

CMIP6 Citation Service and Long-Term Archival / DDC

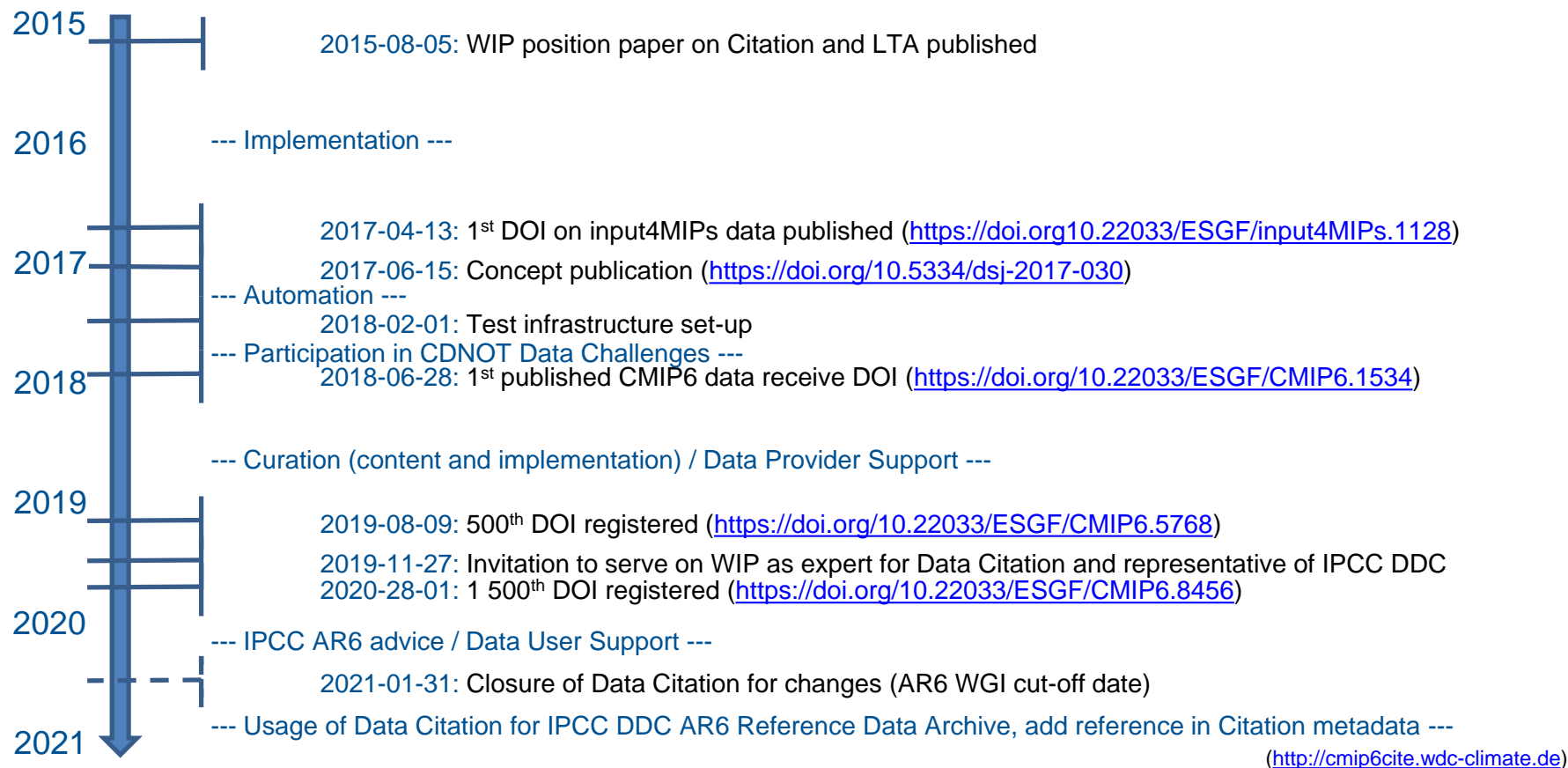
DICAD Final Meeting, 2020-06-22

M. Stockhause

