Version 0.2 (MS-15.11.2016)	GUI part
Version 0.5 (MS-19.02.2018)	API added
Version 1.0 (MS-26.03.2018)	API error documentation added in section 2.6
Version 1.1 (MS-16.08.2018)	GUI update
Version 1.2 (MS-17.09.2018)	GUI update (person and institute sections)
Version 1.3 (AL-29.01.2019)	API checking tool (Section 2.3)
Version 1.4 (MS-01.08.2019)	GUI update (connect lists; completed citations sections)
Version 1.5 (MS-24.05.2021)	GUI update (visibility and restriction to entries without DOIs)

This user guide gives a step-by-step introduction, how to create and maintain data references for CMIP6 data collections on **model/MIP** and **experiment** granularities. These data citations will be visible in the <u>ESGF CoG portals</u>. Users of CMIP6 data are obliged to use these data citations in the reference lists of their publications (<u>CMIP6 data user guide</u> - under development). Data references can be modified prior to DOI registration to keep data references stable, which have been used in the IPCC AR6 or other publications. The CMIP6 data subset used in the AR6 will form the AR6 Reference Data Archive of the IPCC Data Distribution Centre (IPCC DDC, <a href="https://www.ipcc-data.org">http://www.ipcc-data.org</a>).

#### **Further Information:**

- CMIP6 Participation Guidance for Modelers: <a href="https://pcmdi.llnl.gov/CMIP6/Guide/modelers.html">https://pcmdi.llnl.gov/CMIP6/Guide/modelers.html</a>
- Information on the CMIP6 citation service: http://cmip6cite.wdc-climate.de

## **Citation Service Tools for Citation Managers:**

Graphical User Interface: <a href="http://cera-www.dkrz.de/citeXA">http://cera-www.dkrz.de/citeXA</a>

 Machine access (API client): https://swiftbrowser.dkrz.de/public/dkrz 11279a46963f4201bae564a253d528cc/Citation API Client/

#### **Overview**

The following steps are required or recommended to create citations for your CMIP6 data:

- 1. **GUI Edit Title:** Change the given default titles for the given CMIP6 model and experiment data collections.
- 2. **GUI Create Lists of Persons and Institutes:** Enter lists of persons and institutes to be connected in step 3 with a special role to a data title (step 1).
  - Note: Required for step 3 are a list of *Creators (authors)* and a list of *ContactPersons (corresponding author)*. The position in the list is interpreted as position in the (author) list.
- GUI Build/Edit Data Citation: Connect your Creator (author) list and your ContactPerson (corresponding author) list from step 2 to a citation title from step 1 [mandatory].
   Optionally, a citation can be linked to other lists of contributors from step 2, e.g. Funders or ResearchGroups, who contributed to the creation of the selected model/MIP or experiment data. More information on roles is provided in the DataCite Metadata Schema documentation at http://schema.datacite.org.
- 4. **GUI Add References** to articles, input4MIPs and other data or software [recommended].
- 5. API Provide Data Citations for the experiment granularity (chapter 2)

### **Notes: 1.** Please do not use the API any more.

2. You can control the visibility of the entries in the 'Visibility' tab and check all data citations in the 'Completed Citations' tab of the GUI.

# **Table of Contents**

1.	GUI	- Graphical User Interface	3
	1.1	Login	3
	1.2	Home Area	4
	1.3	CMIP6 Citation area –main tab for providing citation information	5
	Step	o 1 – Provide a title for a data citation entry	6
	Step	o 2 – Provide named author/creator and other lists	7
	Step	o 3 – Connect author list and other lists to the data reference	9
	Step	o 4 – Connect paper or other reference to the data reference (recommended)	11
	1.4	Control Visibility of data citation entries	12
	1.5	List of Completed Citations and landing pages	13
	1.6	Create a person and institute entries	14
	a.	Create a new person entry	14
	b.	Add ORCID Researcher ID to person record (recommended)	15
	1.7	Add Paper and Data References	16
	1.8	Upload Area	17
2.	API	- Application Programming Interface	18
	2.1	Preconditions	18
	2.2	Best Practice	18
	2.3	Checking tool (without an entry in the database)	19
	2.4	Run the API	20
	2.5	Check citation information	20
	2.6	Overview of top-level elements (JSON file) for citation	21
	2.7	JSON Example (some unread sections removed)	22
	2.8	Error codes and messages	23

nobody Feedback Logout

# 1. GUI - Graphical User Interface

# 1.1 Login

Login is provided at: <a href="http://cera-www.dkrz.de/citeXA">http://cera-www.dkrz.de/citeXA</a>.

