CMIP6 Infrastructure Status and Future

Karl E. Taylor and V. Balaji

With contributions from The WGCM Infrastructure Panel

Presented at the WGCM 22nd Session

Barcelona, Spain 25 March 2019

Purpose of CMIP Infrastructure



- Facilitate access to model output produced by MIPs (e.g., CMIP and the CMIP6 endorsed MIPs; obs4MIPs; CORDEX)
- Impose standards that make datasets self-describing
- Define "controlled vocabularies" to enable automated management of data and user-friendly search capabilities
- Provide information and services for
 - Documenting models and experiments
 - Assigning persistent i.d.'s for data citation purposes
 - Reporting errors in data
 - Helping contributors and users of CMIP data to do their work
 - •



Purposes of the WGCM Infrastructure Panel*



- To set requirements ensuring the infrastructure will serve its purpose
- To write documents defining specifications for the infrastructure and data it hosts.
- To help coordinate development work done under independently-funded projects so that the infrastructure elements work well together
- To communicate and coordinate with data managers at modeling groups via a "CMIP Data Node Operations Team" (CDNOT; Sébastien Denvil, chair)

^{*} Y. Bai, V. Balaji, L. Cinquini, S. Denvil, P. Durack, F. Guglielmo, E. Guilyardi, M. Juckes, S. Kharin, M. Lautenschlager, B. Lawrence, M. Mizielinski, K. Taylor



Objectives of this talk

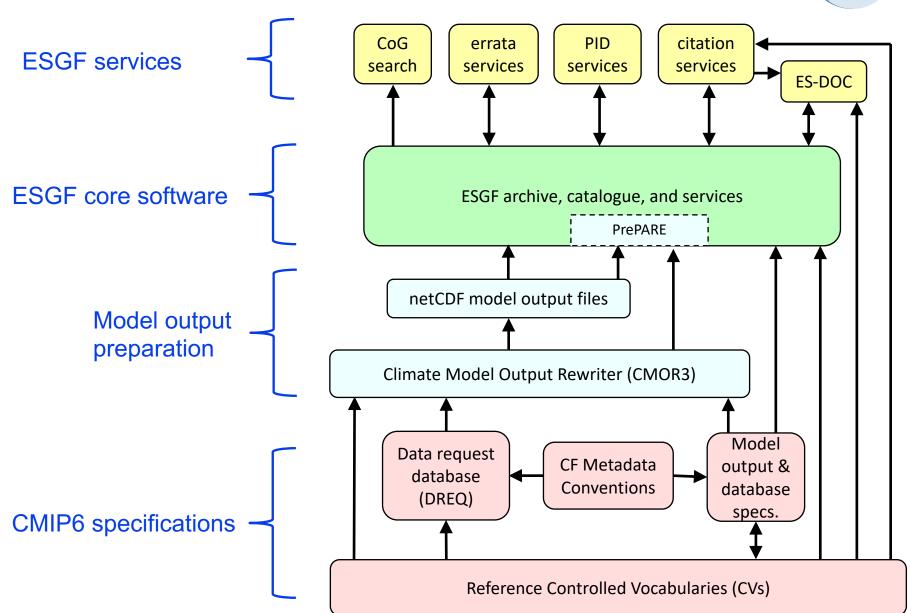


- To provide an overview of the components of the CMIP infrastructure and their status
- To highlight tools and information that may help researchers obtain the data
- Look to the future ... cloud-based opportunities



Infrastructure components and dependencies





Formally registered experiments, institutions, and models are recorded by "controlled vocabularies" (CV's)



CMIP6_CVs

Core Controlled Vocabularies (CVs) for use in CMIP6

Registering Institutions, Models, or requesting changes to CVs:

To register your institution or model or to request changes to a CV, please submit an issue/ticket following the instructions on the CMIP6_CVs issue page.