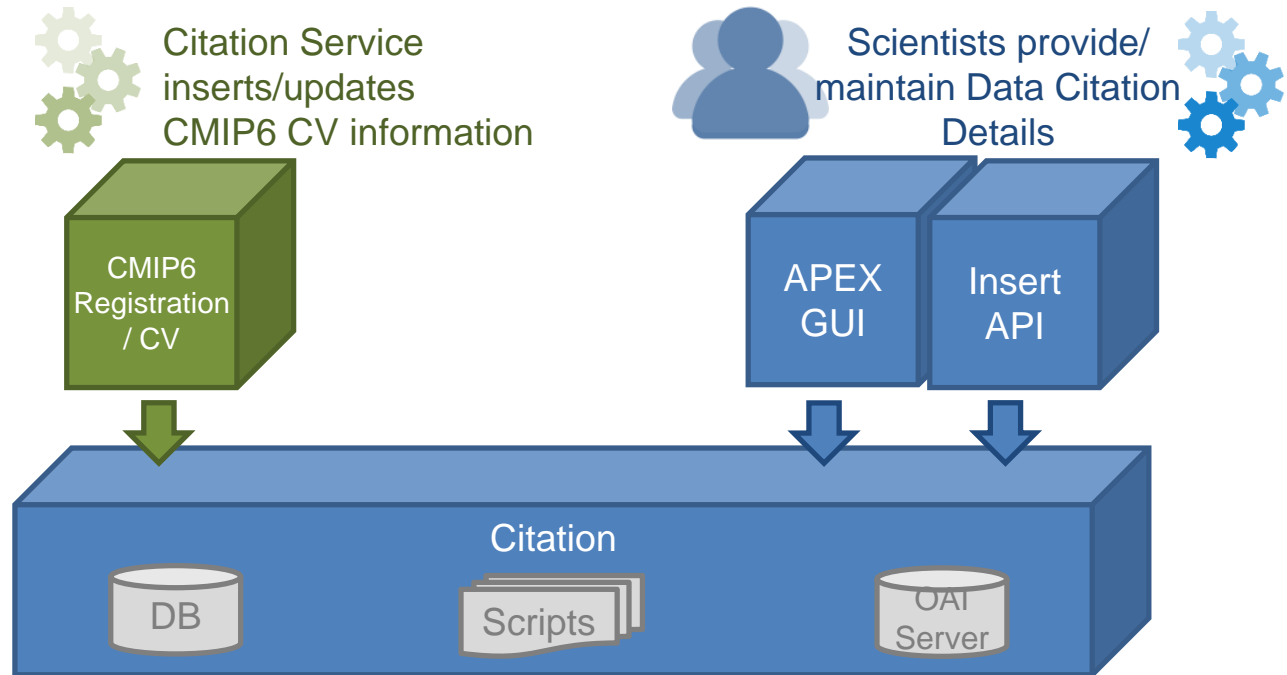
I. CMIP6 Citation Service (~AP7):

For the first time in the history of CMIP, data is citable.

