

# AP7/AP8: Citation and Long-Term Archival

DICAD, 18.04.2018  
AWI, Bremerhaven

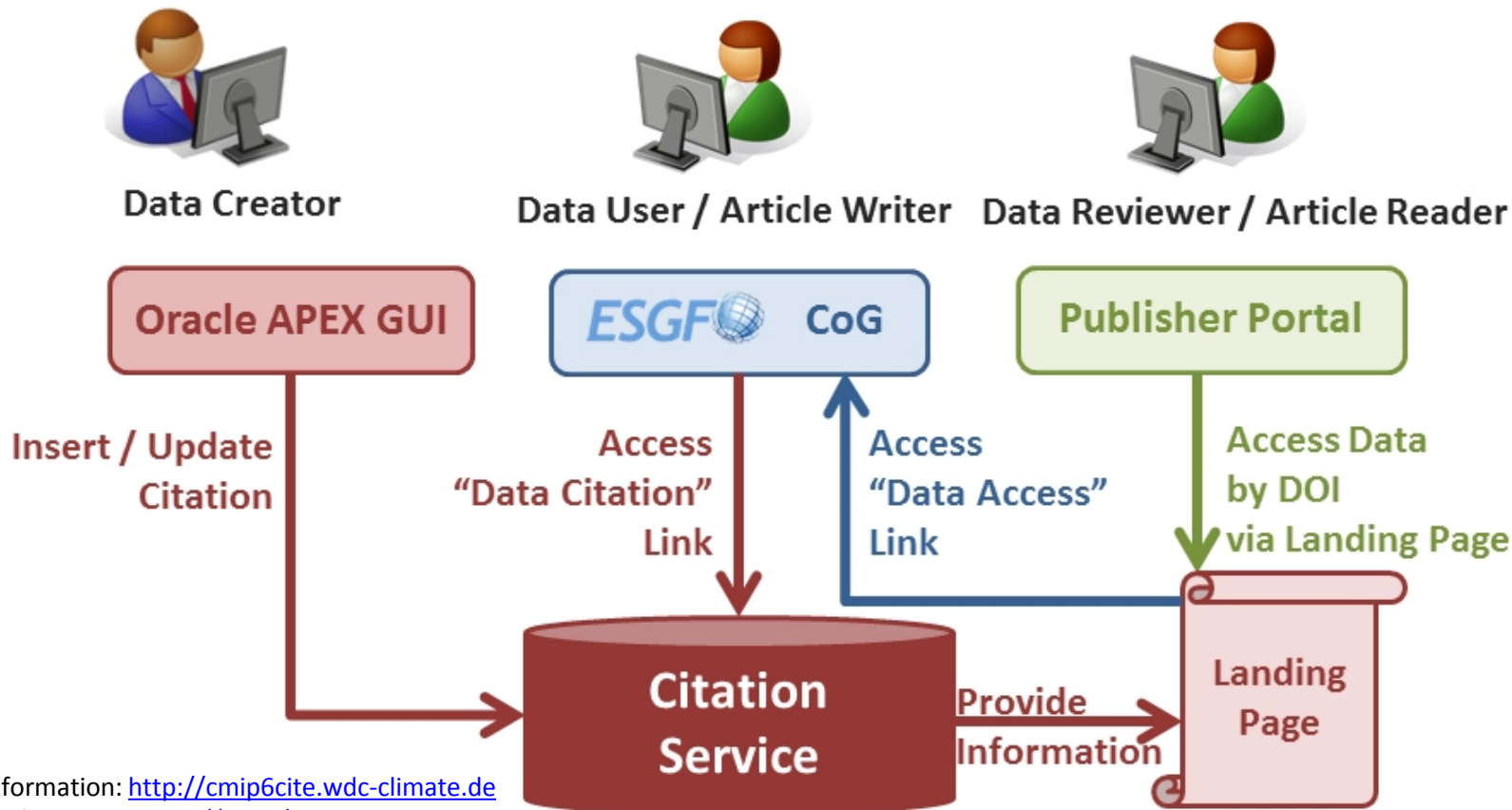
M. Stockhause  
WDC Climate at DKRZ

# AP7/AP8: Progress

## Progress AP7/AP8:

- Position not filled due to illness/pregnancy
- AP7 – Data Citation: work on making prototype operable
- AP7 – Data Citation: input4MIPs republication
- AP8 – Long-Term Archival: IPCC DDC-related tasks

