

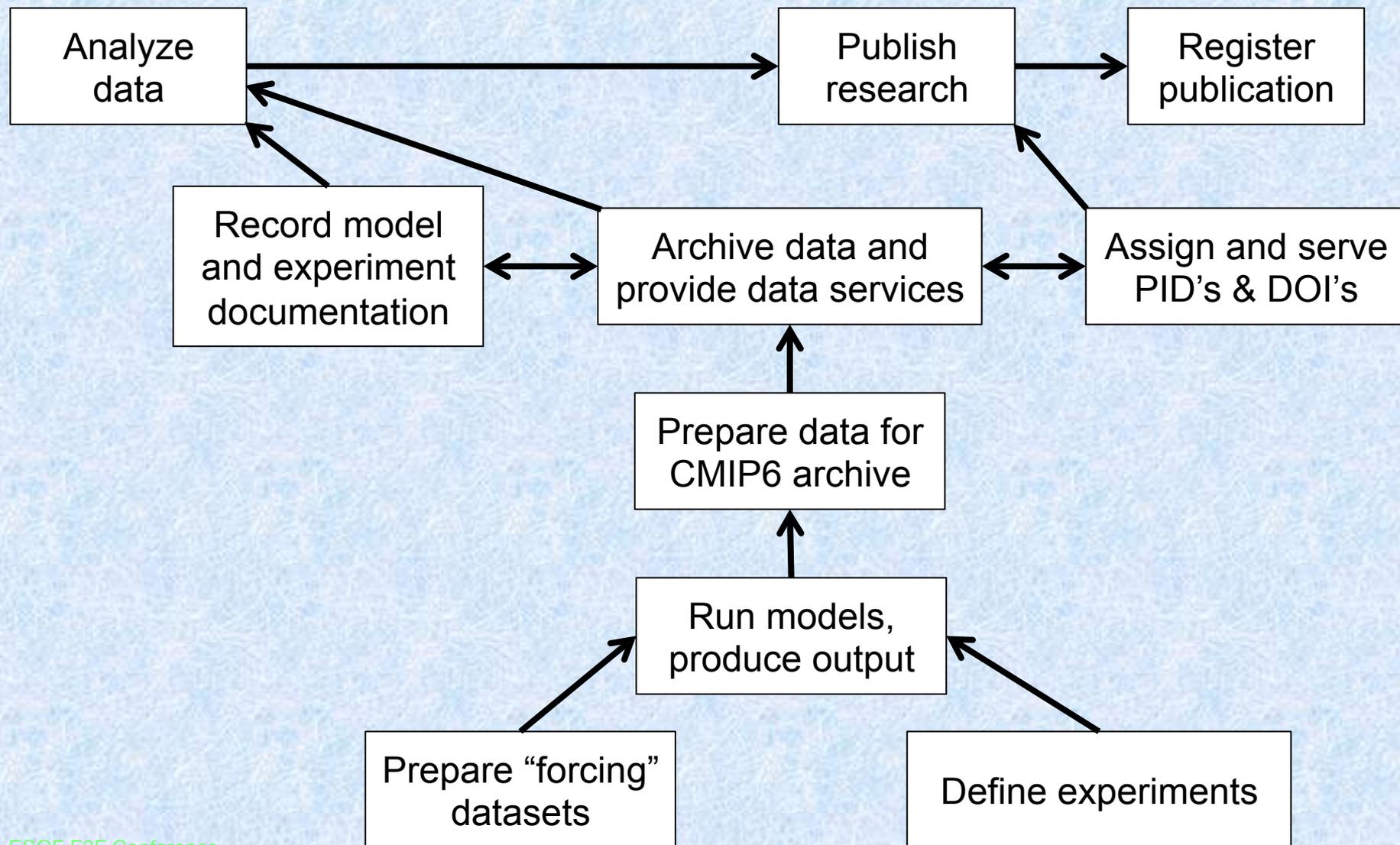
# CMIP6 Standards Enabling Management, Search, and Interpretation of Model Output

