


Documentation of climate change data supporting cross-domain data reuse

EGU 2020
4-8 May 2020

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

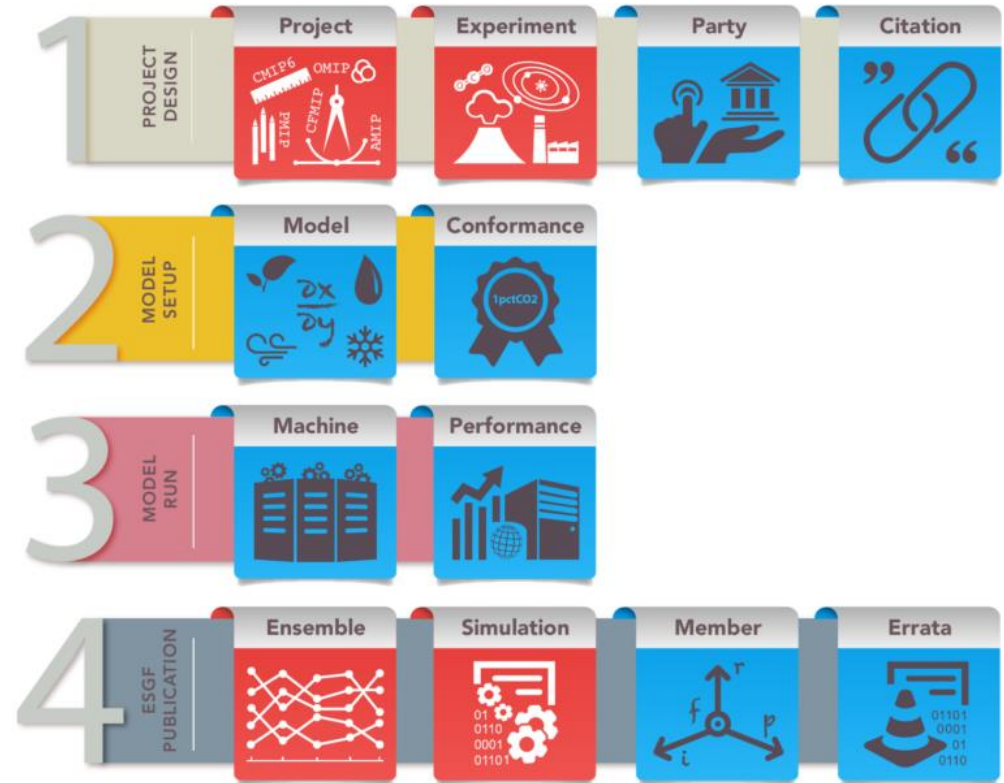
EGU2020-4616: <https://doi.org/10.5194/egusphere-egu2020-4616>

Climate Change Data and Cross-Domain Reuse

- Climate change data is created by multiple modelling centres worldwide in international projects such as [CMIP6 \(Coupled Model Intercomparison Project Phase 6\)](#).
- Climate change data is analysed under various aspects and reused for a wide range of questions.
- Users of climate change data need to be supported by detailed documentations on data creation and experiment design to understand the data and information on license and data citation defining the terms of use. These are provided by:
 - [ES-DOC \(Earth System Documentation\)](#)
 - [Citation Service](#)
- **Furtherinfo** pages provide the entries to these documentation. Links are included in every CMIP6 NetCDF file header.

Documentation workflow:

- About half of the documents **automated or ES-DOC generated**
- The others **produced by groups** when ready
- **Further Info URL**
Binds all available documentation together in a “bottom up” approach