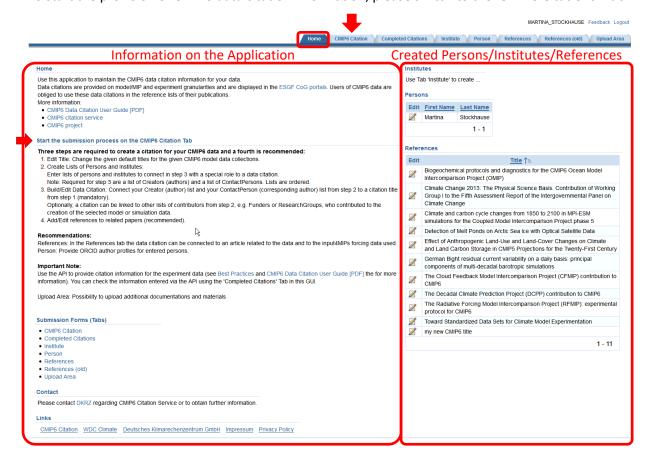
Use your CERA username and password to login. In case you have no CERA account or encounter login problems, please write an email to <a href="mailto:stockhause@dkrz.de">stockhause@dkrz.de</a> with subject 'CMIP6 Citation...'



# 1.2 Home Area

After login you are redirected to the 'Home' area, which provides information around the CMIP6 citation service (left part). If you have entered new persons, institutes or references, these will be listed on the right. After a while these might disappear.

To start the provision of CMIP6 data citation information, please switch to the 'CMIP6 Citations' Tab.



# 1.3 CMIP6 Citation area -main tab for providing citation information

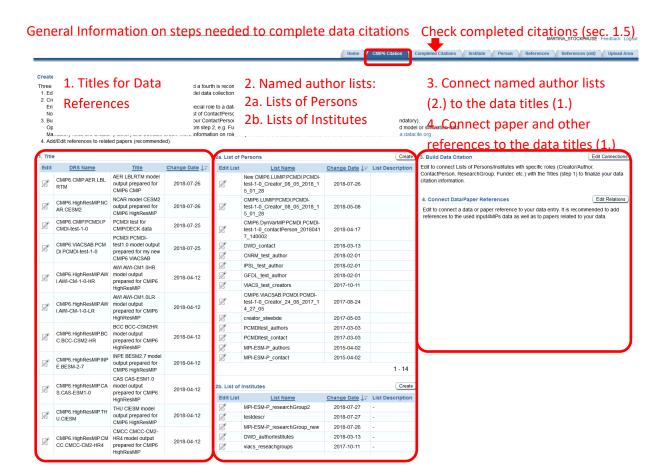
The CMIP6 Citation tab is the main entry point to maintain your data citation information. It is organized as follows:

- **top:** short information on the process steps for providing complete data citation information
- left step 1: area to edit provided default titles for DRS data collection entries on model/MIP granularity.
- center step 2: area to create/delete named author lists and other lists of persons (upper part) and to create/delete lists of institutes, e.g. Funders or research groups (lower part).
  Notes: A copy of an existing list is possible via edit and saving under a new name. List entries are ordered.
- **right step 3:** area to compose citations out of the titles from step 1 and the lists of step 2, e.g. list of authors/creators, list of funders, list of contact persons and list of research groups.

  Note: Mandatory list types to connect to a data citation entry are 'Creator' and 'ContactPerson'.
- **right step 4:** area to add references and connect them to the data, e.g. relevant paper, used data and/or software.

Note: Recommended are references to papers and the used input4MIPs datasets.

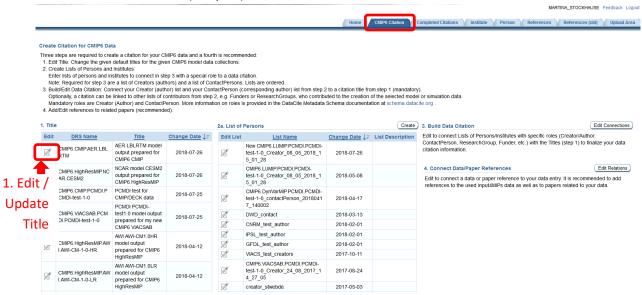
**Note:** From 2021-05-24 on, changes of title (1.) and authors (3.) are allowed again but restricted to entries prior to DOI registration in order to keep data references stable for citation in the IPCC AR6. Adding paper references (4.) is possible for all entries including those with a registered DOI. You can control, which data entries are visible in (1.) and (3.) yourself, using the application under the 'Visibility' tab.



6

# Step 1 - Provide a title for a data citation entry

To edit the titles for your data references click on the 'Edit' symbol next to the data reference in the table in the '1. Title' area (left part).