Karl E. Taylor, Paul Durack, Denis Nadeau, and  
Sasha Ames

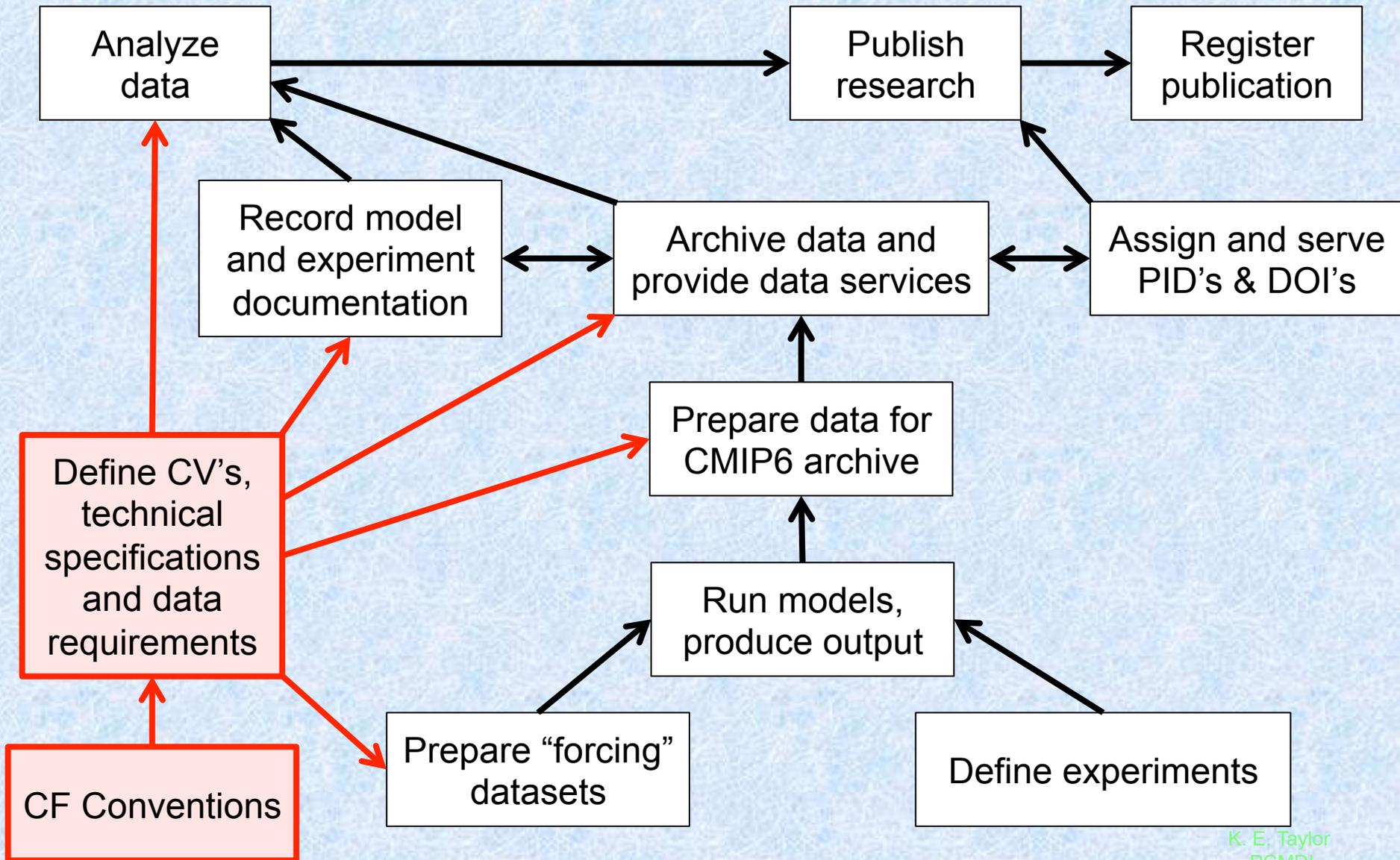
Presented at the 2016 Earth System Grid Federation  
(ESGF) Face-to-Face Conference