## AP7 - Overview



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

Technical information: <http://bit.ly/1XsVOoz>

# AP7 – Data Citation Service (1)

## Data Citation Service Components:

Split operation from development by set-up of test instance (used in DC)

- **Database:** Duplication of schema in testdb  
+ filling with content related to Data Challenge (DC)
- **Landing Page:**
  - Duplication for test instance
  - Content reviewed by WIP (Karl Taylor)
- **Inserts from CMIP6 Controlled Vocabulary:**
  - **Model information:** insert/update automated
  - **Citation Entries:** model connection included in automated insert process
- **Export to DataCite:** DOI registration and metadata update procedures consolidated
- **OAI Server:** test and operable sets tidies up and in use for DC and input4MIPs

# AP7 – Data Citation Service (2)

## Machine-accessible Interfaces (machine accessible):

Successfully tested within DC2 using test instance

- **ESGF interfaces tested within DC:**
  - Publish citation URL template
  - Include citation URL in Solr index
  - Display citation information: included in CoG;  
**Dissemination API** request and rendering response
- **ES-DOC:** furtherInfoURL – still not set-up by ES-DOC
- **CV:** Discussion how to include DOI information in version information for input4MIPs
- **Data Publication:** machine-readable landing page (schema.org/JSON-LD) - ongoing

# AP7 – Data Citation Service (3)

## Services for data creators:

Insert API tested, improved and announced

- GUI (<http://cera-www.dkrz.de/citeXA>): no further development
- Insert API:
  - API client (python) made available on DKRZ cloud
  - Tested and reviewed by GFDL colleague
- Documentation improved and API usage added
- User Support
- planned Webinar postponed
- Scholix (<http://www.scholix.org>): provide information on data usage in literature
  - early stage of Scholix API testing and quality aspects

# AP7 – Data Citation Service (4)

## CMIP6 Citation Special Cases:

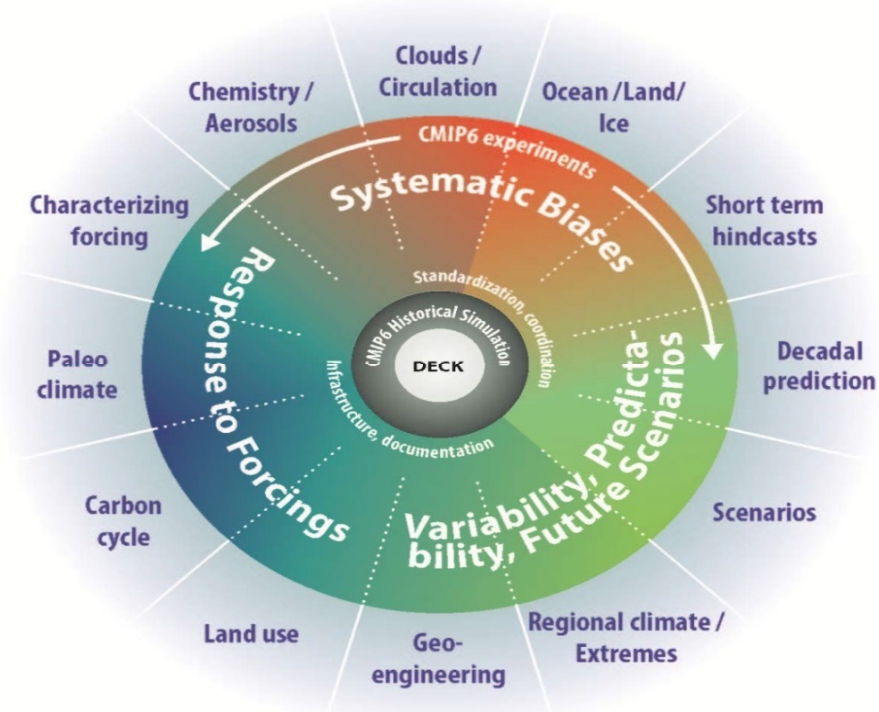
### Data citation ‘submitted’ status

- input4MIPs: re-publication with revised DRS
- HighResMIP / PRIMAVERA: data citation ‘submitted’ status (supported by CEDA):
  - mandatory data citation
  - successful ESGF publication within DC