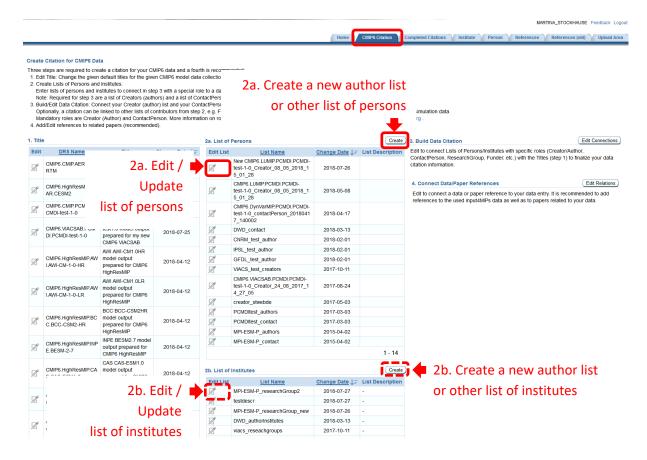


For CMIP6 a CMOR-like default title is provided for each entry. Save your changed title by pressing the 'Apply Changes' button.



## Step 2 - Provide named author/creator and other lists

In the second step, author/creator lists and list of other contributing persons and contributing institutes are prepared.



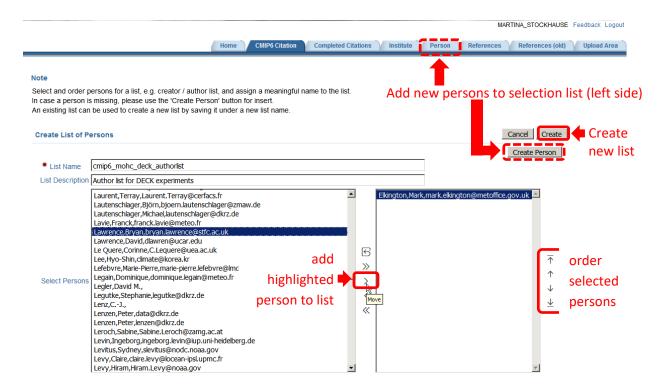
To create a list of persons click on the 'Create' button in the middle part.

You are redirected to a page where you can select persons from a list provided by the database (left side) and add them to the author/Creator, ContactPerson or researcher list (right side) you are creating. A reordering of selected persons in the created list (right side) is possible using the arrow buttons.

Enter a 'List Name', which indicates the use of this list and finalize the list by clicking on the 'Create' button. When you return to the main page using the 'Cancel' button, the name of your entered person list will appear in the table in the 'List of persons' area (upper middle part).

The list of persons can be edited by clicking on the 'Edit' symbol next to it. Lists can be created and deleted but not updated. In case you want to change/update a list, please:

- 1. Edit the existing list of persons and save it under a new 'List Name';
- 2. Replace the old list, e.g. author list, connected to the data by the newly created one (see Step 3 Connect author list and other lists to the data reference for details).

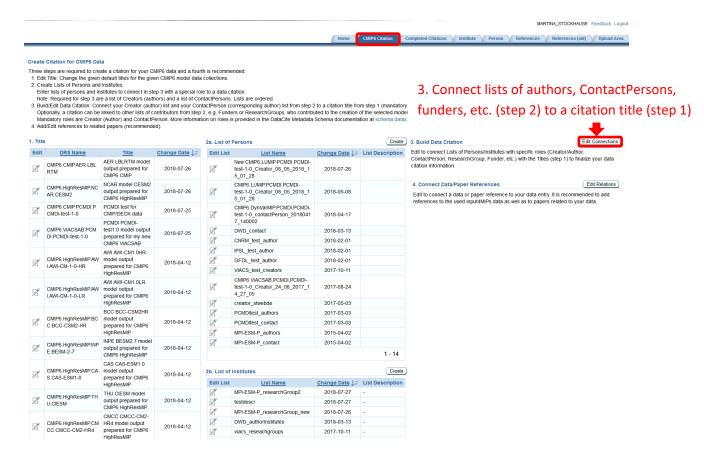


Lists of institutes for Funders, research groups etc. can be created, analogously, in the 'List of Institutes' section located in the lower middle part of the page.

In case a required person (institute) is not available in the selection list on the left side of the person (institute) list dialog, please create the person (institute) first at the 'Person' ('Institute') tab. The button 'Create Person' ('Create Institute') leads you to that tab (see section 0).

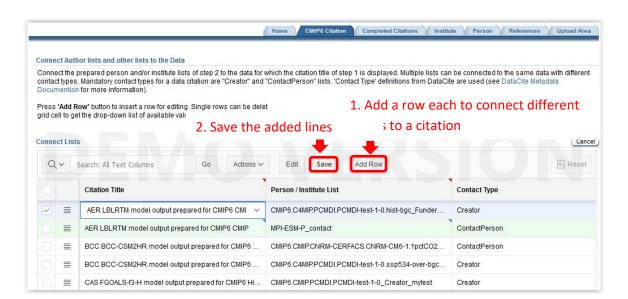
## Step 3 - Connect author list and other lists to the data reference

After providing titles for your data entities in step 1 and lists of authors and other contributors, such as Funder lists, in step 2, these lists are connected to the data entities in step 3. Please note that at least an author list and a ContactPerson list are required for a completed data citation.