Washington D.C.  
5-9 December 2016

# Major components of CMIP6 modeling and data archive infrastructure must communicate and be coordinated

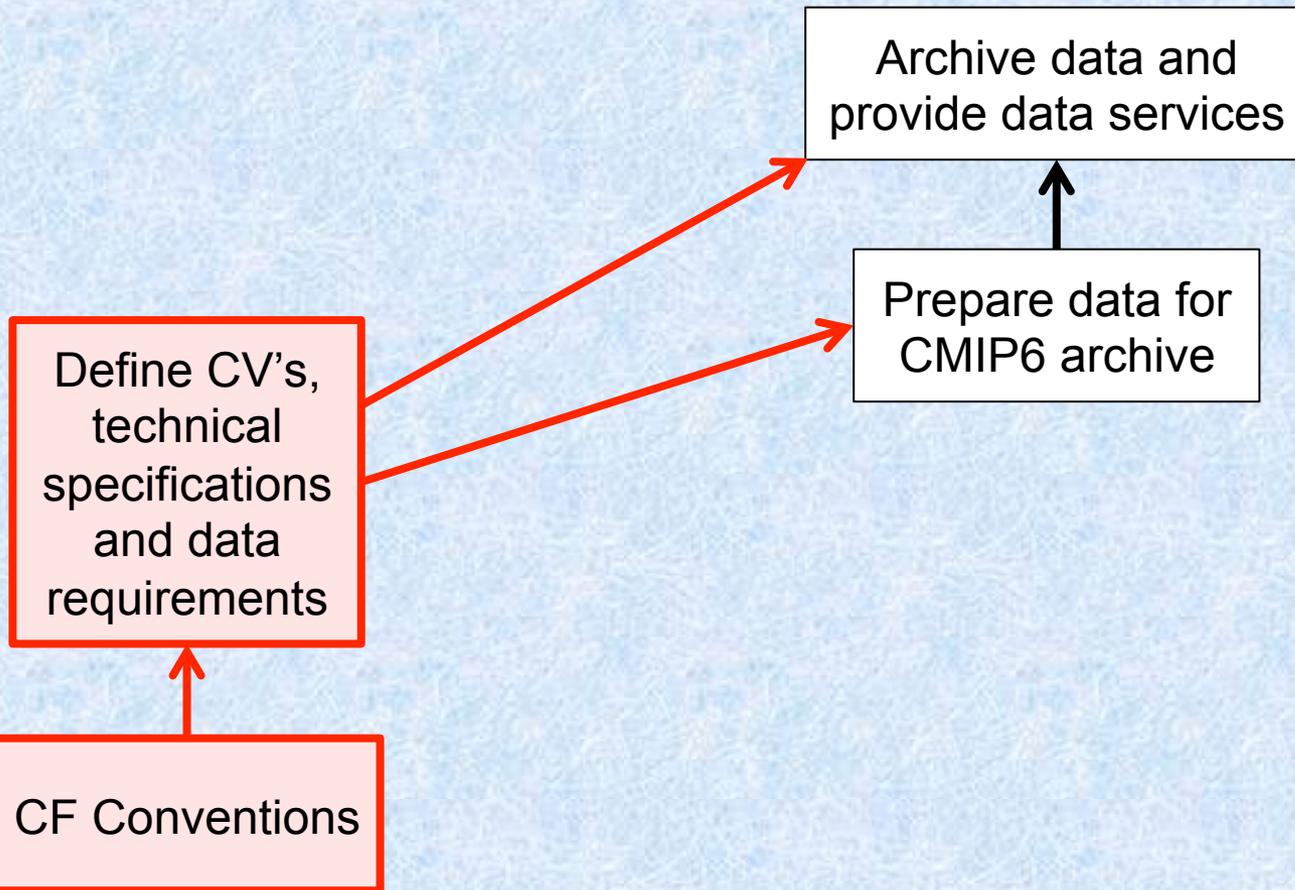


