

A webtool for accessing CMIP6 Data Citations

Karsten Peters (peters@dkrz.de) and Martina Stockhause (stockhause@dkrz.de)
(DKRZ, Hamburg, Germany)

June 2020

Table of Contents

Introduction	1
CMIP6 controlled vocabulary used for filtering.....	2
Simple search (“google-like”).....	3
Searching by CMIP6 contribution	3
Searching by CMIP6 categories.....	3
Using the advanced search (Actions -> Filter)	4
Filtering for multiple entries at once	5
Downloading the search results.....	6
CMIP6 citation search: use cases	7
Selecting a subset of experiments	7
Use case	7
Solution	7
Searching for distinct contributions to one specific experiment.....	7
Use case	7
Solution	7
Searching for contributions by institution and experiment	8
Use case	8
Solution	8

Introduction

Accessing CMIP6 data citations after having already performed a scientific study can be time consuming. A new web tool provided within the framework of the CMIP6 Citation Service (<https://cmip6cite.wdc-climate.de/>) makes the search for the CMIP6 data citations straightforward.

The web tool can be accessed via http://bit.ly/CMIP6_Citation_Search and provides quick and easy subsetting functionalities of the entire CMIP6 citation database.

When accessing the online tool, the **citation details in clear text** are displayed in the column “Data Reference”. If you want to access the official data citation landing page, which also

offers the opportunity to **download the citation in BibTeX format**, click on the DOI-link the the column “DOI”.

Data citations are available at two levels of granularity – either for model contributions to an entire MIP, e.g. all contributions of one particular model to AerChemMIP, or for contributions to a specific experiment, e.g. the contribution of one particular model to the piClim-CH4 experiment of AerChemMIP. Depending on the specific scientific use case, one or the other citation may be required.

The **query results** of the search requests **can be downloaded** as csv, HTML or PDF document.

This userguide contains **documentation of the basic functionalities** as well specific documentation of **hands-on use cases**.

Further help topics, particularly on the technical aspects of the search tool which are non-specific to the CMIP6 citation database, can be accessed via **Actions -> Help**.

CMIP6 controlled vocabulary used for filtering

The CMIP6 citations in the database can be filtered by the CMIP6 controlled vocabulary:

- DRS_ID
 - CMIP6 identifier of the dataset, e.g. CMIP6.AerChemMIP.BCC.BCC-ESM1.piClim-CH4
- MIP_ERA
 - Identifier of the CMIP-organised iteration, e.g. CMIP6
- ACTIVITY_ID
 - Identifier of the CMIP6 endorsed MIP, e.g. CFMIP
- INSTITUTION_ID
 - Identifier of the participating scientific institution, e.g. CMCC
- SOURCE_ID
 - Identifier of the participating model, e.g. BCC-ESM1
- EXPERIMENT_ID
 - Identifier of the experiment performed in the corresponding MIP (selected by “ACTIVITY_ID”), e.g. “aqua_control” in “CFMIP”

The controlled vocabulary can be filtered so that only the exact dataset(s) you require the citation(s) for remain in the list.

Two approaches can be used for arriving at the citation you need: searching through the database using a standard “google-like” search functionality (🔍) or using an advanced search to filter through a given list of field entries. That option also allows for subsetting by multiple search expressions at once.

Simple search (“google-like”)

The simplest way to explore the citation database is by using a google-like free text search using the 🔍 tool.

Two approaches are available

- Searching through all the available CMIP6 contributions
- Searching through the available entries of the CMIP6 categories

Searching by CMIP6 contribution

The screenshot shows a search bar with 'csiro' entered. Below it, a search report is displayed with the filter 'Row text contains 'csiro''. Below the report, it states '50 CSIRO-related CMIP6 contributions (as of May 2020)'. Blue arrows indicate the flow from the search input to the report and then to the result count.

The default search is a complete search over all entries possible by free text (not case sensitive)

Subsetting the search results then requires a new search request

Searching by CMIP6 categories

The screenshot shows a search interface with a table of results. A red circle highlights the 'All Columns' dropdown menu, which lists various columns: DRS_ID, MIP_ERA, ACTIVITY_ID, INSTITUTION_ID, SOURCE_ID, EXPERIMENT_ID, DOI, and Data Reference. The table below shows columns MIP_ERA, ACTIVITY_ID, INSTITUTION_ID, and SOURCE_ID with corresponding values.