Some support for CMIP participating modeling groups is available: pcmdi-cmip@llnl.gov

To view the current experiment_id entries point your browser to CMIP6 experiment_id html

To view the current institution_id entries point your browser to CMIP6_institution_id.html

To view the current source_id entries point your browser to CMIP6 source_id html

The CVs build on logic that is described in the CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV's document

Lists of registered:

- experiments
- institutions
- models

All CV's are hosted at:

https://github.com/WCRP-CMIP/CMIP6 CVs

Tabl

Show 5 entries

ble of Registered	Models		
3,5,5,5		Search:	

source_id	institution id	release year	activity participation	cohort	label	label extended	atmos	natNomRes atmos	ocean	natNomRes ocean	lan
ACCESS- CM2	CSIRO- ARCCSS- BoM	2018	MIPS CMIP FAFMIP OMIP RFMIP ScenarioMIP	Registered Mod	ACCESS- CM2	Australian Community Climate and Earth System Simulator Climate Addel Crision 2	MetUM- HadGEM3- GA7.1 (N96; 192 x 144 longitude/latitude; 85 levels; top level 85 km)	250 km	ACCESS-OM2 (GFDL-MOM5, tripolar primarily 1deg; 360 x 300 longitude/latitude; 50 levels; top grid cell 0-10 m)	100 km	none
ACCESS- ESM1-5	CSIRO	2018	C4MIP CDRMIP CMIP OMIP RFMIP ScenarioMIP	Registered	ACCESS- ESM1.5	Australian Community Climate and Earth System Simulator Earth System Model Version 1.5	HadGAM2 (r1.1, N96; 192 x 145 longitude/latitude; 38 levels; top level 39255 m)	250 km	ACCESS-OM2 (MOM5, tripolar primarily 1deg; 360 x 300 longitude/latitude; 50 levels; top grid cell 0-10 m)	100 km	none
ARTS-2-3	UНН	2015	RFMIP	Registered	ARTS 2.3	ARTS 2.3 (Current development version of the Atmospheric Radiative Transfer Simulator)	none	none	none	none	none
AWI- CM-1-1-HR	AWI	2018	CMIP CORDEX HighResMIP OMIP SIMIP VIACSAB	Registered	AWI-CM 1.1 HR	AWI-CM 1.1 HR	ECHAM6.3.04p1 (T127L95 native atmosphere T127 gaussian grid; 384 x 192 longitude/latitude; 95 levels; top level 80 km)	100 km	FESOM 1.4 (unstructured grid in the horizontal with 1306775 wet nodes; 46 levels; top grid cell 0-5 m)	25 km	none
AWI- CM-1-1-LR	AWI	2018	CMIP CORDEX HighResMIP OMIP SIMIP ScenarioMIP VIACSAB	Registered	AWI-CM 1.1 LR	AWI-CM 1.1 LR	ECHAM6.3.04p1 (T63L47 native atmosphere T63 gaussian grid; 192 x 96 longitude/latitude; 47 levels; top level 80 km)	250 km	FESOM 1.4 (unstructured grid in the horizontal with 126859 wet nodes; 46 levels; top grid cell 0-5 m)	50 km	none
source_id	institution id	release year	activity participation	cohort	label	label extended	atmos	natNomRes atmos	ocean	natNomRes ocean	lan
Showing 1 to 5	of 102 entries						Previous	1 2	3 4 5	21 N	ext

Major infrastructure components are operational



- Model output & data base specifications and global metadata requirements
 - Provides the foundation for much of the infrastructure
 https://pcmdi.llnl.gov/CMIP6/Guide/modelers.html#5-model-output-requirements
- CMIP data request tools and documentation (M. Juckes' poster)
 - Specifies variables to save and at what frequency for each experiment
 - Specifies variable-specific metadata to include in output files.

https://www.earthsystemcog.org/projects/wip/CMIP6DataRequest

CMOR can be used to write model output in conformance with CMIP requirements

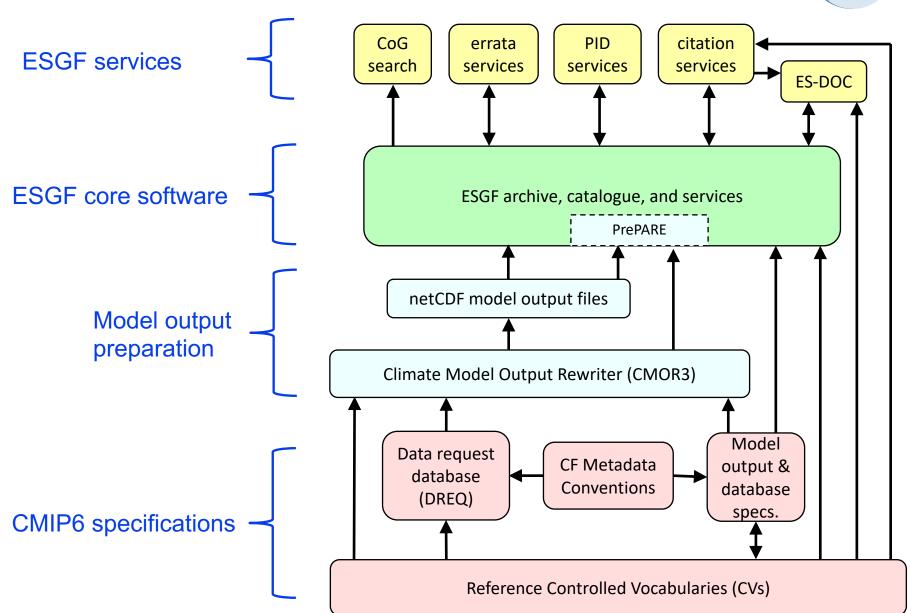
https://github.com/PCMDI/cmor and http://cmor.llnl.gov/