# Standardized output, metadata and CVs are relied on in managing and interpreting CMIP6 results



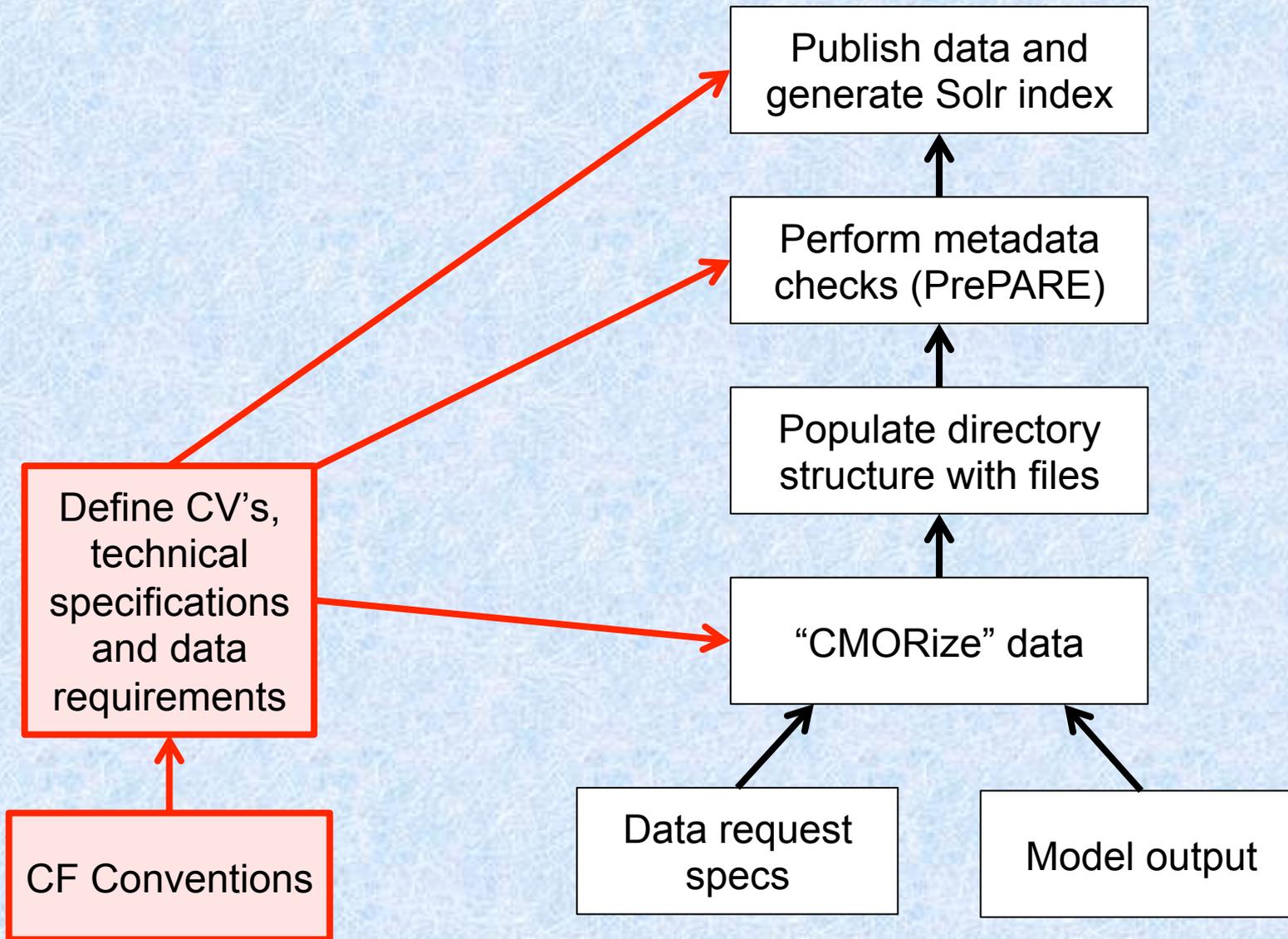
# Data preparation and publication steps.