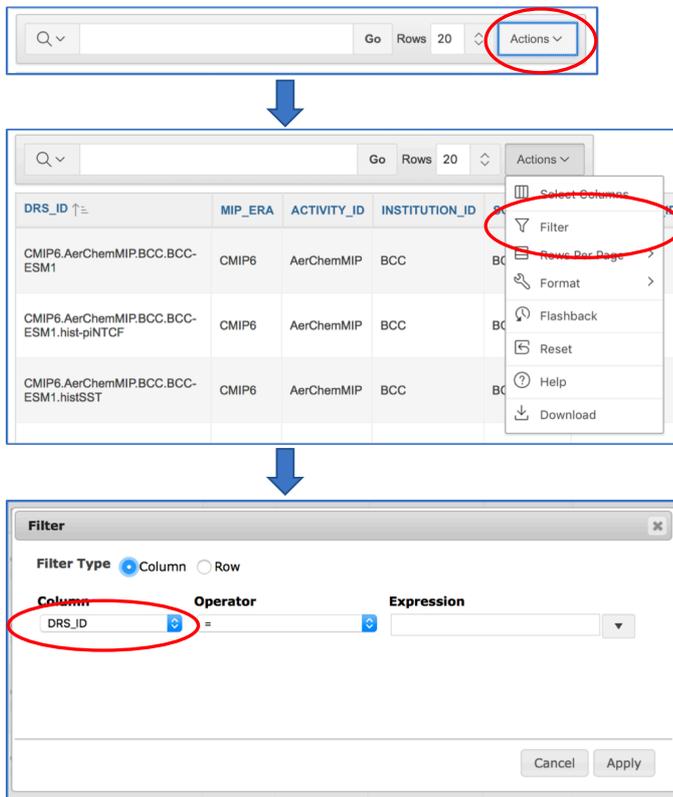
MIP_ERA	ACTIVITY_ID	INSTITUTION_ID	SOURCE_ID
CMIP6	AerChemMIP	BCC	BCC-ESM1
CMIP6	AerChemMIP	BCC	BCC-ESM1
CMIP6	AerChemMIP	BCC	BCC-ESM1
CMIP6	AerChemMIP	BCC	BCC-ESM1

A faceted search over distinct columns is also possible by selecting the appropriate entry from the drop-down menu.



Further subsetting is possible by executing multiple search requests, e.g. search for individual institutions (via "INSTITUTION_ID") in the already subsetting list

Using the advanced search (Actions -> Filter)



Click on Actions

Select „Filter“ from the dropdown list

The Filter window pops up.

Columns contain the CMIP6 controlled vocabulary entries.

Filter dialog box. Filter Type: Column. A list of columns is shown on the left, with 'ACTIVITY_ID' circled in red. The Operator is set to '=' and the Expression field is empty.

Select the entry to filter for, e.g. ACTIVITY_ID

Filter dialog box. Column: ACTIVITY_ID. Operator: =. A dropdown list of values is shown on the right, with 'FAFMIP' circled in red. The Expression field contains 'ACTIVITY_ID = FAFMIP'.

Expression lists the corresponding entries.

Select the desired activity from the dropdown list, e.g. FAFMIP, and click „Apply“

Table header area. A filter tag is visible: ACTIVITY_ID = 'FAFMIP' x

Filter is applied

Table with 3 rows of filtered data. Each row has a green filter icon and a close button. The rows are:

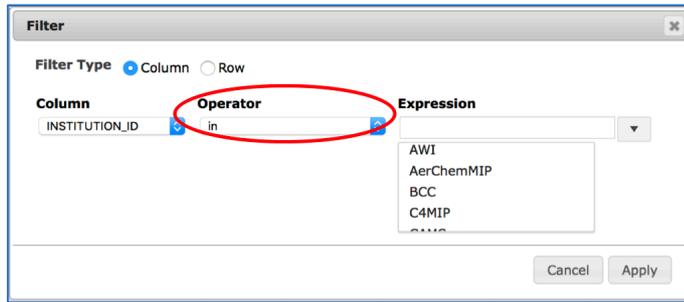
- ACTIVITY_ID = 'FAFMIP' x
- EXPERIMENT_ID = 'faf-heat' x
- SOURCE_ID = 'CanESM5' x

Continue until the list of datasets is reduced so far that you can access the one you need.