## CMIP6 Citation Numbers (2018-04-16):

### Ongoing citation manager registrations

- 23 MIPs + DECK with ca. 260 experiments
- 33 institutions with 89 models registered
- 509 model/MIP citation entries (experiment entries postponed to May/June – end of DC)



(Source: Eyring, V. et al. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, [doi:10.5194/gmd-9-1937-2016](https://doi.org/10.5194/gmd-9-1937-2016).)

# AP7 – Data Citation Service (5)

## Request for Information:

Provision of data citation information prior to ESGF publication

- Please provide data citation information via the GUI / API; you can change them (e.g. add further authors) until 10/2020
- Recommendation to provide:
  - ORCIDs for persons
  - Reference to related input4MIPs forcing data
  - References to related papers
- Please give me feedback for GUI / API to help to improve the API functionality and documentation

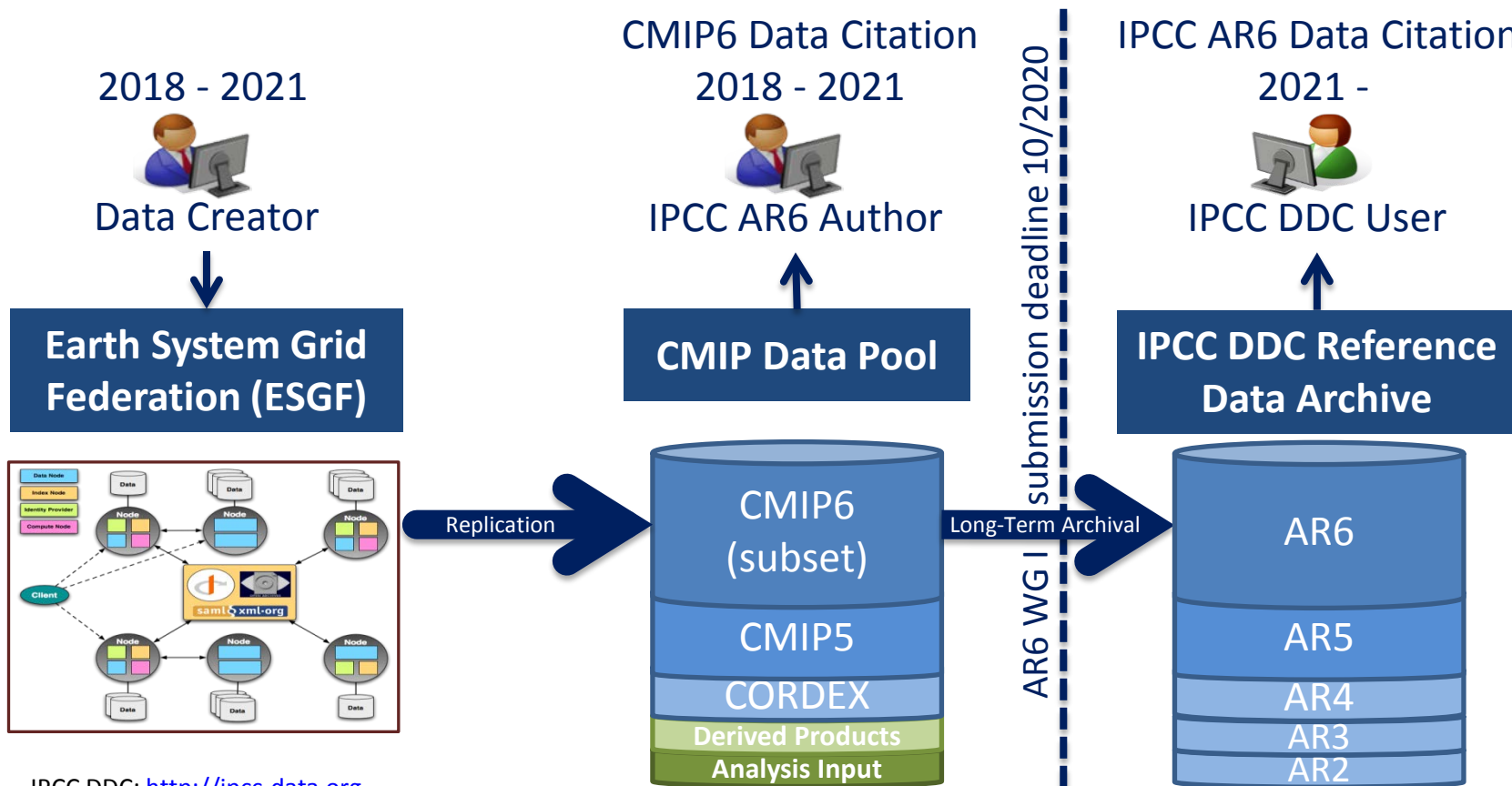


# AP7 – Data Citation Service (6)

## Best Practice to provide data citation information:

1. Use the GUI at <http://cera-www.dkrz.de/citeXA>
  - to provide information on persons, institutes, references
  - to create a single detailed data citation template for API use by exporting the JSON representation from the landing page
2. Prepare for API usage (requirements: python 2.6+):
  - download the [python API client](#)