Documentation task carried out by: ● Institute ● ES-DOC

Furtherinfo (2): ES-DOC Documentations

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



CMIP6 Further Information v1.1.2

Support

Help

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

Information on applied model

Model: UKESM-1-0-LL

Information on CMIP6

CMIP6

Experiment: historical

Furtherinfo (3): Data Documentation

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



CMIP6 Further Information v1.1.2

Support

Help

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

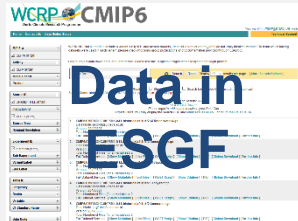
Dataset Documentation

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

Data Access and Errata



Data Citation and License



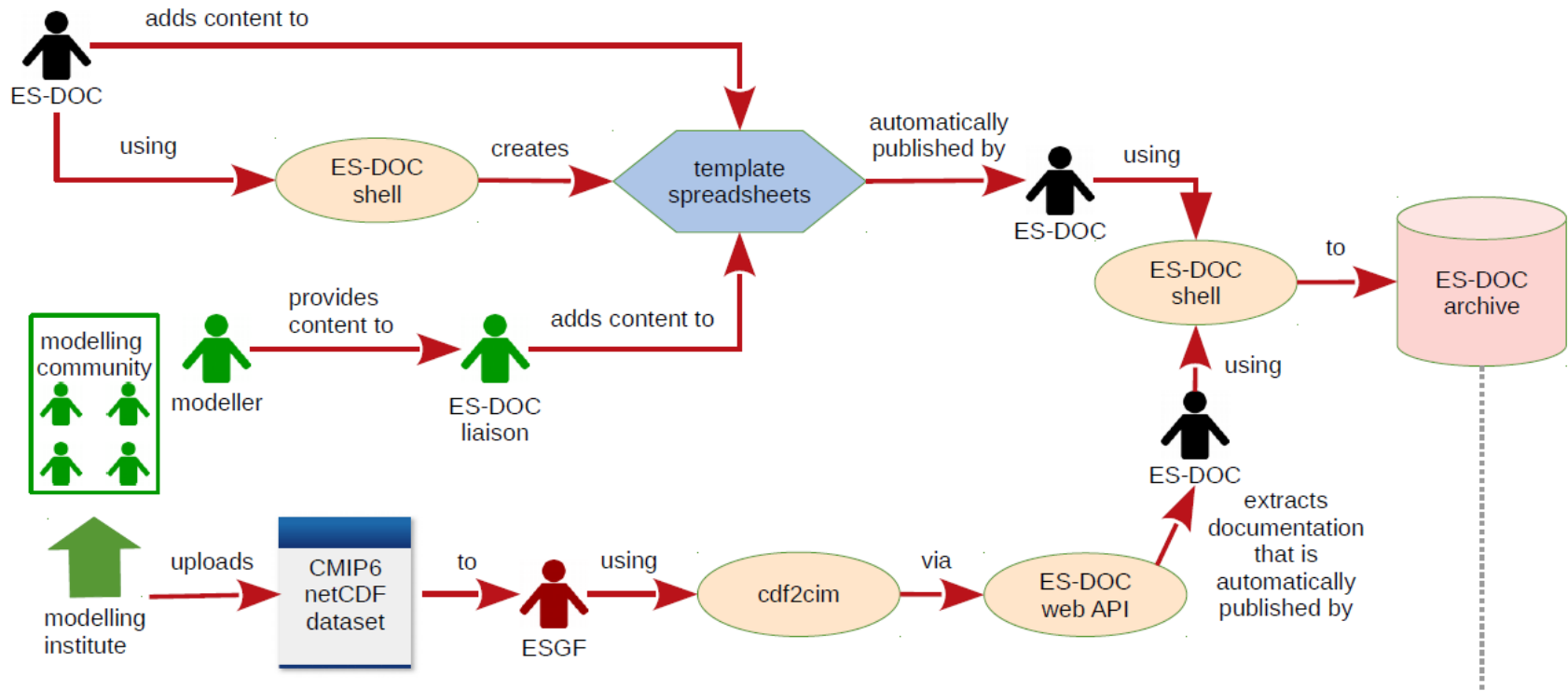
DOI landing page

ES-DOC (1): Overview and Status

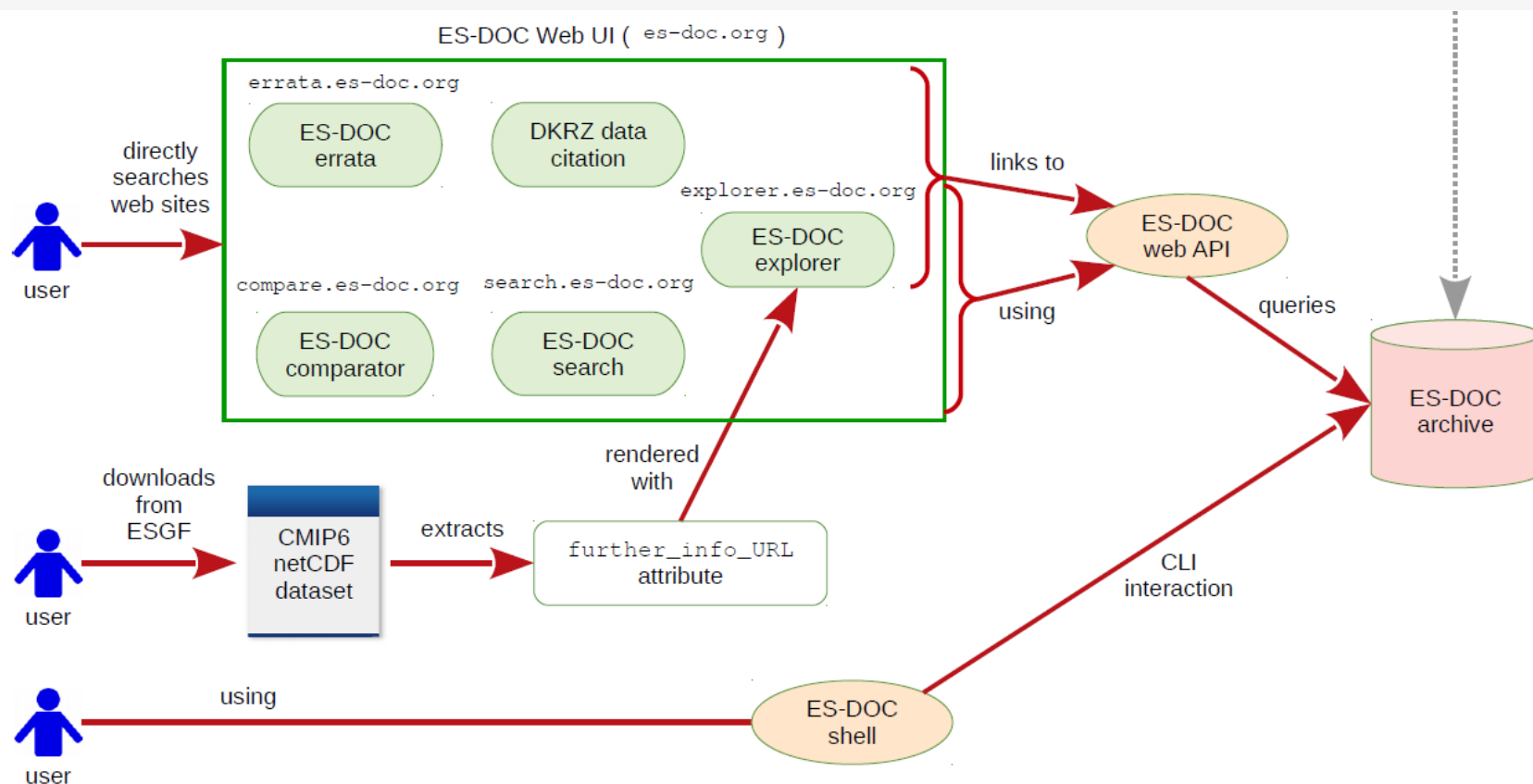
ES-DOC is Operational for CMIP6 since December 2018

- **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
 - **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

ES-DOC (2): Collecting Documentation



ES-DOC (3): Accessing Documentation



The possibility to cite CMIP6 data for the first time in the history of CMIP was established upon a request from WGCM-CMIP.