- Pre-Publication Attribute Reviewer for ESGF (PrePARE)
 - new module checking file metadata before publication



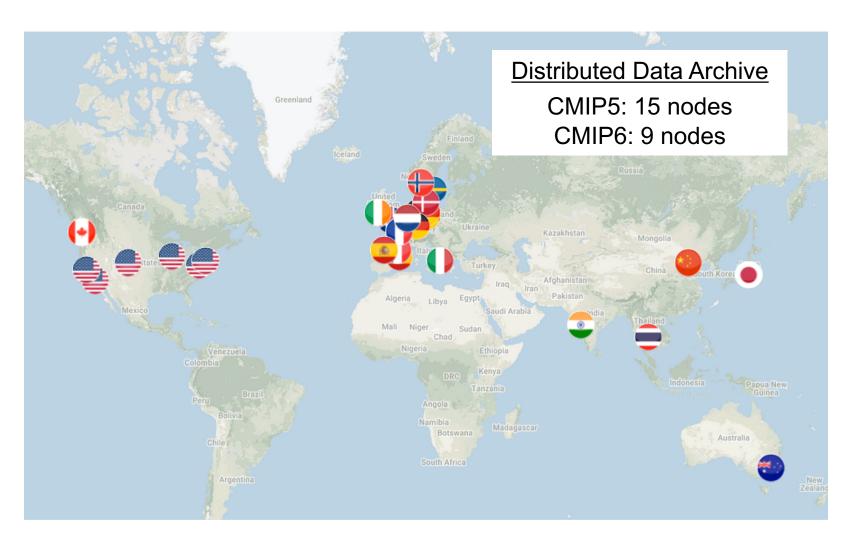
Infrastructure components and dependencies





The Earth System Grid Federation (ESGF) develops and manages CMIP and other data

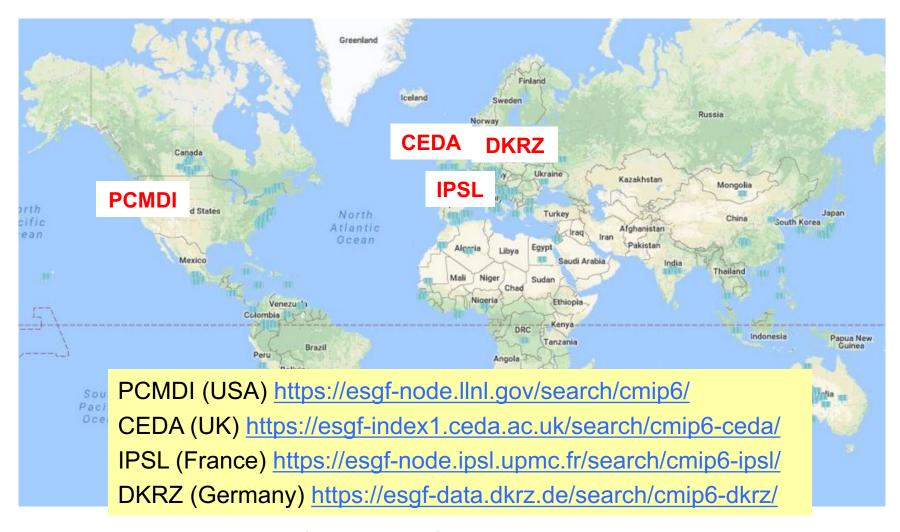






All CMIP6 data can be accessed from each of the 4 CMIP6 portals







ESGF status for CMIP6



- Summary of data holdings available
 https://pcmdi.llnl.gov/CMIP6/ArchiveStatistics/esgf_data_holdings/
- More detailed info. available from portal search window
 - https://esgf-node.llnl.gov/search/cmip6/
 - 6 institutions; nearly 100 experiments
- Replication is ongoing at PCMDI and DKRZ
- To do (high priority): expand availability of Globus grid ftp to all sites.
- To do (high priority): provide server side computational services