Citation Service Development

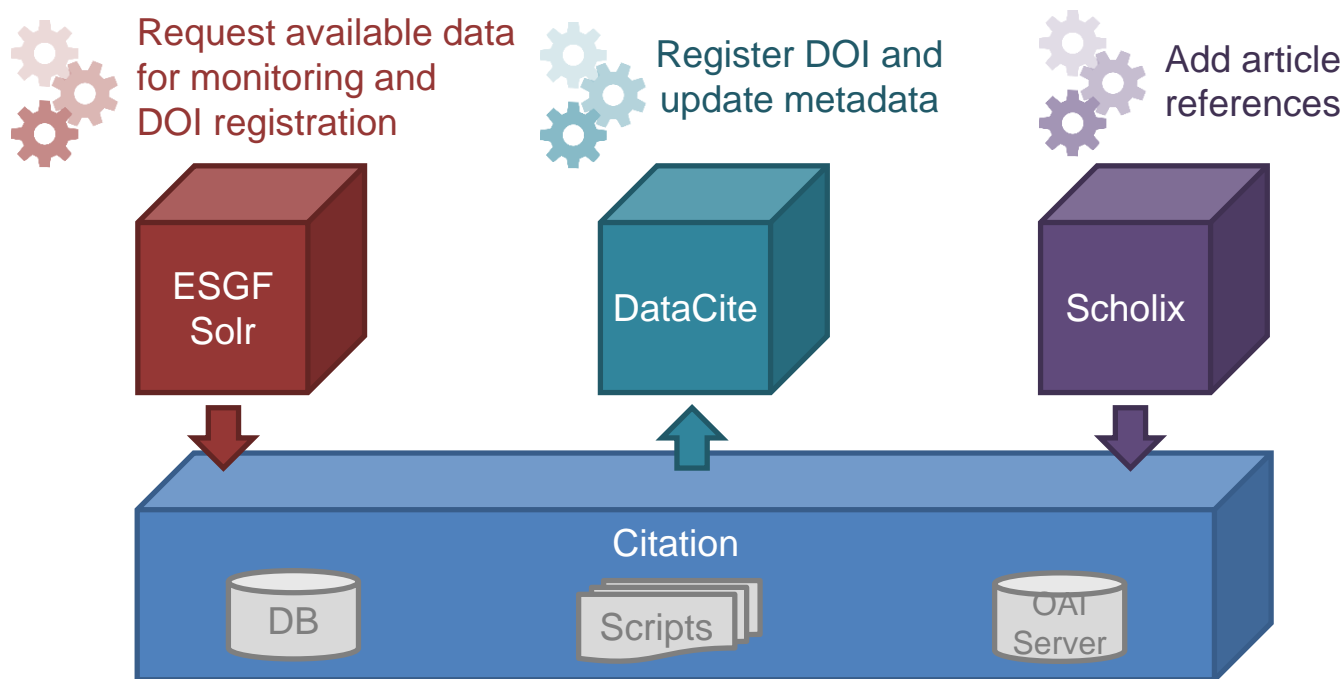


Citation Service Functionality (1): Collect Information



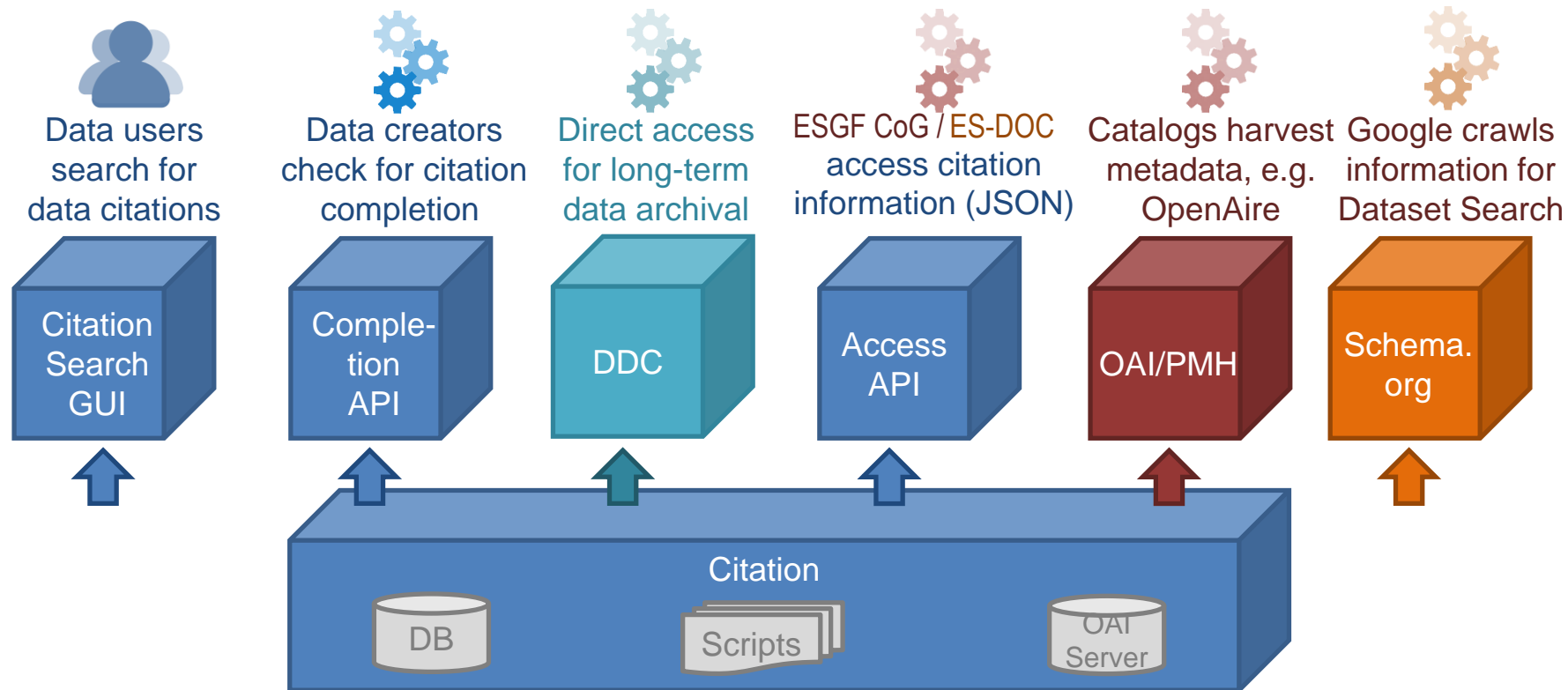
(<http://cmip6cite.wdc-climate.de>)

Citation Service Functionality (2): Process Information



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

Citation Service Functionality (3): Disseminate Information



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

DOI Status (2020-06-18):

