit the metadata resources that they so generously provide.

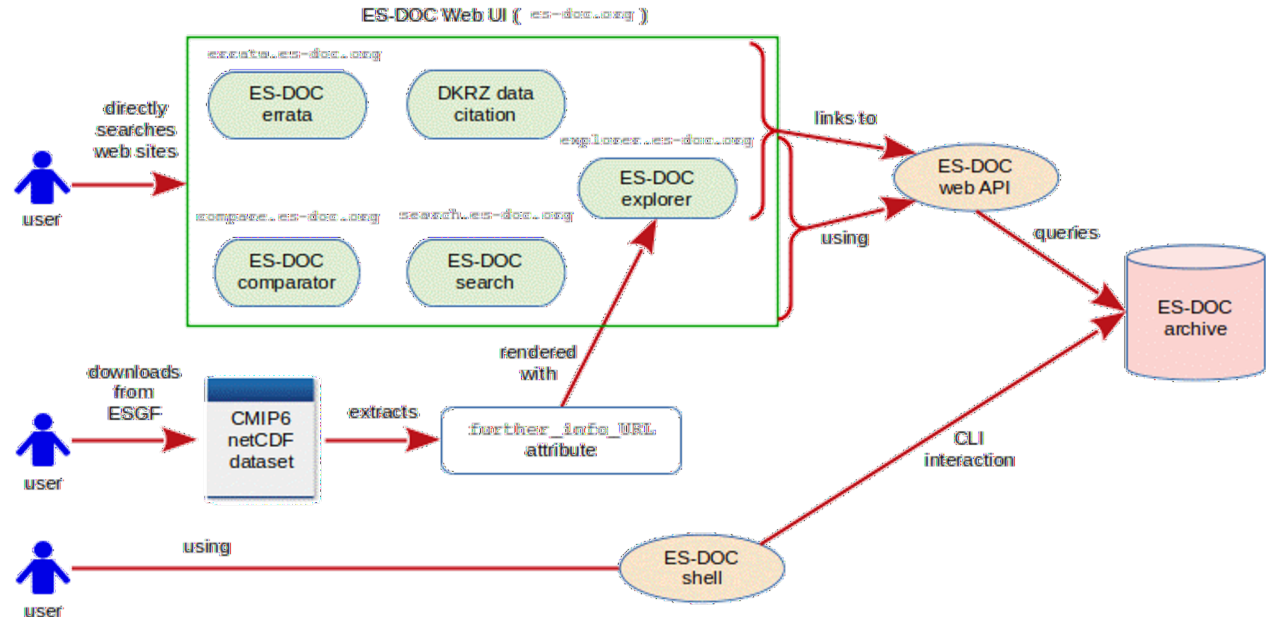
ES-DOC services and the Citation Service depend upon structured documentation provided by the scientific community.



Scientists involved in CMIP6 will find the ES-DOC infrastructure helpful.

Whether they are...

- Highlighting dataset errors.
- Exploring experiment protocols.
- Comparing forcings across ensemble runs.
- Understanding Model Intercomparison Project objectives.
- Reviewing citations.
- Exploring component properties of configured models.
- Visualising inter-model relationships.



ES-DOC has been operational for CMIP6 since December 2018

- **Experiments** fully documented by ES-DOC at <http://search.es-doc.org>
- Modelling groups have started to use the ES-DOC infrastructure to document their **model formulations: 10/49** Groups have (partially) published **22/130** model descriptions
- Spreadsheets disseminated via GitHub are filled in, returned to GitHub, and then automatically published to <http://search.es-doc.org>
- This methodology has been designed to be generalisable to the other types of documentation that will be rolled out later in 2020 including:
 - **Conformance, Machine, performance, Differences between ensemble members**
- **Simulations** are documented by an automated script running as part of the ESGF publisher
- Metadata from **995** institute/model/experiment combinations spanning from at least **2230267** publications have been collected, more arriving all the time.
- These are currently being reduced to $O(1000)$ unique publishable simulation descriptions
- **Further Info URL** found in each CMIP6 netCDF files is fully functional, including links to errata citation services

99% of the registered CMIP6 data is citable

The DKRZ Data Citation Service provides users of CMIP6 data with the formal citation that should accompany any use of the datasets that comprise their analysis.

Data DOI landing pages are also part of the Citation Service, provided by DKRZ.

The data citation information is also available independently through the ESGF Search portal, in the DataCite search and via Google's dataset search.

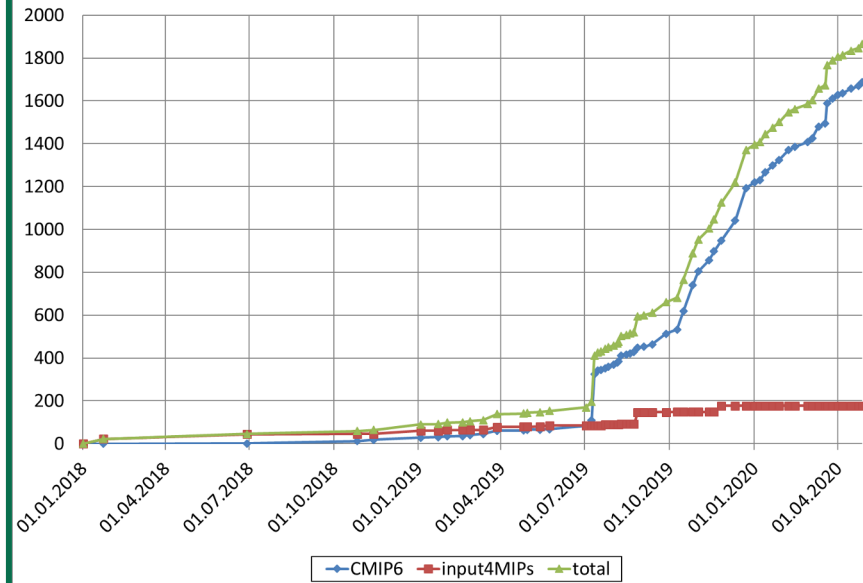
DOI Status (2020-04-01)

- Number of DOIs: 1 865
 - CMIP6: 1 688
 - input4MIPs: 177
- Number of MD updates: 8 164
- Missing Citations (total): 27
 - CMIP6: 24
 - input4MIPs: 3

References:

http://bit.ly/CMIP6_DOI_Statistics
http://bit.ly/CMIP6_Citation_Search

DataCite DOI Registration for CMIP6



Discover more about the DKRZ citation service on Wednesday in EGU2020-4616: <https://doi.org/10.5194/egusphere-egu2020-4616>



The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°824084

<https://is.enes.org>