  - check if the API server is reachable: 'ping ceracite.dkrz.de'
  - store your user credentials in \$HOME/.netrc
3. Use the API
  - modify the JSON template, esp. the 'subject' of type 'DRS'
  - run the API client with the modified JSON file
  - check the result on the landing page (the link is part of the API response)

# AP8 - Overview



IPCC DDC: <http://ipcc-data.org>

# AP8 – Long-Term Archival (1)

## IPCC TG-Data (former TGICA):

Review of IPCC TG-Data / DDC finalized (IPCC-47, 13-16 March 2018)

→ Membership renewal for TG-Data and probably new Co-Chairs

- in intensive cooperation with Christiane Textor (DE-IPCC)
- **New Name:** Task Group on Data Support for Climate Change Assessments (TG-Data)
- **Revised Mandate for TG-Data and DDC Guidance focusing on data:**  
TG-Data supports the DDCs and the DDC Support Group;  
DDC Guidance: *“The purpose of the IPCC DDC is to archive and provide transparency, traceability, and stability of data and scenarios that are relevant in the context of the IPCC.”*
- **Importance for Data within IPCC:**  
*“[IPCC-47] ... 8. Expresses its gratitude to members that support the Technical Support Units (TSUs) and a number of IPCC activities, including data centres, meetings and outreach actions;...”*
- **DDC Resourcing:**  
*“IPCC Bureau to seek additional and more sustainable resourcing options.”*