---



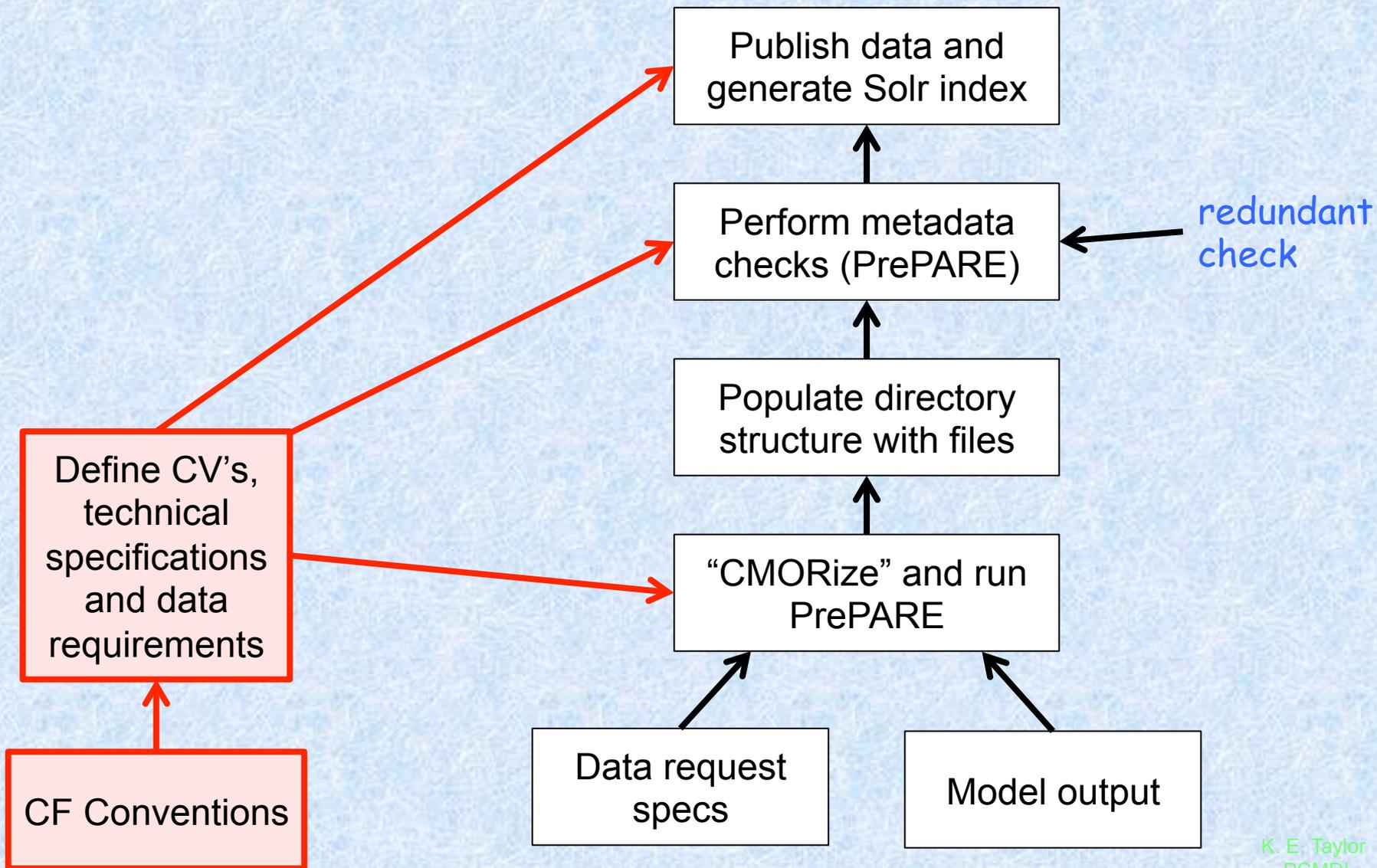
# Expanded view of data preparation and publication steps

---



Recommend that PrePARE be run at time of file creation, so errors can be corrected immediately

---



# Define global attributes that are used in identifying datasets, constructing filenames, and defining search facet

---

<u>DRS Element</u>	<u>Examples</u>	<u>Controlled Vocabulary?</u>
activity_id	PMIP, CFMIP, ScenarioMIP	CV
product	=output	only 1 option
institution_id	IPSL, CCCma	CV
source_id	EC-Earth-3-LR, NorESM2-LM	CV
experiment_id	piControl, historical, 1pctCO2	CV
variable_id	tas, pr, hur	data request
table_id	Amon, 3hr, Oday	CV
variant_label	r2i1p1f1, r1i1p1f2	template
version	v20160218, v20170821	template
sub_experiment_id	1980, 1981, 2001	CV
grid_label	gn, gr, gr1, gr2	CV
mip_era	=CMIP6	CV
frequency	mon, day, 6hr	CV
realm	atmos, ocean, land, seaice	CV

## 46 Global attributes are defined in a table (with notes)

The attributes provide critical information needed to interpret the model output and are key attributes are relied on by the infrastructure.

CMIP6 global attribute see note 1	description	examples	corresponding attribute in CMIP5	form see note 2	when required?	further information and rationale
<b>activity_id</b>	activity identifier(s)	“CMIP”, “PMIP”, “LS3MIP LUMIP” see note 3	project_id	CV	always	renamed more generically, since not all activities are projects; also multiple activities may now be listed separated by single spaces.
branch_method	branching procedure	“standard”, “none provided”, “no parent” see note 4	-	free form	whenever parent exists	in CMIP6 some branching methods will involve short spin-up periods or other non-standard procedures which need to be described. See note 4. If no parent, omit or set to “no parent”
branch_time_in_child	branch time with respect to child’s time axis	365.0D0, 0.0D0 see note 5	-	double precision float	whenever parent exists	aids in interpreting branch times; units are the same as the units used for the child’s time axis. If no parent, omit (preferred) or set to start time of the run.
<b>branch_time_in_parent</b>	branch time with respect to parent time axis	3650.0D0 see note 5	branch_time	double precision float	whenever parent exists	changed name to explicitly distinguish it from branch_time_in_child; units are specified in the attribute: parent_time_units. If no parent, omit or set to 0.0D0.