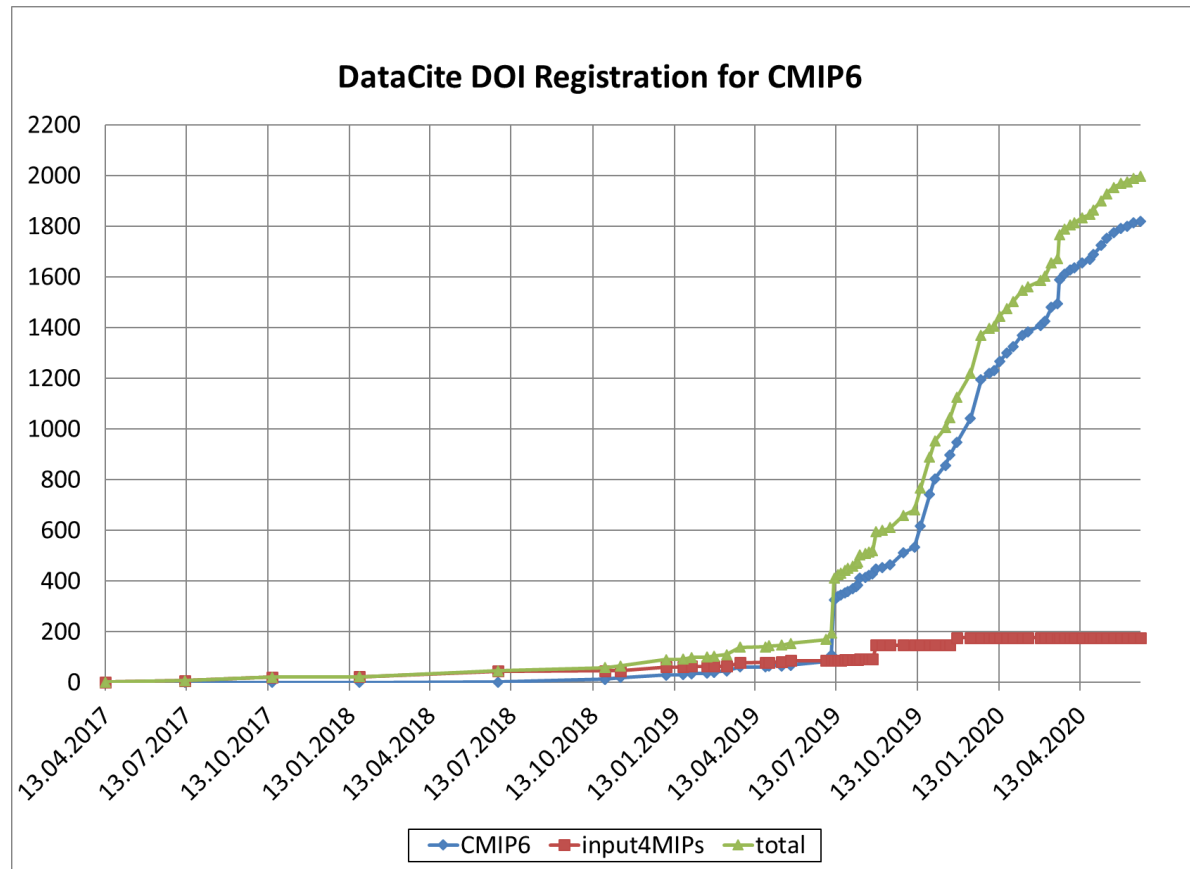
- Number of DOIs: 1 996
 - CMIP6: 1 819
 - input4MIPs: 177
- Number of MD updates: 8 502
- Number of Scholix ref.: 36

- Missing Citations (total): 50
 - CMIP6: 47
 - input4MIPs: 3

→ 97.5 % of the registered CMIP6 data is citable.

References:

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



Citation Service 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

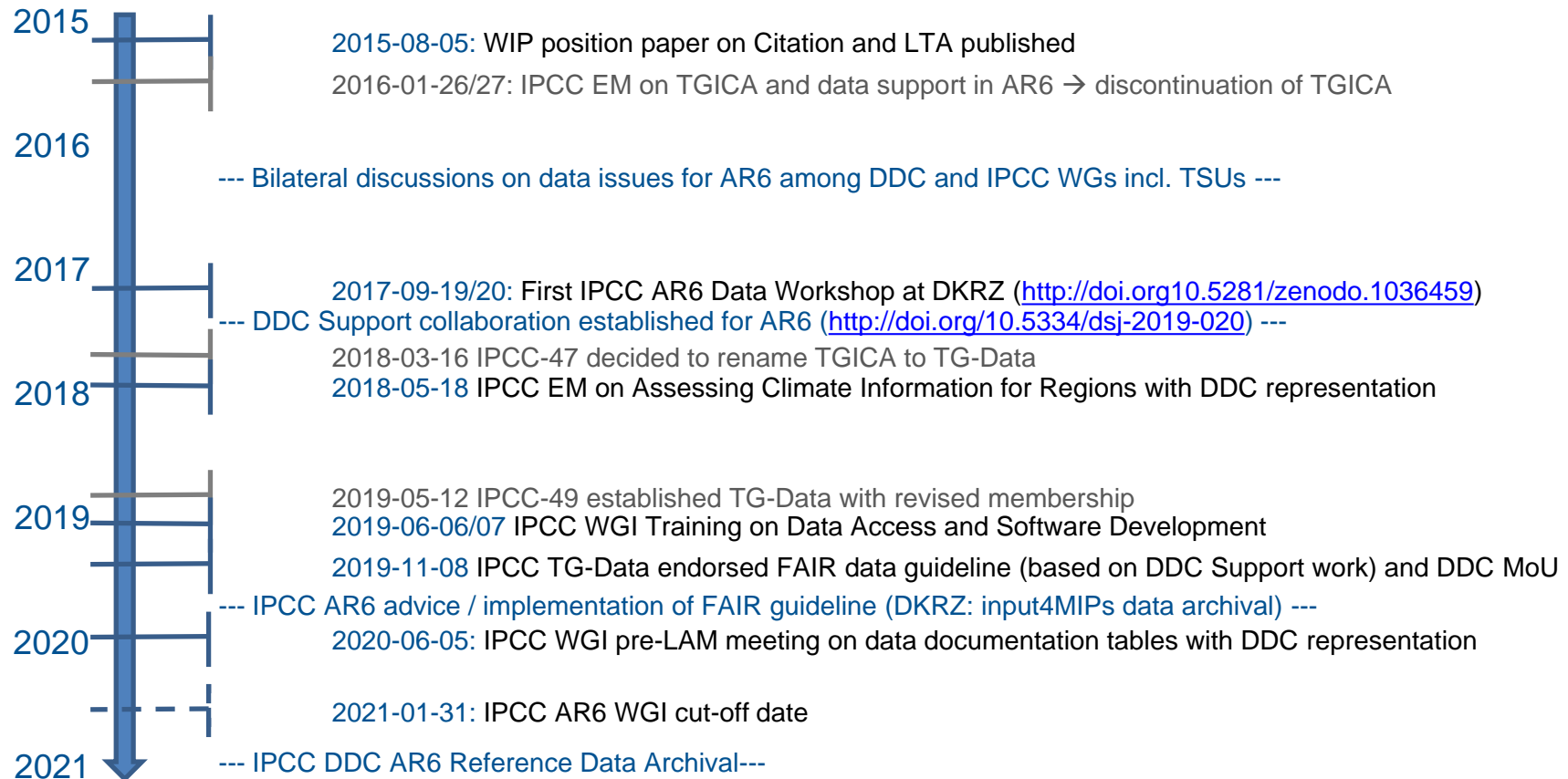
- Sustainable funding or institutional support required; as user service this is staff-intensive
- Improved integration within CMIP infrastructure
- Integration into DKRZ ESGF data publication workflow

II. CMIP6 Long-Term Archival / DDC (~AP8)

IPCC DDC AR6 Reference Data Archive

Contributors: A. Kaiser, H. Widmann, H. Ramthun

DDC Support for AR6



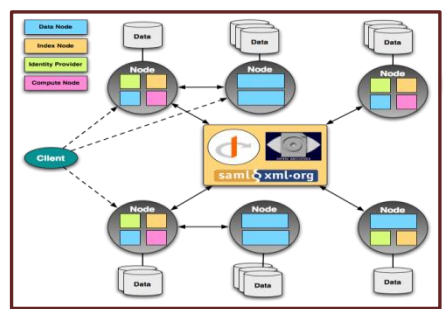
CMIP6 Data Creation
2018 -



Data Creator



Earth System Grid Federation (ESGF)