ESGF services: CoG

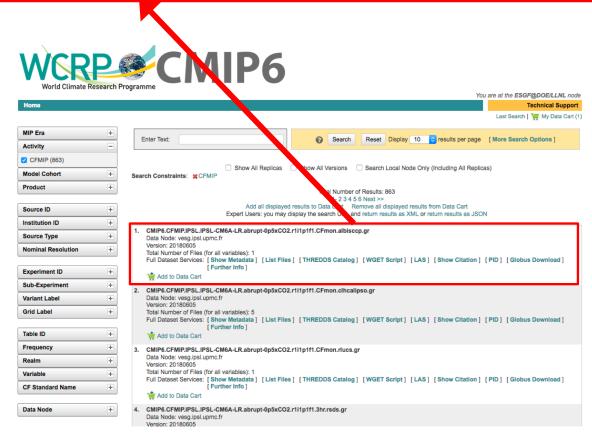


- Provides users with an interface to browse and download all available data
- All 4 gateways have identical look and feel



Data citation services





Citation page







Metadata for 'CMIP6.CFMIP.IPSL.IPSL-CM6A-LR.abrupt-0p5xCO2'

General Information

General Information

Name Abstract

CMIP6.CFMIP.IPSL.IPSL-CM6A-LR.abrupt-0p5xCO2

Coupled Model Intercomparison Project Phase 6 (CMIP6) data sets.

These data includes all datasets published for 'CMIP6.CFMIP.IPSL.IPSL-CM6A-LR.abrupt-0p5xCO2' according to the Data Reference Syntax defined as

'mip_era.activity_id.institution_id.source_id.experiment_id.member_id.table_id.variable_id.grid_label.version'.

The Earth System Model IPSL-CM6A-LR, released in 2017, includes the components:

atmos: LMDZ (NPv6, N96; 144 x 143 longitude/latitude; 79 levels; top level 40000 m), land: ORCHIDEE (v2.0, Water/Carbon/Energy mode), ocean: NEMO-OPA (eORCA1.3, tripolar primarily 1deg; 362 x 332 longitude/latitude; 75 levels; top grid cell 0-2 m), ocnBgchem: NEMO-PISCES, seaIce: NEMO-LIM3.

The model was run by the Institut Pierre Simon Laplace, Paris 75252, France (IPSL) in native nominal resolutions: atmos: 250 km, land: 250 km, ocean: 100 km, ocnBgchem: 100 km, seaIce: 100 km.

Project: These data have been generated as part of the internationally-coordinated Coupled Model Intercomparison

Project Phase 6 (CMIDE: see also CMD Special Issue) http://www.coosei.model.dov.pot/cposial_issue500.html). The