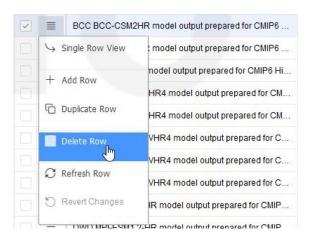


For establishing these connections of lists of persons and institutes to prepared data citation entities please:

- 1. add a row;
- 2. select the data by its title, the list of persons or institutes by its list name and the list's type, e.g. 'Creator'=Author, 'ContactPerson', 'Funder' etc. from the provided drop down menus.



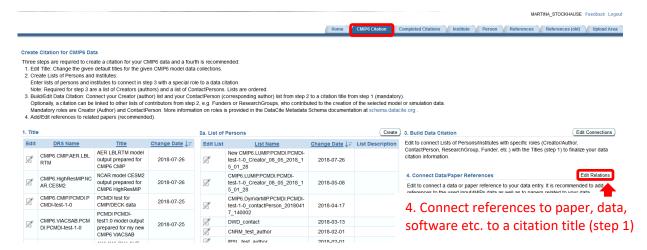
In case you need to update a list, use the drop down menu to change the connected list. For deleting a connection, please use the row menu:



Finalize your changes by pressing the 'Save' button.

## Step 4 - Connect paper or other reference to the data reference (recommended)

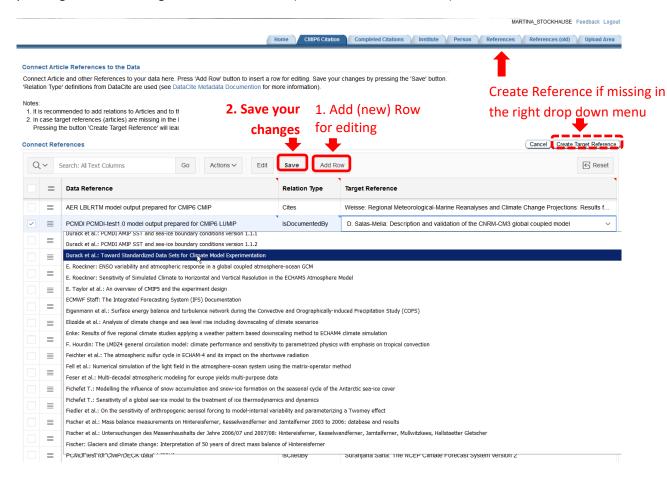
It is recommended to add references to papers or to data (e.g. input4MIPs datasets) and software.



Press the 'Edit References' button to get redirected. Add a row and select data references and the target (paper) reference you want to connect to the data from the given drop down menus. 'Relation Type's from DataCite are used to specify the type of connection (see <a href="schema.datacite.org">schema.datacite.org</a>). Read reference entries from left to right: 'Data Reference' 'Relation Type' 'Target Reference', e.g. 'PCMDI PCMDI-test 1.0 model output for CMIP6 LUMIP' 'isDocumentedBy' 'Durack et al. (2018)'. For deleting a connection, please use the row menu (see Step 3).

### Finalize your changes by pressing the 'Save' button.

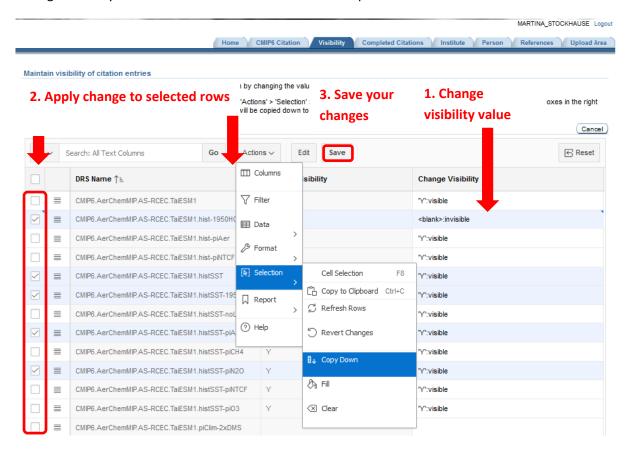
In case a 'Target Reference' (e.g. a paper) is missing in the drop down menu, please create it first by pressing the 'Create Target Reference' button (see section 1.7 for details).



# 1.4 Control Visibility of data citation entries

Use this form to control the visibility of your entries in this application by changing the value in the 'Change Visibility' column and 'Save' the changes.

For changing multiple rows, please use the 'Copy Down' function in 'Actions' > 'Selection' > 'Copy Down' by selecting all rows you want to change using the check boxes in the right most column. The 'Change Visibility' value of the first selected row will be copied down to the other selected rows.