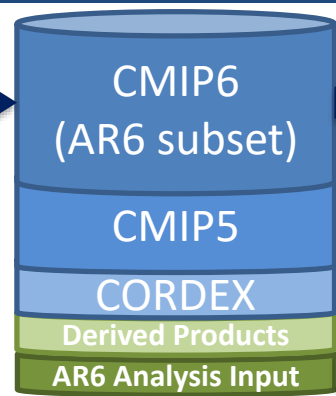
CMIP6 Data Use
2018 - 2020



IPCC AR6 Author



Virtual Workspace



submission deadline 2021-01-31

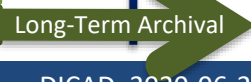
IPCC AR6 Ref Data Archive
2021 -



IPCC DDC User



IPCC DDC Reference Data Archive



1. input4MIPs and CMIP6 (subset) LTA with adapted “**ESGF Workflow**”:
 - CMIP6 subset specified by IPCC AR6 authors
 - Adaptations for ancillary metadata and the “new” CERA GUI plus CMIP6 agreements and the elimination of other manual inserts/updates
 - Extended quality checks after each step in the LTA workflow
2. Additional AR6 source dataset LTA with “**WDCC Workflow**”
3. IPCC DDC management
 - DDC Support collaborated on data aspects and support for AR6 and developed the framework of the FAIR data guidelines, which IPCC TG-Data endorsed in 11/2019.
 - Virtual Workspace as DDC Support activity

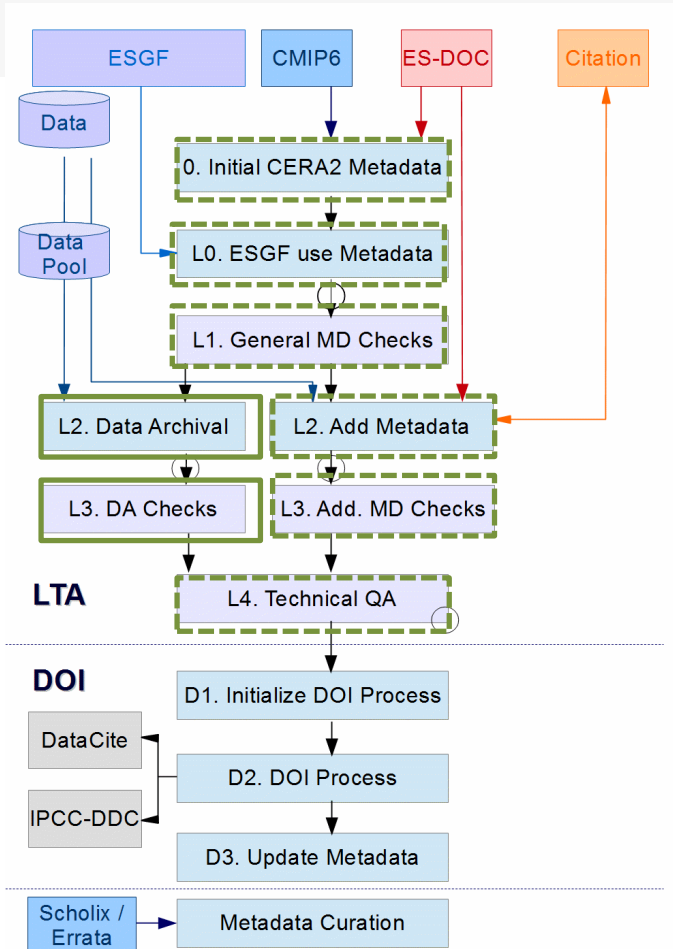
LTA / DDC: ESGF + WDCC Workflows – Required Timeline

start	end	to dos
	12/2019	Adaptations of LTA Metadata (CMIP6_Metadata): * Creation of XMLs * XML load * Insertion of metadata out of cmip6_cite_test Other Adaptations: * Set-up of DOI landing page * Adapt TQA * Adapt DataCite interface
01/2020	07/2020	LTA/DOI for input 4MIPs data as test case for CMIP6
08/2020	09/2020	Preparation for CMIP6 LTA (CMIP6_LTA_Coordination): * Revision of LTA workflow and tools * Extensions for ES-DOC, ESMVal etc. * Discussions with replication team about snap-shot date (30.09. or 18.10.2020), ESGF search attributes, etc. Adaptation of shown data reference recommendation: * ESGF API * CMIP6 landing page
10/2020	10/2021	LTA/DOI for CMIP6 and Creation of IPCC DDC AR6 Reference Data Archive
ca. 01/2021	03/2021	Past-Project functionalities (CMIP6_LTA_Coordination): * Define ES-DOC-Errata interface and code functionality to add errata information to LTA * Define Update Functionality for cmip6_cite_test and implement references to LTA

Additional AR6 source data LTA

Thanks to COVID-19 and the postponement of WGI's cut-off date, the LTA is back on schedule!