# ESGF search facets are largely based on DRS elements governed by CVs.

---

- DRS elements

▶▶▶ activity_id	experiment_id
▶▶▶ subexperiment_id	source_id
▶▶▶ run_variant_id	frequency
▶▶▶ realm	table_id
▶▶▶ variable_id	grid_label
▶▶▶ institution_id	product

- Other

▶▶▶ Model cohort	Source_type
▶▶▶ CF standard_name	long_name
▶▶▶ grid_resolution	Data_node

See CMIP6 search requirements:  
<https://goo.gl/rAvXIB>

# Controlled vocabularies are used to define filenames and directory structure and to reach documentation

---

- The CVs are defined by JSON files on github.
- Filenames template:
  - `<variable_id>_<table_id>_<source_id>_<experiment_id>_<member_id>_<grid_label>[_<time_range>].nc`
- Directory structure template:
  - `<mip_era>/<activity_id>/<institution_id>/<source_id>/<experiment_id>/<member_id>/<table_id>/<variable_id>/<grid_label>/<version>`
- ES-DOC documentation reachable via the `further_info_url` global attribute constructed from a template:
  - `http://furtherinfo.es-doc.org/<mip_era>.<institution_id>.<source_id>.<experiment_id>.<sub_experiment_id>.<variant_label>`

# JSON files are hosted by github

[https://github.com/WCRP-CMIP/CMIP6\\_CVs](https://github.com/WCRP-CMIP/CMIP6_CVs)

Issues

WCRP-CMIP / CMIP6\_CVs

Unwatch 6 Star 0 Fork 3

Code Issues 14 Pull requests 0 Projects 0 Wiki Pulse Graphs Settings

Controlled Vocabularies (CVs) for use in CMIP6 — Edit

828 commits 1 branch 1 release 5 contributors

Branch: master New pull request Create new file Upload files Find file Clone or download

durack1 committed on GitHub Issue156 durack1 revise source\_id NorESM various (#167) Latest commit b6f52dd 6 days ago

.github	Source_id format reorder	a month ago
src	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
.gitignore	Further formatting - deal with xlsx quirks	5 months ago
CMIP6_activity_id.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_experiment_id.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_frequency.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_grid_label.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_institution_id.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_license.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_nominal_resolution.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_realm.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_required_global_attributes.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_source_id.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_source_type.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
CMIP6_table_id.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
README.md	Added source_id html	15 days ago
mip_era.json	Issue156 durack1 revise source_id NorESM various (#167)	6 days ago
README.md		

CMIP6\_activity\_id.json

CMIP6\_institution\_id.json

# Some JSON file contents are viewable through web browsers

experiment\_id.json rendering:

http://rawgit.com/WCRP-CMIP/CMIP6\_CVs/master/src/CMIP6\_experiment\_id.html

experiment\_id activity\_id parent\_experiment\_id model components description

experiment_id	activity_id	experiment	tier	sub_experiment_id	sub_experiment	parent_experiment_id	required model components	additional allowed model components	start year	end year	min number yrs per sim	parent activity id	description
1pctCO2	CMIP	1 percent per year increase in CO2	1	none	none	piControl	AOGCM	AER CHEM BGC			150	CMIP	DECK: 1pctCO2
1pctCO2-bgc	CMIP	biogeochemically-coupled version of 1 percent per year increasing CO2 experiment	1	none	none	piControl	AOGCM BGC	AER CHEM			150	CMIP	Biogeochemically-coupled specified concentration simulation in which CO2 increases at a rate of 1% per year until quadrupling
1pctCO2-rad	CMIP	radiatively-coupled version of 1 percent per year increasing CO2 experiment	2	none	none	piControl	AOGCM BGC	AER CHEM			150	CMIP	Radiatively-coupled specified concentration simulation in which CO2 increases at a rate of 1% per year until quadrupling
1pctCO2ndep	CMIP	1 percent per year increasing CO2 experiment with increasing N-deposition	2	none	none	piControl	AOGCM BGC	AER CHEM			150	CMIP	Fully-coupled specified concentration simulation in which CO2 increases at a rate of 1% per year until quadrupling, plus an additional scenario of anthropogenic nitrogen deposition
1pctCO2ndep-bgc	CMIP	biogeochemically-coupled version of 1 percent per year increasing CO2 experiment with increasing N-deposition	2	none	none	piControl	AOGCM BGC	AER CHEM			150	CMIP	Biogeochemically-coupled specified concentration simulation in which CO2 increases at a rate of 1% per year until quadrupling, plus an additional scenario of anthropogenic nitrogen deposition
1pctCO2o4s-withism	ISMIP6	simulation with interactive ice sheet forced by 1 percent per year increase in CO2 to 4xCO2 (subsequently held fixed)	1	none	none	piControl-withism	AOGCM ISM	AER CHEM BGC			350	ISMIP6	Idealized 1%/yr CO2 increase to 4xCO2 over 140yrs and kept constant at 4xCO2 for an additional 200 to 400 yrs simulation that includes interactive ice sheets
a4SST	CFMIP	as piSST but with SSTs from abrupt4xCO2	2	none	none	AGCM	AER CHEM				30		As piSST, but with monthly-varying SSTs taken from years 111-140 of each model's own abrupt4xCO2 experiment instead of from piControl. Sea-ice is unchanged from piSST
a4SSTice	CFMIP	as piSST but with SSTs and sea ice from abrupt4xCO2	2	none	none	AGCM	AER CHEM				30		As piSST, but with monthly-varying SSTs and sea-ice taken from years 111-140 of each model's own abrupt4xCO2 experiment instead of from piControl
a4SSTice-4xCO2	CFMIP	as piSST but with SSTs and sea ice from abrupt4xCO2, and 4xCO2 seen by radiation and vegetation	2	none	none	AGCM	AER CHEM				30		As a4SSTice, but CO2 is quadrupled, and the increase in CO2 is seen by both the radiation scheme and vegetation
abrupt-0p5xCO2	CFMIP	abrupt halving of CO2	2	none	none	piControl	AOGCM	AER CHEM BGC			150	CMIP	Identical to the DECK abrupt-4xCO2, but at 0.5xCO2
abrupt-2xCO2	CFMIP	abrupt doubling of CO2	2	none	none	piControl	AOGCM	AER CHEM BGC			150	CMIP	Identical to the DECK abrupt-4xCO2, but at 2xCO2
abrupt-4xCO2	CMIP	abrupt quadrupling of CO2	1	none	none	piControl	AOGCM	AER CHEM BGC			150	CMIP	DECK: abrupt4xCO2
abrupt-solm4p	CFMIP	abrupt 4% decrease in solar constant	2	none	none	piControl	AGCM	AER CHEM			150	CMIP	Conceptually similar to abrupt 4xCO2 DECK experiment, except that the solar constant rather than CO2 is abruptly reduced by 4%
abrupt-solp4p	CFMIP	abrupt 4% increase in solar constant	2	none	none	piControl	AGCM	AER CHEM			150	CMIP	Conceptually similar to abrupt 4xCO2 DECK experiment, except that the solar constant rather than CO2 is abruptly increased by 4%
amip	CMIP	AMIP	1	none	none	AGCM	AER CHEM		1979	2014	36		DECK: AMIP
amip-4xCO2	CFMIP	AMIP SSTs with 4xCO2	1	none	none	AGCM	AER CHEM		1979	2014	36		As CMIP5/CFMIP-2 amip4xCO2 experiment. AMIP experiment where SSTs are held at control values and the CO2 seen by the radiation scheme is quadrupled
amip-a4SST-4xCO2	CFMIP	as AMIP but with warming pattern from abrupt4xCO2 added to SSTs and 4xCO2 seen by radiation and vegetation	2	none	none	AGCM	AER CHEM				36		Same as amip, but a patterned SST anomaly is applied on top of the monthly-varying amip SSTs. This anomaly is a monthly climatology, taken from each model's own abrupt4xCO2 run minus piControl (using the mean of years 111-140 of abrupt4xCO2, and the parallel 30-year section of piControl). CO2 is quadrupled, and the increase in CO2 is seen by both the radiation scheme and vegetation
amip-future4K	CFMIP	AMIP with patterned 4K SST increase	1	none	none	AGCM	AER CHEM		1979	2014	36		As CMIP5/CFMIP-2 amip/future experiment. AMIP experiment where SSTs are subject to a composite SST warming pattern derived from coupled models, scaled to an ice-free ocean mean of 4K
amip-hist	GMMIP	AMIP-style simulation covering the period 1870-2014	1	none	none	AGCM	AER CHEM		1870	2014	144		Extended AMIP run that covers 1870-2014. All natural and anthropogenic historical forcings as used in CMIP5 Historical Simulation will be included. AGCM resolution as CMIP6 Historical Simulation. The HadISST data will be used
amip-hid	GMMIP	same as "amip" run, but surface elevations of the East African Highlands in Africa, Sierra Madre in N. America and Andes in S. America reduced to 500m	3	none	none	AGCM	AER CHEM		1979	2014	35		The topography of the highlands in Africa, N. America and S. America TP is modified by setting surface elevations to a certain height (500m). Same model as DECK
amip-lfmp-pdLC	LS3MIP	prescribed land (from current climatology) and AMIP SSTs	2	none	none	AGCM	AER CHEM		1980	2100	121		Prescribed land conditions 1980-2014 climate; AMIP SSTs
amip-lfmp-pObs	LS3MIP	prescribed land (from pseudo-observations) and AMIP SSTs	2	none	none	AGCM	AER CHEM		1980	2100	121		Land-hist land conditions; AMIP SSTs
amip-lfmp-rmLC	LS3MIP	prescribed land conditions (from running mean climatology) and AMIP SSTs	2	none	none	AGCM	AER CHEM		1980	2100	121		Prescribed land conditions 30yr running mean; AMIP SSTs
amip-lwoff	CFMIP	AMIP experiment with longwave cloud-radiative effects off	2	none	none	AGCM	AER CHEM		1979	2014	36		As amip experiment, but with cloud-radiative effects switched off in the LW radiation code
amip-m4K	CFMIP	AMIP with uniform 4K SST decrease	2	none	none	AGCM	AER CHEM		1979	2014	36		As amip experiment but SSTs are subject to a uniform cooling of 4K