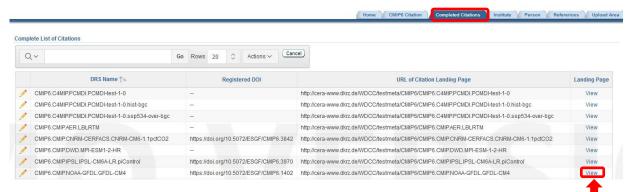
# 1.5 List of Completed Citations and landing pages

After finalizing steps 1-3, your data citations will appear under the 'Completed Citations' tab together with the registered DOI and a link to the citation landing page displaying the full citation information.

Note: This is the only place in the GUI, where you can see the data references for data references

**Note:** This is the only place in the GUI, where you can see the data references for data reference with a registered DOI.

## Check completed citations



Access Landing Page to view full content



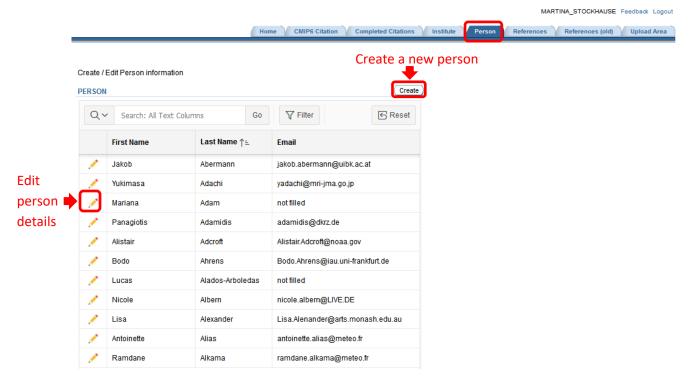
# 1.6 Create a person and institute entries

Creation and update of person and institutes work the same way. Here the example for the creation of a person is described. Go to the 'Person' Tab and

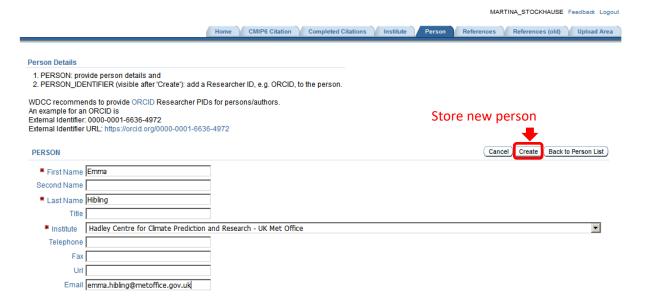
- a. Create a person and
- b. Add ORCID details (recommended)

## a. Create a new person entry

Press the 'Create' button.

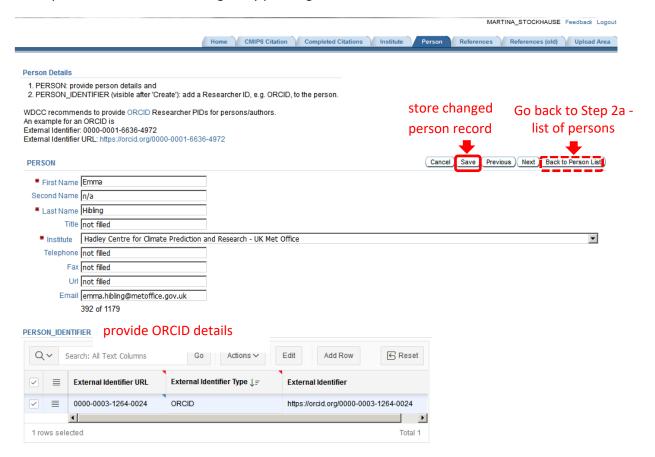


Enter at least the mandatory information in the form: first name, last name, institute and email. Finalize your edits by pressing the 'Create' button.



# b. Add ORCID Researcher ID to person record (recommended)

After saving the person record the person identifier section will appear at the bottom. Add the ORCID to the person record and store it again by pressing the 'Save' button.

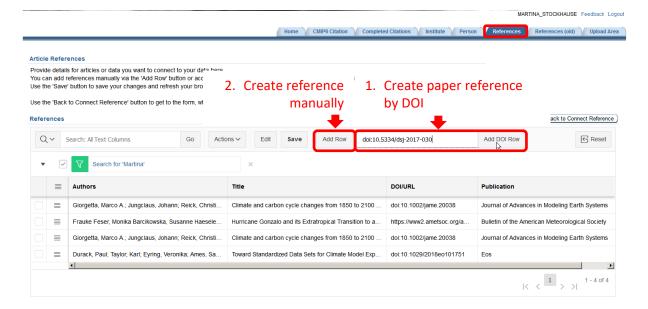


# 1.7 Add Paper and Data References

Under the references tab you have the opportunity to add scholarly publications, which you then can connect in Step 4 to the data. This information will be published together with the data reference and therefore enable data-interlinking services to direct readers from data to articles and vice versa.