Date: 09.08.2019



Current State:

Core LTA components exist; LTA procedure and QC-Software is under review (based on first input4MIPs experiment)

0. project, code list, and summaries done;

ES-DOC or CV exp. information to be added to these summaries; model + resol. in CERA2?

L0. Minor adaptations remaining

L1. QC checks under review

L2. data existing and citation metadata existing; ES-DOC metadata missing, ESMVAL?

L3. data existing; citation metadata existing; further metadata missing

L4. TQA-Software under final review

DOI (D1-D3): Adaptations pending

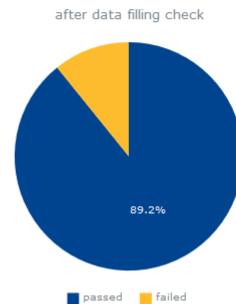
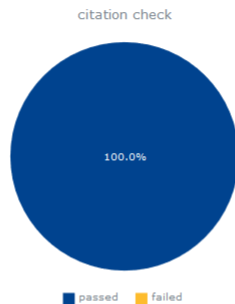
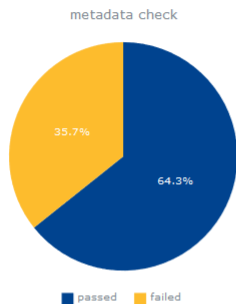
Scholix: existing but to be improved

Errata: missing

Author: Amandine Kaiser
 Date: 2020-06-09 10:47
 TQA Version: 1.0
 Run Type: final
 Performed checks: metadata check, citation check, after data filling check
 Input entry name: input4MIPs CMIP6 HighResMIP MOHC
 Input entry type: experiment

QC Result of first input4MIPs experiment

Passed checks by category:



Passed checks in detail:

Type something in the input field to search the table:

Explanation: When a check passed error = 0. Otherwise, error = 1. Moreover, a distinction is made between three levels of error severity: error, warning, info. By clicking on the table headers the columns can be sorted.

entry id	entry type	error	check	procedure name	message	level	entry name	entry acronym
3878783	dataset	1	metadata check	CMIP6_METADATA_CHECK.spatial_coverage_check	ORA-20000: error raised in: APP_QUALITY_CONTROL_CMIP6_METADATA_CHECK.SPATIAL_COVERAGE_CHECK - 278 - PACKAGE BODY - The spatial_coverage.min_lat value is not equal to the location.min_lat value. The spatial_coverage.max_lat value is not equal to the location.max_lat value. The spatial_coverage.min_lon value is not equal to the location.min_lon value. The spatial_coverage.max_lon value is not equal to the location.max_lon value.	error	input4MIPs CMIP6 HighResMIP MOHC MOHC-HadISST-2-2-0-0-0 ocean day tos gn v20170505	CMIP6_input4MIPs_3878783
3878783	dataset	0	metadata check	CMIP6_METADATA_CHECK.format_check	The format is correct.	error	input4MIPs CMIP6 HighResMIP MOHC MOHC-HadISST-2-2-0-0-0 ocean day tos gn v20170505	CMIP6_input4MIPs_3878783
3878783	dataset	1	metadata check	CMIP6_METADATA_CHECK.parameter_check	ORA-20000: error raised in: APP_QUALITY_CONTROL_CMIP6_METADATA_CHECK.PARAMETER_CHECK - 526 - PACKAGE BODY - Either the parameter entry does not exists and or the access_id is not 0. Note: CERA2_ADM.archive_input4mips.DATA_FILLED is null.	error	input4MIPs CMIP6 HighResMIP MOHC MOHC-HadISST-2-2-0-0-0 ocean day tos gn v20170505	CMIP6_input4MIPs_3878783