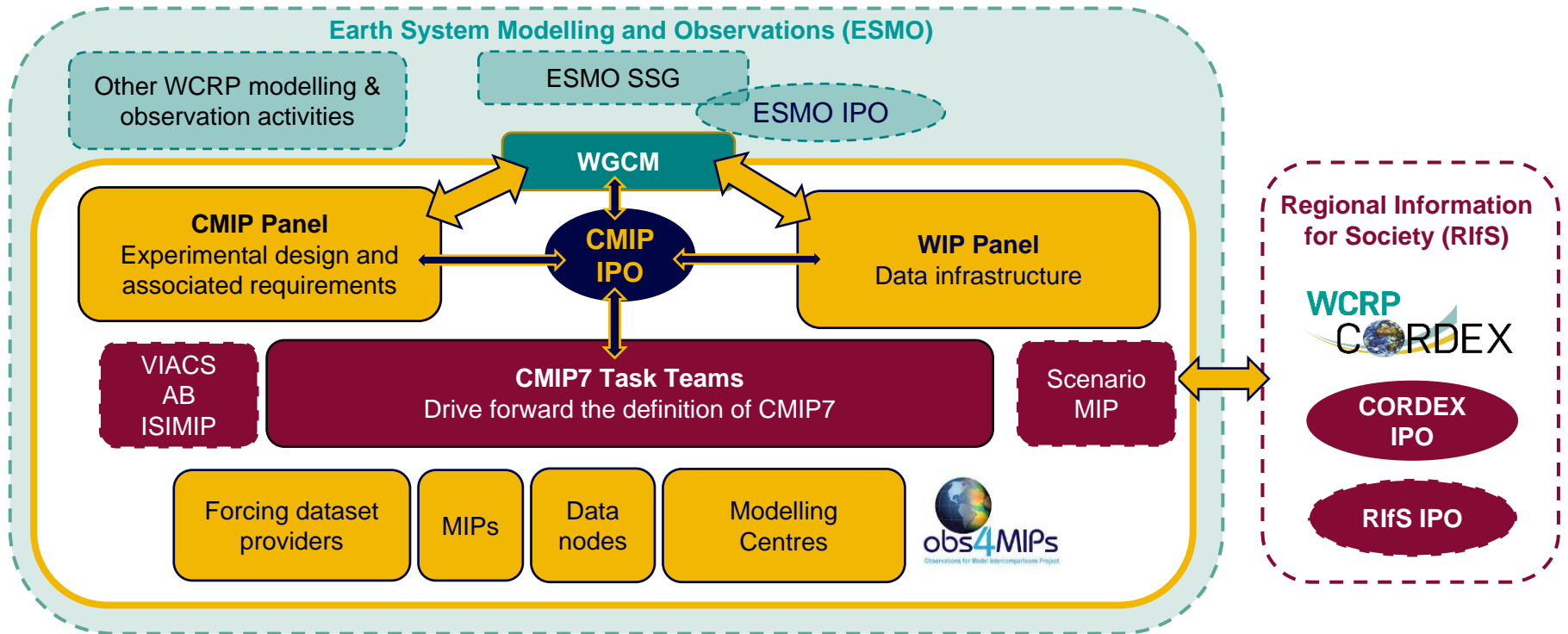# The CMIP7 TT Data Citation for a sustainable continuation of a requested and established service

Martina Stockhause, Sasha Ames, and TT Members

Supported by Eleanor O'Rourke and Elisabeth Dingley (CMIP IPO)