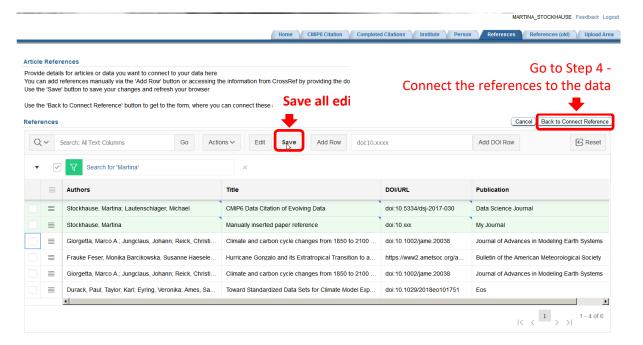
The form provides two possibilities to add a reference:

- 1. 'Add DOI Row' Insert a paper reference by its DOI from CrossRef (a 2<sup>nd</sup> try might be required);
- 2. 'Add Row' Insert a paper/data/software reference manually.



## Finalize your edits by pressing the 'Save' button.

To connect this reference to your data please press the 'Back to Connect Reference' button to be redirected to the right form.

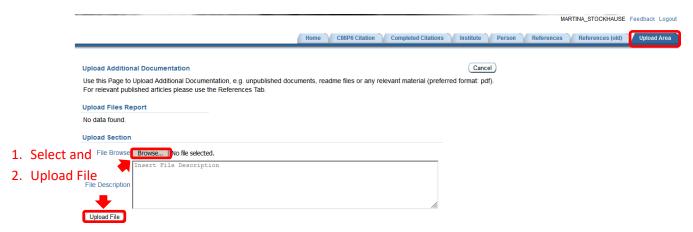


# 1.8 Upload Area

A possible use case for the upload area is:

- Provision of materials which are relevant for the cited CMIP6 data but not formally published in a scholarly publication. These materials will be long-term archived together with the data.

# Please do not upload data files.



# 2. API - Application Programming Interface

Please do not use the API any more as it is not maintained any longer. It's functionality is also restricted to adding paper references (step 4).

The API uses the JSON representation of the DataCite 4.1 schema (<a href="http://doi.org/10.5438/0014">http://doi.org/10.5438/0014</a>) to insert and update citation information.

Please download the python API client from DKRZ's cloud at:

https://swiftbrowser.dkrz.de/public/dkrz 11279a46963f4201bae564a253d528cc/Citation API Client/and check if you can reach the API server with: `ping ceracite.dkrz.de`.

The API client is provided as

- python 3.5+ plus httplib2 (citation\_client.py),
- python 2.7 plus httplib2 (citation\_client.py), and as
- python 2.6+ / python 2.5 plus simplejson (citation\_client\_httplib2.py).

### 2.1 Preconditions

To use the API client please:

- Use the GUI to provide details for all persons, institutes and paper references, which you want to use in the JSON input file for the API insert of citation information (section 1)
- Store your user credentials in the \$HOME/.netrc file:

machine cera
login <user name>
password <password>

## 2.2 Best Practice

1. **Download a template JSON file** from a landing page:

You can find a landing page using the 'View' links on the 'Completed Citation' tab of the GUI (see section 1.5 for details). Use the 'JSON' button to get the JSON representation of your citation information on this example on the coarse citation. To have all optional sections present in the downloaded JSON, *it is best to use a detailed metadata example, i.e. with additional contributors such as funders and paper references as template.* Such a detailed example is given in section 0.

- 2. **Modify the template JSON file** as required, but at least:
  - 2.1. DRS (subject: "CMIP6.<activity\_id>.<institution\_id>.<source\_id>.<experiment\_id>"): This is the information for the citation service, to which data your citation information belongs.

    Note that this subject has to be the first in the list.
  - 2.2. Title

Section 0 gives an overview of the elements and in section 0 you find a shortened JSON example with grey parts indicating ignored and **red** parts indicating mandatory content. Please note that many project sections are ignored, e.g. description, identifier, license, subject, publisher.

- 3. Check your JSON file (see Section 0).
- 4. After a successful check please run the API with your JSON file (Section 0).

# 2.3 Checking tool (without an entry in the database)

# python citation\_client.py <json\_file> check

The checking tool goes through the json file and gives error messages back without any inserts into the database.