# AP8 – Long-Term Archival (2)

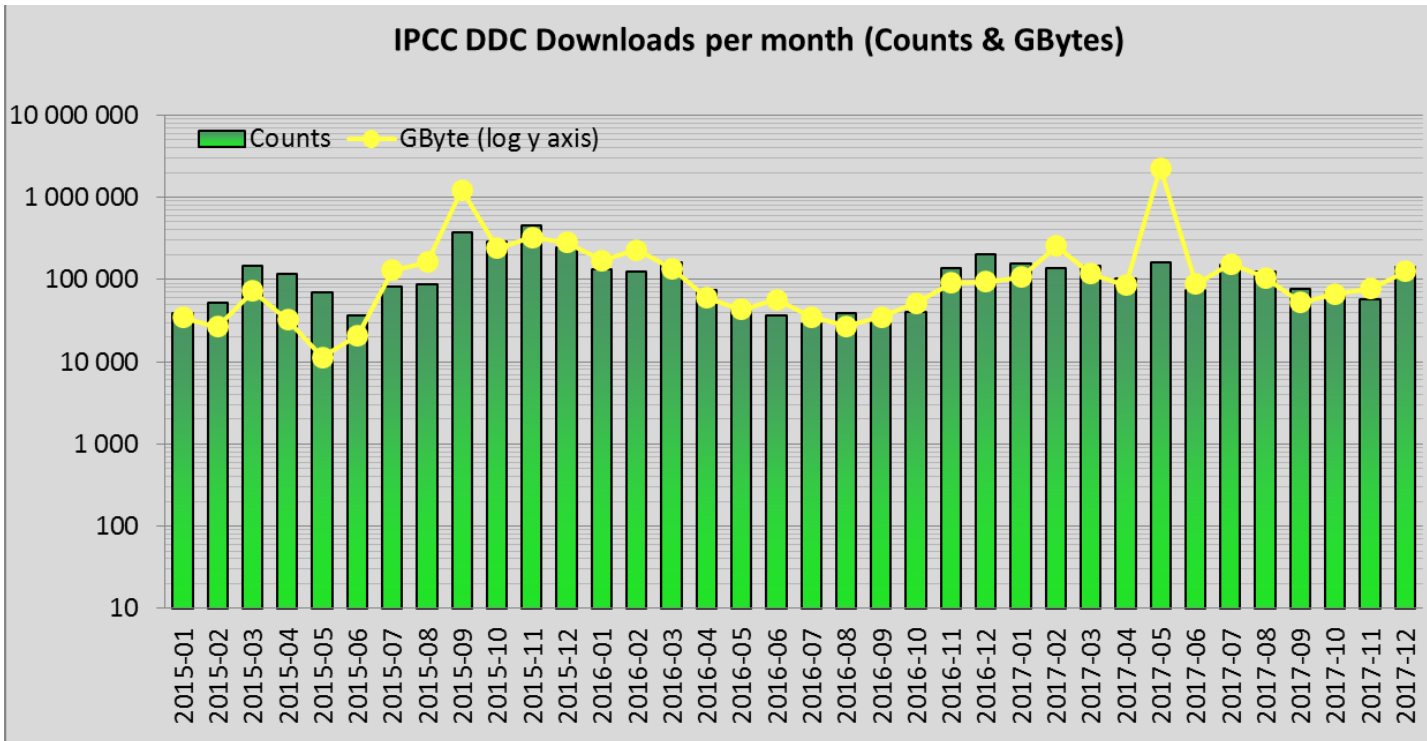
## DDC Support (Group of IPCC WG TSUs and DDC Managers):

Ongoing discussions on support for the IPCC WGs on data

- **DDC Support:**
  - Meetings: WebEx in Feb. 2018 and upcoming F2F with IPCC authors in May 2018 at an IPCC EM
  - Virtual workspaces for IPCC authors at CEDA and DKRZ (“CMIP Data Pool”)
  - Basic provenance of data underlying tables and figures of AR6 (at least part 1) by CMIP6 data (DDC at DKRZ) + analysis script + table/figure data
  - WGI TSU works on data policies / best practices for CLAs
- **DDC:**
  - DDC’s Annual Report 2017 published
  - HAPPI-MIP (SR1.5) LTA to be started in May 2018