CMIP
Coupled Model
Intercomparison
Project

WCRP
World Climate Research Programme

Task Team in context

Data Citation Service in a Nutshell

Data Citation in CMIP6:
- Citation Service became a new component in CMIP6.
- It was requested by WGCM as data citation possibility prior to long-term data archival.
- It was endorsed by WIP in 2015 and has been running as unfunded stand-alone service.
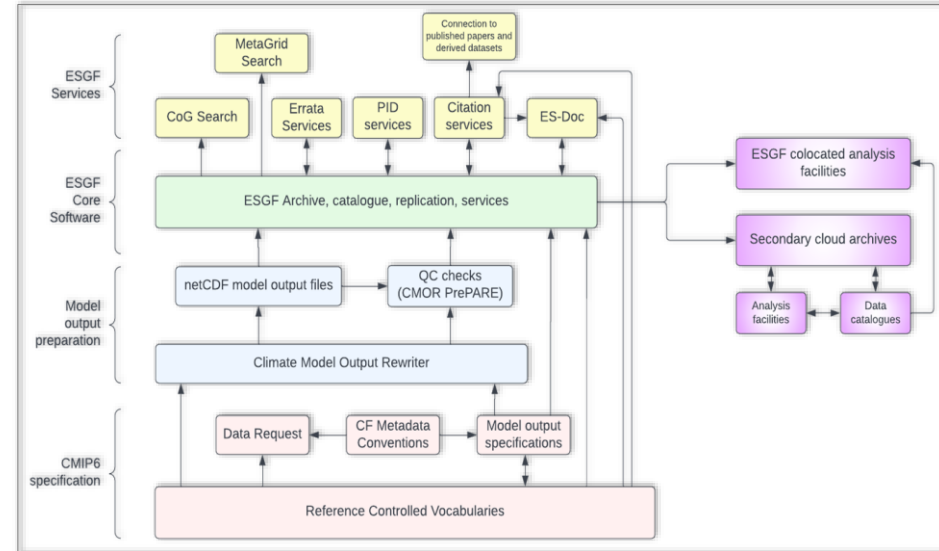
Data Citation Functionalities:
- Collecting/maintaining data citation information
- Automated processing of information including DOI registration
- Disseminating data citation information within and outside of CMIP

Benefit:
- Data is citable in scientific publications.
- Receive credit for data creation
- Enhanced data discoverability

Data Citation Service as infrastructure component:
- Integrated in ESGF and ES-DOC
- Interfaces to DataCite and Scholix
- GUI [API] for maintaining citation information
- APIs for disseminating / GUI for discovery of citation information
- Schema.org compliance enables integration in e.g. Google Dataset Search

(Credit: Matt Mizielinski & Paul Durack)

Stockhause and Ames: CMIP TT Data Citation, EGU 2023, https://doi.org/10.5194/egusphere-egu23-4809

3

Metrics: Registered Data DOIs (2023-04-12)

Data DOI Statistics:

- 2606 CMIP6 & 210 input4MIPs DOIs have been registered
- At WGI AR6 literature and data cut-off date, all datasets were citable.

CMIP6 DOI registration

Statistics

This page   http://bit.ly/CMIP6_DOI_Statistics

Data usage

Further Information

DataCite
https://search.datacite.org/repositories/dkrz.esgf

DataCite statistics for repository ESGF
https://commons.datacite.org/repositories/8orcv25

Citation Service
http://cmip6cite.wdc-climate.de

Citation Search
http://bit.ly/CMIP6_Citation_Search

Stockhause and Ames: CMIP TT Data Citation, EGU 2023, https://doi.org/10.5194/egusphere-egu23-4809

4
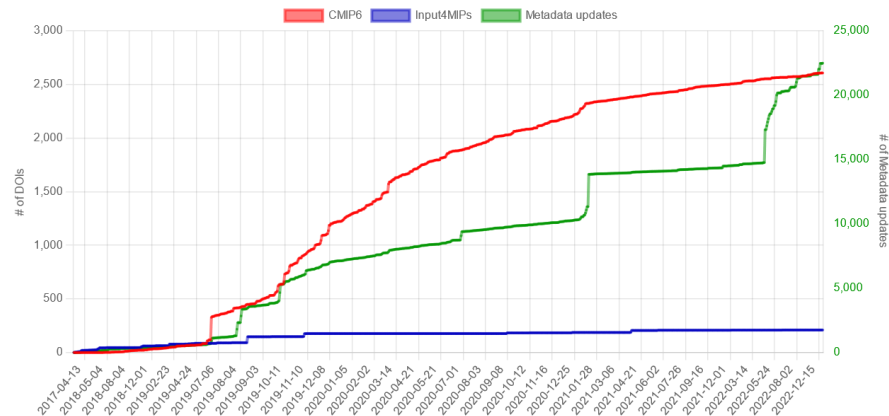
Metrics: Data Impact (2023-04-12)