## "subjects":

- **Interruption** in the case of missing attribute "subjects" (citation could not be checked without the subject!)
- error messages:

ERROR type	Message/Meaning
NOT FOUND	You have to register your "subject" first at https://github.com/WCRP-CMIP/CMIP6_CVs
AMBIGUOUS	Subject is not unique
NOT ALLOWED TO EDIT	Subject is aligned with another user

### "titles":

• error messages:

ERROR type	Message/Meaning	
IN USE	The title is already in use for another citation ("subject"). The	
	title has to be unique!	
WRONG NUMBER	just one-line title allowed	

# "contacts" (creators, contributors, funders):

- "Creators": one person or institute is mandatory!
- "Contributors": "contributorType" is mandatory!
- error messages:

ERROR type	Message/Meaning
NO CREATOR	No creators provided! A Creator is mandatory
WRONG CONTRIBUTOR	wrong contributor type – a list of valid types is part of the
	response
INFORMATION MISSING	A person must be identified by firstName, givenName, and
	email
AMBIGUOUS DATA	person or institute or personIdentifier or instituteIdentifier is
	ambiguous
UNKNOWN	Person or Institute is not available in the database. Please
	insert via GUI!
COMBINATION NOT ALLOWED	A combination of persons and institutes is not allowed in the
	same list / for the same contributorType (see next Table for
	allowed list member entries per contributorType)
PERSON INSTEAD OF INSTITUTE	Wrong list member! Fill in one or more persons instead of
OR VICE VERSA	institutes / institutes instead of persons (depending on
	contributorType)

• "contributorTypes": Please note that it is not possible to combine persons and institutes in the same contributorType! The allowed list members for the different Types are (in the case of both, person and institute are possible, you have to select one opportunity: persons OR institutes!):

Table 1: List of possible Contributor types

<b>Contributor Type</b>	Person	Institute
Funder	N	Υ
Researcher	Υ	Ν
ResearchGroup	N	Υ
ContactPerson	Υ	Ν
Creator	Υ	Υ
RightsHolder	Υ	Υ
Editor	Υ	N
Other	Υ	Y

### "relatedIdentifiers":

• Paper, model documentation etc.

ERROR type	Message/Meaning
UNKNOWN	new references have to be filled in the data base via GUI
	before using them with the API

## 2.4 Run the API

python citation\_client.py <json\_file>

## 2.5 Check citation information

In case of success the response includes a link to the landing page. Please use it to check your inserted citation information, e.g.:

(200, 'SUCCESS: Your data is visible now at: http://cera-www.dkrz.de/WDCC/meta/CMIP6/CMIP6.ScenarioMIP.PCMDI-test-1-0')

It is also possible to check the citation information via the 'Completed Citations' tab in the GUI (see section 1.5).

# 2.6 Overview of top-level elements (JSON file) for citation

Attribute	Sub-element	Definition	Obligation	used
contributors	contributorType	"ContactPerson" or	mandatory	yes
		"ResearchGroup", (see Table 1)	optional	yes
	givenName	if contributor is a person	mandatory	yes
	familyName		mandatory	yes
	email		mandatory	yes
	affiliation		optional	no
	contributorName	"familyName, givenName(s)" or institute name	mandatory	yes
	nameIdentifier	schemeURI, nameIdentifierScheme, PID	optional	yes
creators	creatorName	Minimum ONE creator is mandatory with	mandatory	no
	givenName	<ul> <li>givenName, familyName and email!</li> </ul>	mandatory	yes
	familyName	or	mandatory	yes
	email	institute name	mandatory	yes
	affiliation		optional	no
	nameIdentifier		optional	yes
dates	dateType	"Created"	automated	no
	date	date as YYYY-MM-DD	automated	no
descriptions	descriptionType	"Abstract"	automated	no
	text	"Coupled Model Intercomparison Project	automated	no
	text	Phase 6 (CMIP6)"		
formats	Format	"application/x-netcdf"	automated	no
fundingReferences	funderName	Institute	optional	yes
	nameIdentifier	with "schemeURI", "nameIdentifierScheme", and "pid"	optional	yes
identifier	identifierType	"URL" or "DOI"	automated	no
	id	URL of landing page or DOI	automated	no
language		"en"	automated	no
publisher		"Earth System Grid Federation"	automated	no
publicationYear		year	automated	no
relatedIdentifiers	relatedIdentifier	recommended input4MIPs and paper	optional	yes
	relatedIdentifierType	references with DOI number	'	,
_	relationType			
resourceType	resourceTypeGeneral	"Dataset"	automated	no
	resourceType	"Digital"	automated	no
rightsList	rightsURI	"http://creativecommons.org/licenses/by-sa/4.0/"	automated	no
	rights	"Creative Commons Attribution 4.0 International License (CC BY-SA 4.0)"	automated	no
subjects	subject	predefined subject as proposed in the CV	mandatory	yes
		(needs to be first subject!)		
	schemeURI	"http://github.com/WCRP-	mandatory	no
		CMIP/CMIP6_CVs"		
	subjectScheme	"DRS"	mandatory	no
	subject	further subjects are not used but have to be	optional	no
		positioned <b>after</b> the DRS subject!		
titles		Mandatory. Just one line! Has to be unique.	mandatory	yes