Service Features

- Data references for input4MIPs and CMIP6 provided on two granularities: MIP/model and experiment
- DataCite metadata standard used as metadata exchange format
- ESGF/CMIP6: Integration into ESGF / CoG, ES-DOC, and Long-Term Archival
- IPCC AR6 WGI: CMIP6 data references included
- Scholix / Data Impact: Data impact information (use in publication) gathered as added value for data providers (dep. on implementation at publishers)
- External discovery in
 - Google Dataset Search
 - DataCite catalog → Scholix → Article Publishers (e.g. Elsevier: [doi:10.1016/j.ocemod.2018.07.002](https://doi.org/10.1016/j.ocemod.2018.07.002)) → harvested by many catalogs
 - OAI/PMH

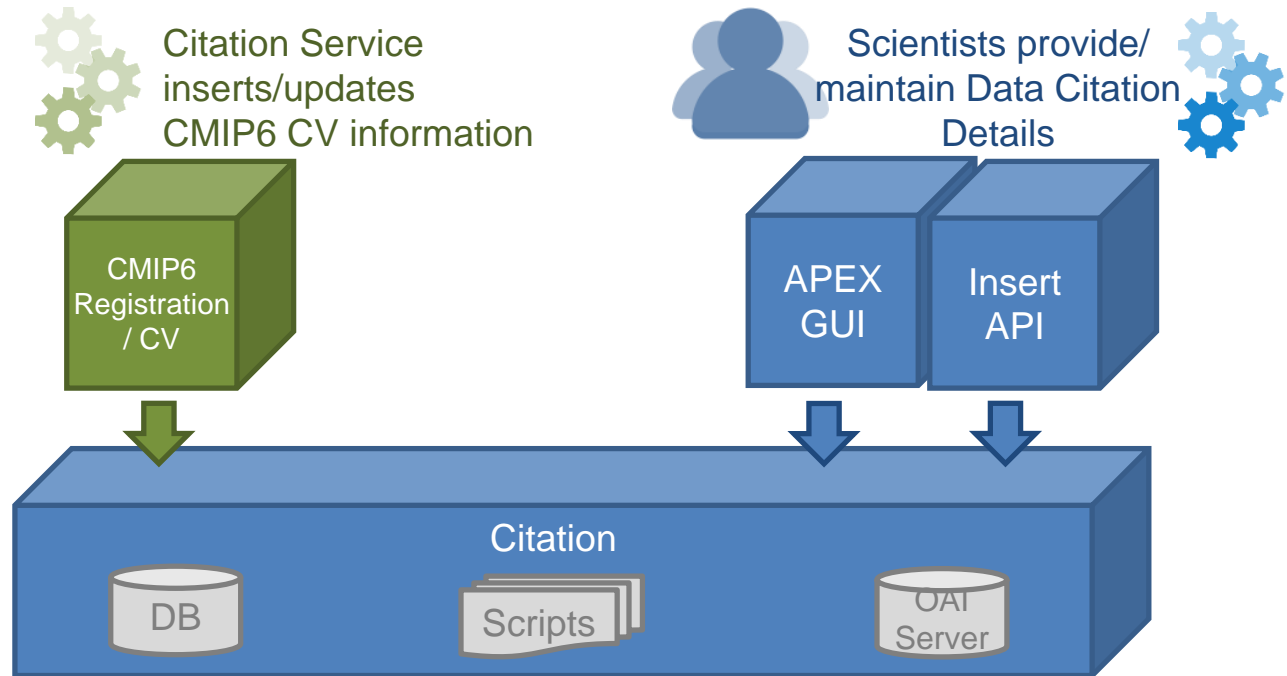
References:

<http://cmip6cite.wdc-climate.de>

WIP paper: [doi:10.5281/zenodo.35178](https://doi.org/10.5281/zenodo.35178)

Publication: [doi:10.5334/dsj-2017-030](https://doi.org/10.5334/dsj-2017-030)