Data Impact/Usage Metrics:
- IPCC WGI AR6 cited CMIP6 data in table AII.10
- 2176 papers/datasets referencing CMIP6 data
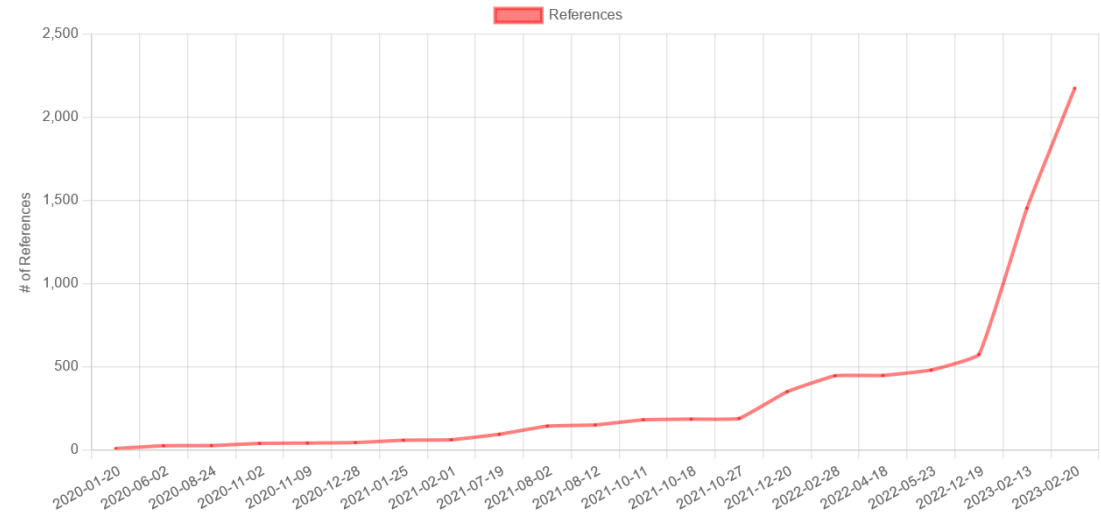 have been added
  - 1603 in 02/2023.

Incomplete data usage due to
- CMIP6 DOI citations not included in reference lists
- Publisher does not publish data references

CMIP6 Data Usage

Based on Scholix interface of OpenAire (http://www.scholix.org; http://scholexplorer.openaire.eu)

Note: Information is incomplete due to data citations not included in reference lists by authors and publishers not publishing data references to crossref.

References

# of References

2,500 — 2,000 — 1,500 — 1,000 — 500 — 0

2020-01-23, 2020-06-02, 2020-08-24, 2020-11-02, 2020-11-09, 2020-12-28, 2021-01-25, 2021-02-01, 2021-07-19, 2021-08-02, 2021-08-12, 2021-10-11, 2021-10-18, 2021-10-27, 2021-12-20, 2022-02-28, 2022-04-18, 2022-05-23, 2022-12-19, 2023-02-13, 2023-02-20

WCRP
World Climate Research Programme

CMIP

Initial community consultation (Survey)

Find/use:

- Improve communication and bring documentation to the users
- Improve findability, e.g. include in NetCDF data header
- Provide citation (and other documentations) together with the data
- Revisit appropriate citation granularity

Provide:

- Better integration in data publication WF,
  e.g. to avoid ESGF datasets without citation
- Improve tools to provide and keep track
- Support for humans and machines
- Support for self-organizing MIPs

Stockhause and Ames: CMIP TT Data Citation, EGU 2023, https://doi.org/10.5194/egusphere-egu23-4809

Data Citation TT started in 01/2023

Objectives for a sustainable citation service for CMIP7 and beyond:
1. Prioritise requirements for CMIP7 and potential wider activities.
2. Identify stakeholders
3. Define a maximum of three options for a future citation service.
4. Develop concepts, specify required services, and estimate costs (financial and human resource) for all options.
5. Carry out stakeholder engagement to determine the preferred option.
6. Develop an appropriate governance structure and process for establishment (e.g. Terms of Reference).

Coordination:
- Across CMIP7 TTs
- With ESGF Data Citation WG (ESGF 2.0: ongoing redesign of ESGF)
- With WGCM's strategy on data citation

Stockhause and Ames: CMIP TT Data Citation, EGU 2023, https://doi.org/10.5194/egusphere-egu23-4809

CMIP

Thank you!

References
WIP White Paper: Stockhause, M., Toussaint, F., & Lautenschlager, M. (2015). CMIP6 Data Citation and Long-Term Archival. Zenodo. https://doi.org/10.5281/zenodo.35178
Publication: Stockhause, M. and Lautenschlager, M., 2017. CMIP6 Data Citation of Evolving Data. Data Science Journal, 16, p.30. DOI: http://doi.org/10.5334/dsj-2017-030
PICO (Interactive poster): Stockhause, M., & Lautenschlager, M. (2017). CMIP6 Data Citation and IPCC Data Distribution Centre Services. Zenodo. https://doi.org/10.5281/zenodo.569646

Links
General Information: cmip6cite.wdc-climate.de
Available data: bit.ly/CMIP6_Citation_Search
References: bit.ly/CMIP6_DOI_Statistic

CMIP at EGU:
daily at 12:30: CMIP drop-in at ESA booth #X202 in Exhibition Hall X2
Thu, 27 Apr, 19:00: TM17 CMIP wants to hear from you! Early career researchers' reflections from CMIP6 (Room 1.34)

Stockhause and Ames: CMIP TT Data Citation, EGU 2023, https://doi.org/10.5194/egusphere-egu23-4809