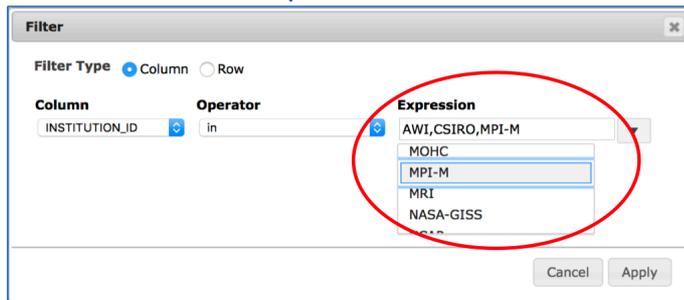
DRS_ID ↑	MIP_ERA	ACTIVITY_ID	INSTITUTION_ID	SOURCE_ID	EXPERIMENT_ID	DOI	Data Reference
CMIP6.FAFMIP.CCCma.CanESM5.faf-heat	CMIP6	FAFMIP	CCCma	CanESM5	faf-heat	https://doi.org/10.22033/ESGF/CMIP6.3590	Swart, Neil Cameron; Cole, Jason N.S.; Kharin, Viatcheslav V.; Lazare, Mike; Schinocca, John F.; Gillett, Nathan P.; Anstey, James; Arora, Vivek; Christian, James R.; Jiao, Yanjun; Lee, Warren G.; Majaess, Fouad; Saenko, Oleg A.; Sellar, Christian; Seinen, Clint; Shao, Andrew; Solheim, Larry; von Salzen, Knut; Yang, Duo; Wintler, Barbara; Sigmund, Michael (2019). CCCma CanESM5 mode output prepared for CMIP6 FAFMIP faf-heat. Earth System Grid Federation. doi: https://doi.org/10.22033/ESGF/CMIP6.3590

Filtering for multiple entries at once

The advanced filtering functionality also offers the capability to select for multiple entries at once, i.e. the classical “.OR.” conditional query.



Select the “in” Operator for conditional subsetting (equivalent to an “OR.” expression)



Select multiple entries from the drop-down list scrolling through it and clicking on the entries.

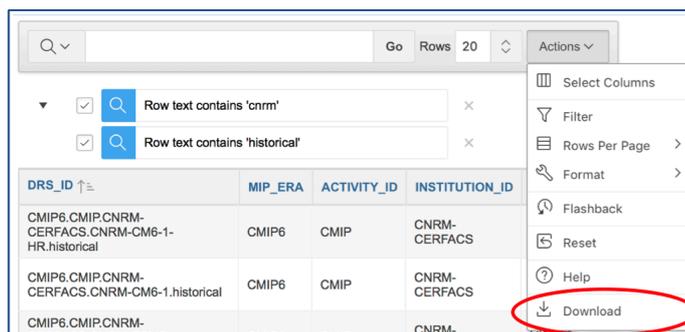


Filter is applied

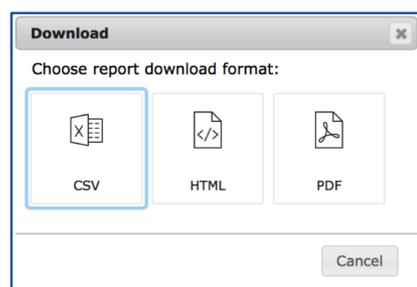
Again, if you want to **filter for more entries**, you can add new subsetting options by executing a **new filter request** on the already subsetting query result.

Downloading the search results

When you have arrived at your desired subset of CMIP6 citations, it can be downloaded for later reference.



Click on “Actions” and select the “Download” option



Select the file format of your preference for download

CMIP6 citation search: use cases

Oftentimes, researcher only use specific parts of the data available in CMIP6. Here, we showcase a few selected use cases.