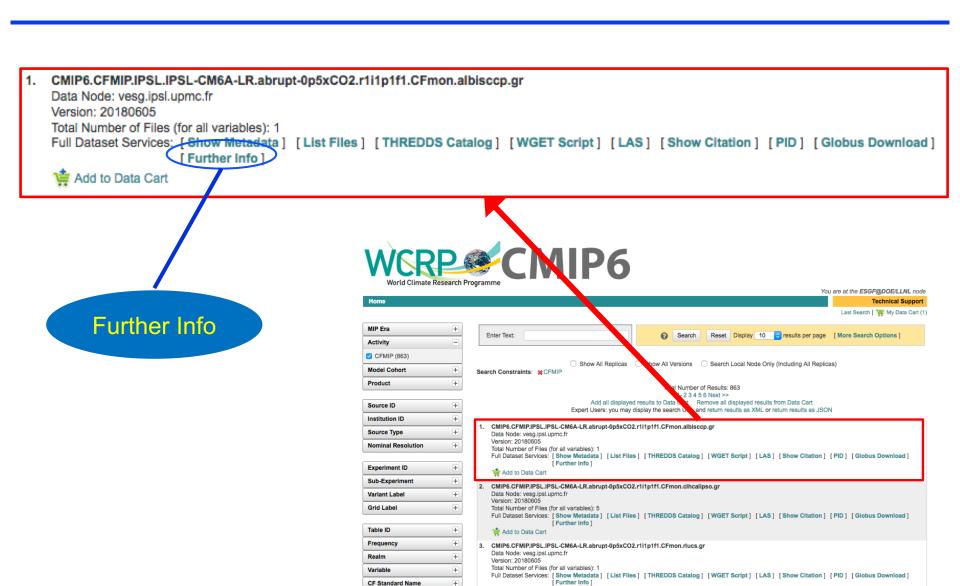
Cite this data

Citation

(2018). IPSL IPSL-CM6A-LR model output prepared for CMIP6 CFMIP abrupt-0p5xCO2. Earth System Grid Federation. http://cera-www.dkrz.de/WDCC/meta/CMIP6/CMIP6.CFMIP.IPSL.IPSL-CM6A-LR.abrupt-0p5xCO2

Model and experiment documentation

Data Node



Add to Data Cart

Data Node: vesg.ipsl.upmc.fr Version: 20180605

CMIP6.CFMIP.IPSL.IPSL-CM6A-LR.abrupt-0p5xCO2.r1l1p1f1.3hr.rsds.gr

further_info_url





CMIP6 Further Information v0.5.1.0



Help

Further Info URL: https://furtherinfo.es-doc.org/cmip6.ipsl.ipsl-cm6a-lr.dcppa-hindcast-niff.s2000.r1

ES-DOC Documentation

MIP Era CMIP6

Institution IPSL

Consortia IPSL

Model IPSL-CM6A-LR

Experiment dcppA-hindcast-niff

Ensemble Description N/A

Machine Performance N/A

Dataset Documentation

Dataset ESGF Search N/A

Dataset Errata N/A

Dataset Citation(s) https://cera-www.dkrz.de/WDCC/meta/CMIP6/CMIP6.DCPP.IPSL.IPSL-CM6A-

LR.dcppA-hindcast-niff

Other Documentation

WCRP CMIP6 Homepage https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6

ES-DOC CMIP6 https://es-doc.org/cmip6

Homepage



further info url





CMIP6 Further Information v0.5.1.0

Support

Help

Further Info URL: https://furtherinfo.es-doc.org/cmip6.ipsl.ipsl-cm6a-lr.dcppa-hindcast-niff.s2000.r1

ES-DOC Documentation

MIP Era CMIP6

Institution IPSL

Consortia IPSL

Model IPSL-CM6A-LR

Experiment dcppA-hindcast-niff

Ensemble Description N/A

Machine Performance N/A

This web page already exists for every ensemble that will be run by every group

Links to ES-DOC documentation, as it becomes available

Dataset Documentation

Dataset ESGF Search N/A

Dataset Errata N/A

Dataset Citation(s) https://cera-www.dkrz.de/WDCC/meta/CMIP6/CMIP6.DCPP.IPSL.IPSL-CM6A-

LR.dcppA-hindcast-niff

Other Documentation

WCRP CMIP6 Homepage https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6

ES-DOC CMIP6 https://es-doc.org/cmip6

Homepage



ES-DOC status (March 2019)



Documents:

- 22 MIPs and 289 experiments fully documented
- 13 groups working on documenting models
- Landing pages for ~200,000 ESGF datasets on the ES-DOC server

Group engagement :

- 37 out of 45 groups (82%) have named an "ES-DOC liaison"
- 8 out of 45 (18%) have not: AS-RCEC, CAMS, CSIRO, DKRZ, E3SM-Project, INM, NERC, NIMS-KMA (requests sent)

Groups who have published at least one model realm :

- 4 out of 45 (9%): IPSL, MIROC, MOHC, NCAR
- These will appear in AR6 FOD model table, the rest in SOD

Groups who have started work on model documentation :

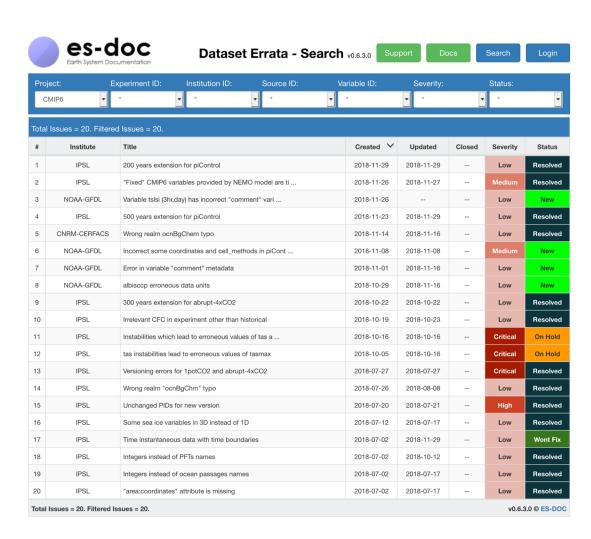
 13 out of 45 (29%): CCCR-IITM, CNRM-CERFACS, ECMWF, HAMMOZ-CONSORTIUM, INPE, IPSL, MESSY-CONSORTIUM, MIROC, MOHC, MPI-M, MRI, NCAR, NOAA-GFDL