# 2.7 JSON Example (some unread sections removed)

```
"creators":
     "creatorName": "Taylor, Karl E."
, "givenName": "Karl E."
, "familyName": "Taylor"
        "email": "taylor13@11n1.gov"
       "nameIdentifier":
               "schemeURI": "http://orcid.org/",
"nameIdentifierScheme": "ORCID",
               "pid": "0000-0002-6491-2135"
        "affiliation": "Lawrence Livermore National Laboratory"
        "creatorName": "Lautenschlager, Michael"
     "creatorName": "Lautenschlager, Michael"
, "givenName": "Michael"
, "familyName": "Lautenschlager"
, "email": "lautenschlager@dkrz.de"
, "affiliation": "Deutsches Klimarechenzentrum (DM)"
        "creatorName": "Lawrence, Bryan"
"givenName": "Bryan"
"familyName": "Lawrence"
"email": "bryan.lawrence@stfc.ac.uk"
"titles":
     "PCMDI PCMDI-test1.0 model output prepared for CMIP6 ScenarioMIP (test)"
"publisher": "Earth System Grid Federation",
"publicationYear": "2017",
"subjects":
        "subject": "CMIP6.ScenarioMIP.PCMDI.PCMDI-test-1-0",
        "schemeURI": "http://github.com/WCRP-CMIP/CMIP6_CVs",
"subjectScheme": "DRS"
"contributors":
     "contributorType": "ContactPerson"
, "contributorName": "Lautenschlager, Michael"
, "givenName": "Michael"
, "familyName": "Lautenschlager"
, "email": "lautenschlager@dkrz.de"
        "affiliation": "Deutsches Klimarechenzentrum (DM)"
        "contributorType": "ResearchGroup"
"contributorName": "Institut Pierre Simon Laplace (IPSL)"
        "nameIdentifier":
               "schemeURI": "http://www.isni.org/isni/",
              "nameIdentifierScheme": "ISNI",
"pid": "00000000990008794"
     }
"relatedIdentifiers":
        "relatedIdentifier": "10.1175/JCLI-D-12-00823.1",
        "relatedIdentifierType": "DOI",
        "relationType": "IsCitedBy"
"fundingReferences":
        "funderName": "Federal Ministry of Education and Research (BMBF)"
      , "funderIdentifier": "http://doi.org/10.13039/501100002347", "funderIdentifierType": "Crossref Funder ID"
         "funderName": "Lawrence Livermore National Laboratory (LLNL)"
```

# 2.8 Error codes and messages

Error Message	Description
Traceback (most recent call last): File  "citation_client.py", line 17, in <module> info = netrc.netrc() File  "/usr/lib/python2.7/netrc.py", line 34, ininit with open(file) as fp: IOError:  [Errno 2] No such file or directory:  '/home/user/.netrc'</module>	You must provide a .netrc file in your home directory.  It must contain a line `machine cera login USERNAME password PASSWORD`
Traceback (most recent call last): File  "citation_client.py", line 18, in <module> login, account, password = info.authenticators("cera") TypeError: 'NoneType' object is not iterable</module>	Your `.netrc` file has no line for cera as mentioned above.
400 No creators provided	Your JSON-File has no `creators`-Object which is necessary.
400 No titles defined	Your JSON-File has no `titles`-array which is necessary.
400 No subjects defined	Your JSON-File has no `subjects`-array which is necessary.
400 Persons must have email, givenName and familyName	Each person inside the objects `creators` or `contributors` must have all of the following fields: `email`, `givenName`, and `familyName`
400 Unknown institute:Chocolate Factory	creators, funders or contributors without `email`, `givenName` and `familyName` are considered as institutes. There is no institute `Chocolate Factory` in the database
400 Unknown or ambiguous person: Alfred Neumann with email neumann@llnl.gov	The mentioned person does not exist in the CERA database, or it exists multiple times
400 Could not update title in table citation: No data found	The first subject in `subjects` does not exist in the database. Misspelled?
socket.error: [Errno 111] Connection refused	The client was not able to establish a connection to the citation server. Please contact <a href="mailto:data@dkrz.de">data@dkrz.de</a> if the problem persists
400 Problem with person identifier: ORCID:0000-0002-6491-2135a: Ambiguous data	This person already has an ORCID, but it does not correspond with the provided ORCID in your JSON-File
400 No contributorType provided	At least one of the items in `contributors` does not have a contributorType
400 Problem with related identifier: doi:10.175/JCLI-D-12-00823.1: No data found	Some of the following fields in `relatedIdentifiers` are missing in the database(misspelled?): `relatedIdentifier`, `relatedIdentifierType`, `relationType`
400 Could not attach list with related unit name id: 384 to citationid 1401 with contacttype ContactPerson in list_connect	One of the contact items ('creators' or 'contributors') in the JSON file has an issue. Probably misspelled the 'contributorType'?