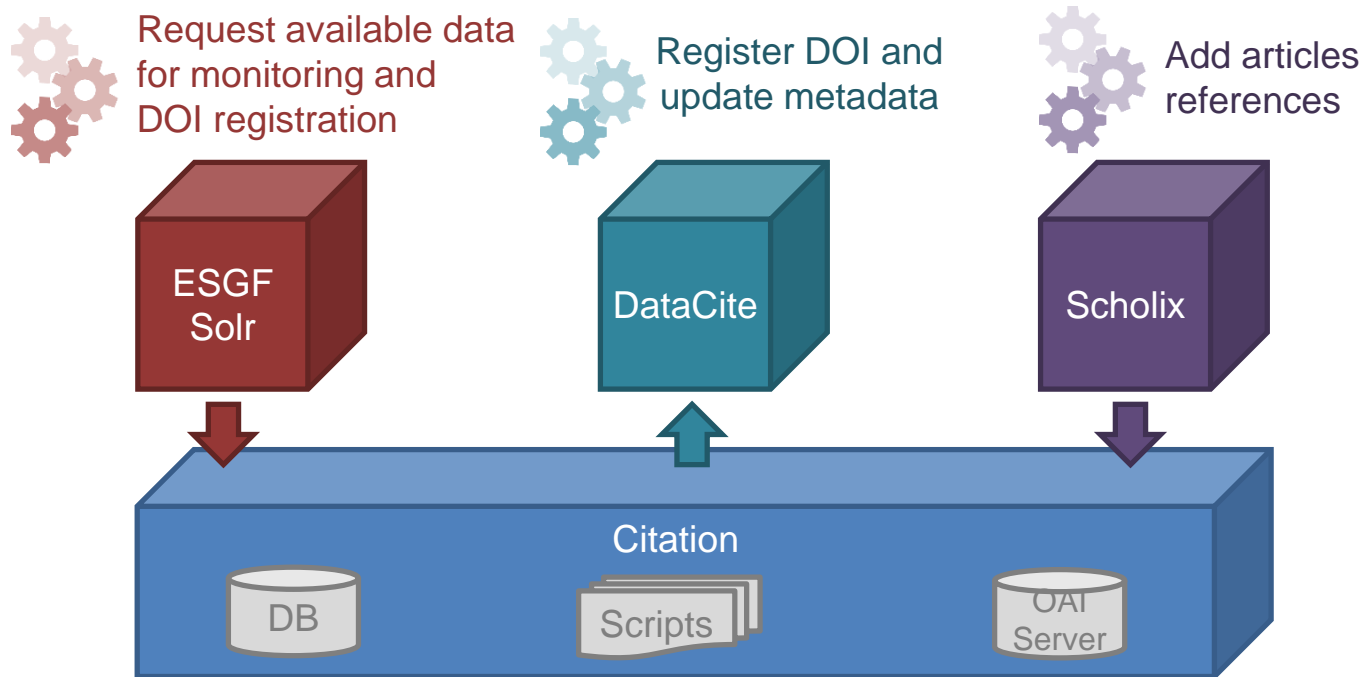
Citation Service (2): Collect Information



Reference:

<http://cmip6cite.wdc-climate.de>

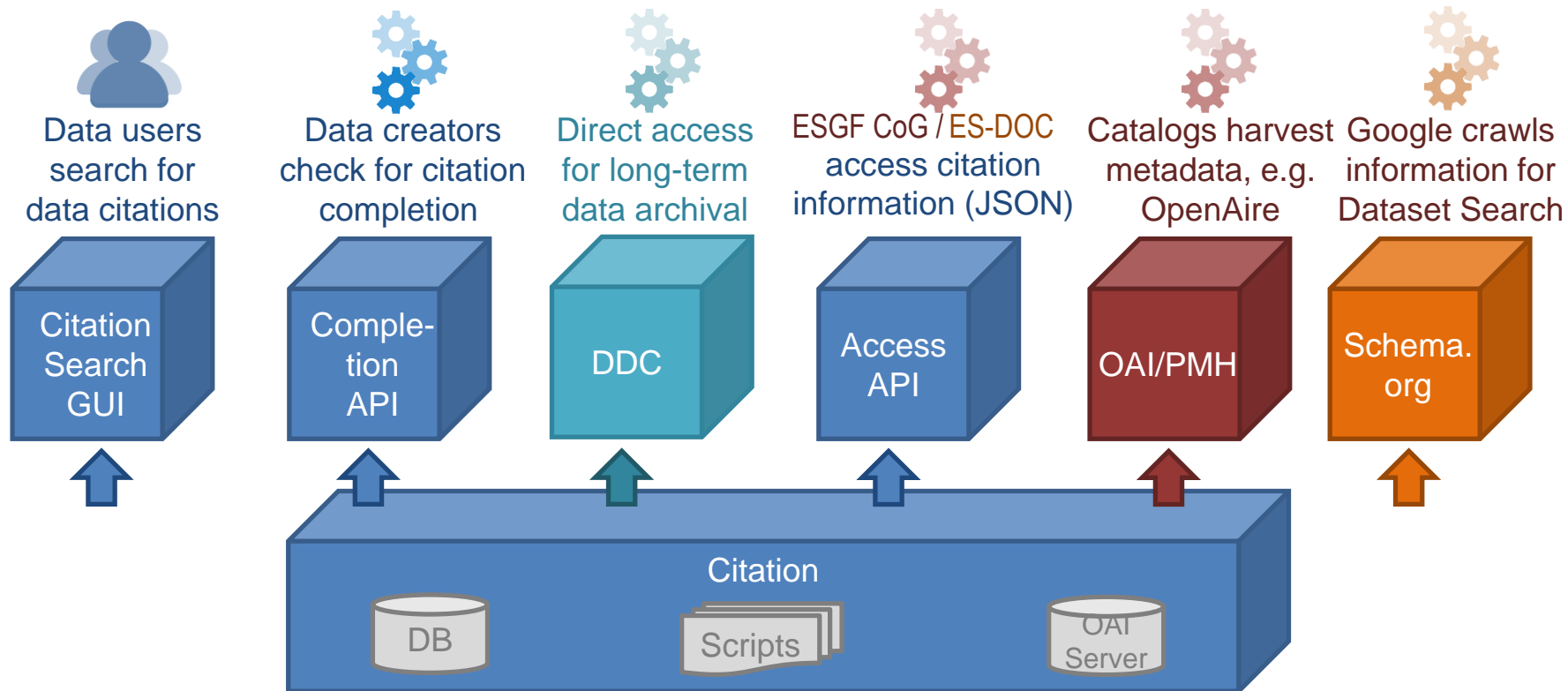
Citation Service (3): Process Information



Reference:

<http://cmip6cite.wdc-climate.de>

Citation Service (4): Disseminate Information



Reference:

<http://cmip6cite.wdc-climate.de>

Citation Service (5): Status

DOI Status (2020-04-27):