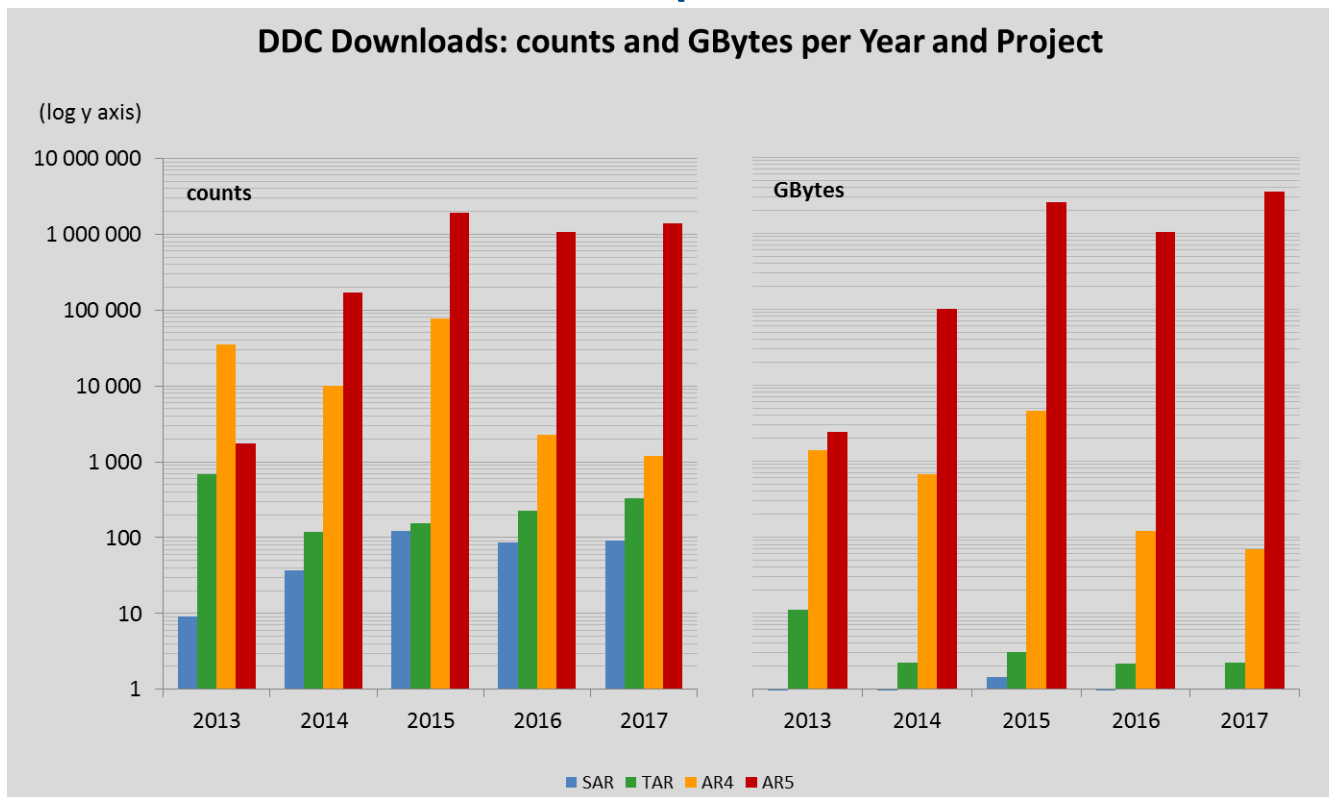
# IPCC DDC Download Rates (1)

## Monthly Data Download Rates:



## IPCC DDC Download Rates (2)

## Annual Data Download Rates per AR:



# AP7/AP8: Status

## Progress AP7/AP8:

- Position not filled due to illness/pregnancy
- AP7 – Data Citation: work on making prototype operable
- AP7 – Data Citation: input4MIPs republication
- AP8 – Long-Term Archival: IPCC DDC-related tasks

## Ongoing AP7/AP8:

- AP7: Data Challenge participation
- AP7: Adaptation of metadata to still-discussed CMIP6 specifications, e.g. facets
- AP7: fine-granular citation entries on experiments (end of DC: May/June 2018)
- AP7: Scholix link access
- AP8: DDC Support activity (IPCC EM: May 2018)
- AP8: Add Reference Data Archive SR1.5/HAPPI-MIP to DDC (starting 15.05.2018)

[cmip6cite.wdc-climate.de](http://cmip6cite.wdc-climate.de)

stockhouse@dkrz.de

 <http://orcid.org/0000-0001-6636-4972>

K. E. Taylor, P. J. Durack, M. Elkin, E. Guilyardi, D. Hassell, M. Lautenschlager, and M. Stockhouse (2017).

CMIP6 Participation Guidance for Modelers. <https://pcmdi.llnl.gov/CMIP6/Guide/modelers.html>

M. Stockhouse, and M. Lautenschlager (2017). CMIP6 Data Citation of Evolving Data.

Data Science Journal. 16, p.30. doi:10.5334/dsj-2017-030. <http://doi.org/10.5334/dsj-2017-030>.

M. Stockhouse, and M. Lautenschlager (2017). CMIP6 Data Citation and IPCC Data Distribution Centre Services.

PICO at EGU 2017, Vienna. Zenodo. <http://doi.org/10.5281/zenodo.569646>.