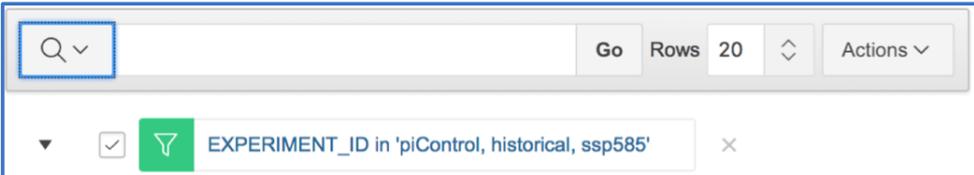
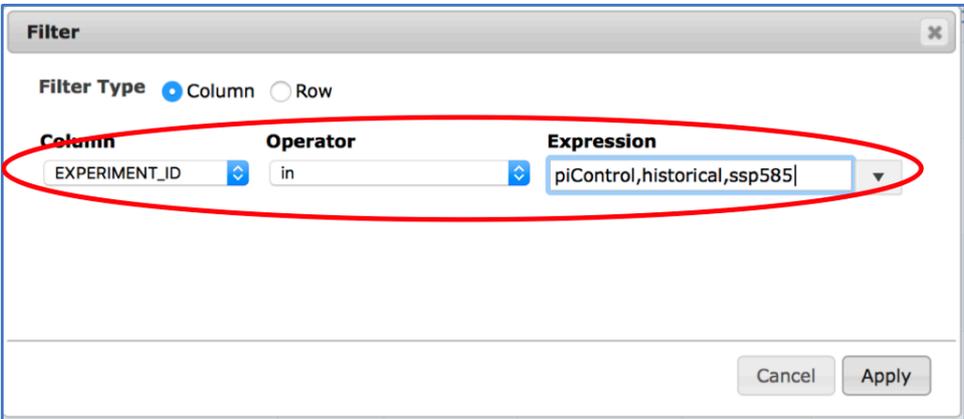
Selecting a subset of experiments

Use case

“I have downloaded lots of CMIP6 data, and I have used a subset of that data. My analysis is limited to three experiments: the piControl, the historical, and the ssp585 simulations.”

Solution

Use the advanced filter option to subset the data by “EXPERIMENT_ID” and using the “in” operator.



145 contributions available (as of May 2020)

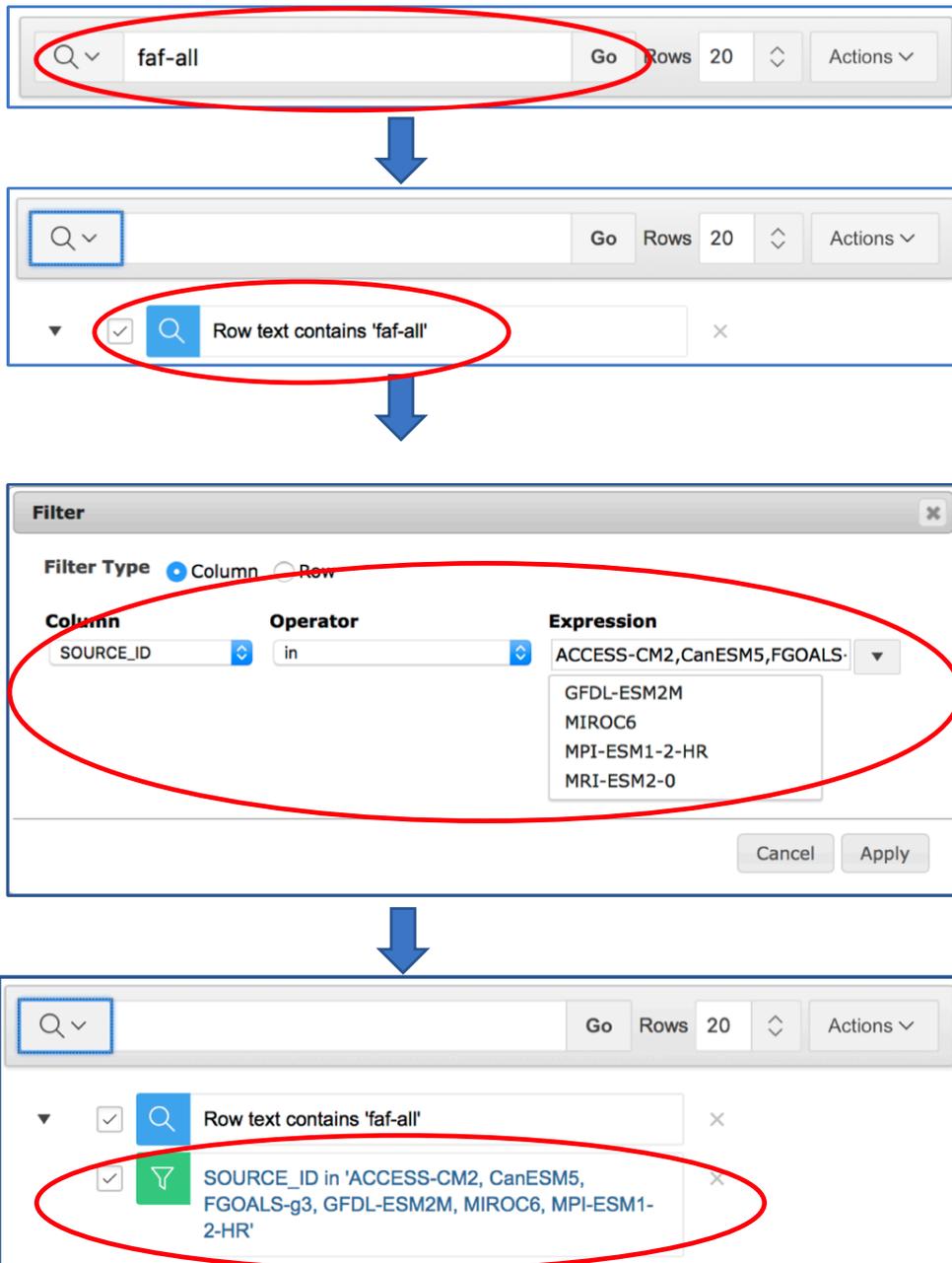
Searching for distinct contributions to one specific experiment