CMIP6 documentation on track - nudge from WGCM helps Good feedback on process by groups (beta testing is key) Feeding in IPCC process and 3 papers in preparations



CMIP6 data errata is provided by ES-DOC

- Entry page https://search.es-doc.org/
- Any simulation/data problems must be reported using the ES-DOC errata service
- This will trigger a response (and resolution) by the contributing modelling group



https://errata.es-doc.org/static/index.html

ESGF provides access to CMIP6 forcing datasets and some observations





CMIP6 forcing datasets (P. Durack's Poster)

https://esgf-node.llnl.gov/search/input4mips/



 Observations conforming to CMIP6 data specifications (P. Gleckler and D. Waliser)

https://esgf-node.llnl.gov/projects/obs4mips/



CMIP6 practical guides for contributors and users



CMIP6 Guide: https://pcmdi.llnl.gov/CMIP6/

Program for Climate Model Diagnosis & Intercomparison

Google Custom Sea



Home About[▼]

Research*

CMIP6 MIPs▼

Publications*

Software*

CMIP Data (ESGF Portal)

CMIP6 - Coupled Model Intercomparison Project Phase 6

Overview:

The WCRP Working Group on Coupled Modelling (WGCM) oversees the Coupled Model Intercomparison Project, which is now in its 6th phase. Background information about CMIP and its phases can be found on WGCM website as well as on the PCMDI-hosted pages. An introductory overview of CMIP6 is also provided by the WGCM.

Practical information for those interested in participating in CMIP6 is provided in three guides, tailored to different groups:

- Modelers carrying out CMIP6 simulations,
- 2. Data managers responsible for data node operations, and
- 3. Data users analyzing and making use of CMIP6 model output

CMIP6 Endorsed MIPs:

- WCRP Endorsed (Model Intercomparison Project) MIPs overview page
- CMIP6 Ocean Model Intercomparison Project (OMIP) overview page

Additional information for CMIP6:

· CMIP6 license and terms of use



Lawrence Livermore National Laboratory 7000 East Avenue • Livermore, CA 94550 Operated by Lawrence Livermore National Security, LLC, for the Department of Energy's National Nuclear Security Administration.

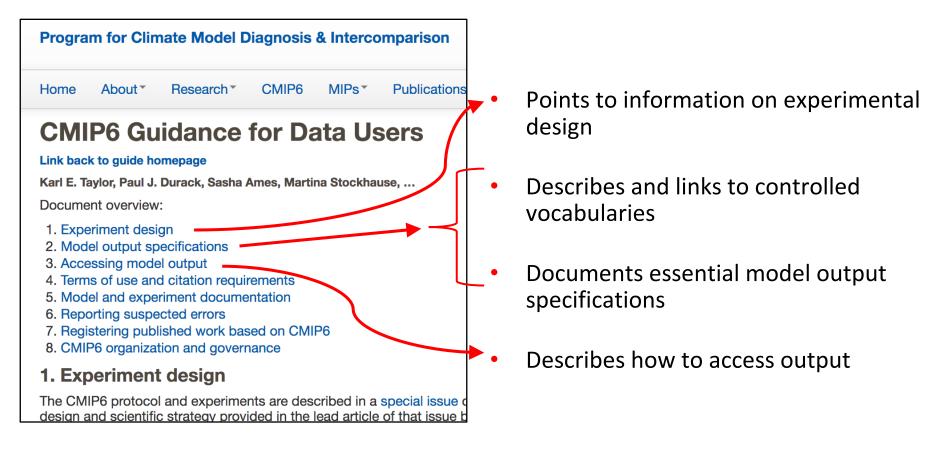






Example: Guidance document for data users





https://pcmdi.llnl.gov/CMIP6/Guide/dataUsers.html



Please record you CMIP publications.



https://cmip-publications.llnl.gov/view/CMIP6/