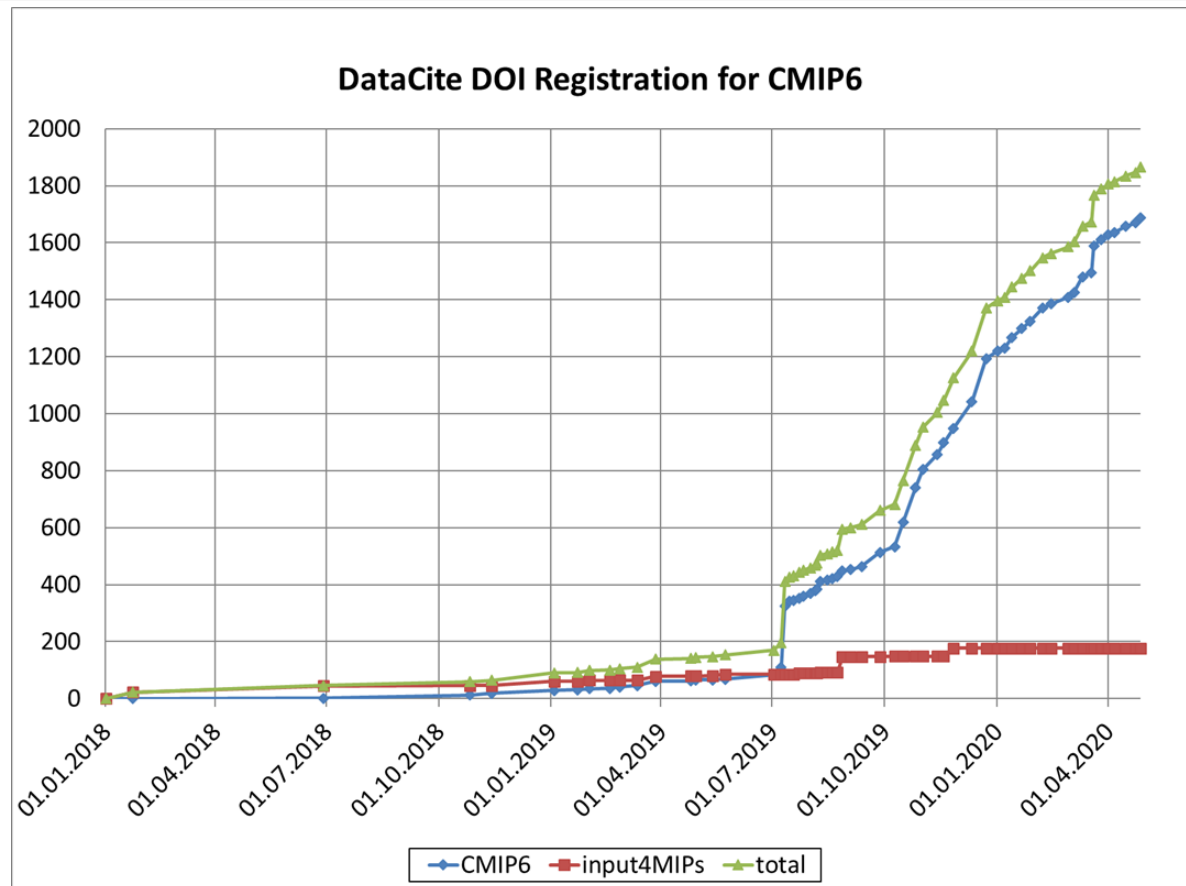
- Number of DOIs: 1 865
 - CMIP6: 1 688
 - input4MIPs: 177
- Number of MD updates: 8 164
- Missing Citations (total): 27
 - CMIP6: 24
 - input4MIPs: 3

→ 99 % of the registered CMIP6 data is citable.

References:

http://bit.ly/CMIP6_Doi_Statistics

http://bit.ly/CMIP6_Citation_Search



Citation Service (6): Feedback and Future

CMIP6 Citation Service has been running as prototype since early 2017.

User Feedback

- Data citations regarded as important
- Few problems with GUI usage
- More difficulties with API usage

Service Requests

- to provide citation for additional projects
- to set up a board for future planning and integration

Requirements

- Technical maintenance and monitoring
- Metadata curation and user support

→ No sustainable funding or institutional support available...

→ Citation Service needs to become: smaller, better integrated with other technical components, and broken down into shareable components

CMIP6 Data Documentation Use Cases

CMIP6 data documentation use cases:

- IPCC WG1 AR6 is going to use model and data references for CMIP6 data to improve the transparency of their findings.
- [Copernicus Data Store \(CDS\)](#) is extended by CMIP6 data. The quality documentation is an important part bringing together QC results, ES-DOC, Errata, Citation, and further CMIP6 project documentations.

- Future plans for the climate data documentation:
 - Improved integration of ES-DOC, citation, and additional documentations, such as the CMIP6 registration
 - Apply ES-DOC / citation for further projects including those without ESGF as data infrastructure
- Preparation for CMIP7:
 - A user survey is in preparation to gather user feedback to guide the WGCM Infrastructure Panel (WIP) in planning CMIP7.
 - The requirements of the technical infrastructure are gathered in the architecture documentation as part of IS ENES3 supporting the WIP.

Questions and / or comments ?