Use case

“Suppose I've analyzed 6 of the 7 models that have contributed output to the FAFMIP faf-all experiment.”

Solution

Use a combination of the simple search functionality (🔍) and the advanced search to quickly search through the entire database first and then fine-tune the results.



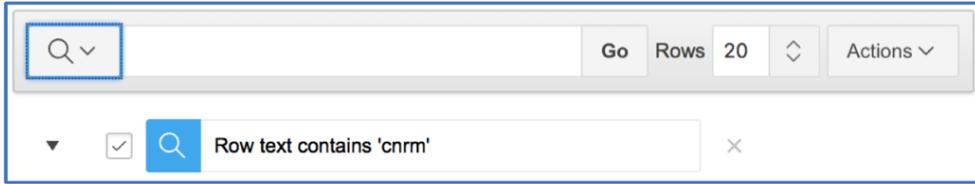
Searching for contributions by institution and experiment

Use case

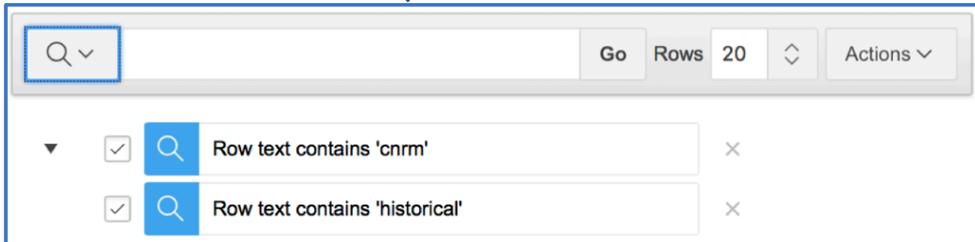
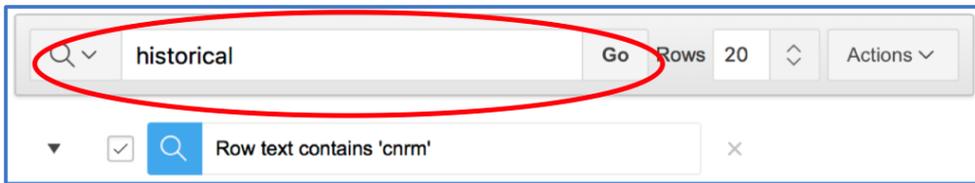
“Suppose I analyzed the CMIP6 historical simulations of the available CNRM models”

Solution

Use the 🔍 functionality in two separate requests.



176 contributions available (as of May 2020)



3 contributions available (as of May 2020)