# CMIP data request software and requirements

---

- Through an API, you can determine what variables to save by specifying
  - An experiment
  - A year of the simulation
  - The experiment suite planned for your model
- Metadata associated with each variable are retrievable:
  - e.g., `standard_name`, `units`, `cell_methods`
  - CMOR tables are generated based on the metadata recorded by the data request

# CMIP data request tools and documentation

---

- Primary source found at the WIP CoG site:

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

## CMIP6 Data Request

The CMIP6 experimental design and organization has been agreed at the WGCM 18th Session in October 2014, see details on the CMIP Panel website at <http://www.wcrp-climate.org/index.php/wgcm-cmip/about-cmip>. Part of this covers the creation and timeline of the *CMIP6 Data Request*.

The data request is available through a repository, and the latest version is available here (updated October 21st, 2016):

<http://proj.badc.rl.ac.uk/svn/exarch/CMIP6dreq/tags/latest>

An overview of the pressure levels proposed for atmospheric diagnostics is [available for discussion \(here\)](#).

Key documents describing the request (in the "docs" directory of the repository) are:

- [Examples](#)
- [Python Library \(dreqPy\)](#)
- [The Request XML document and Schema](#)
- [Spreadsheet view of the variable definitions](#)
- [A searchable list of variables in the request, linking to](#)
- [A browsable HTML view of the request](#)
- [Overview tables for tier 1, priority 1 and all tiers and priorities](#)
- [Discussion of issues: old forum, new github pages](#)
- [Registration for email list: CMIP6-DATAREQUEST@JISCMAIL.AC.UK](#)
- [Installation and usage of the python package](#)

**When problems are found, raise an issue!**  
"CMIP6\_DataRequest\_VariableDefinitions"

# What is your status/timeline?

---

- Agreement on
  - ➔ Global attributes
  - ➔ Filename template and directory structure
  - ➔ Search facets
- Global attribute reference CVs and supporting CVs established
  - ➔ Sources and institutions being added
  - ➔ Decadal prediction sub-experiments need to be defined
- Data request first (non-beta) release is imminent
  - ➔ Some additional revision of variable lists
  - ➔ Need to correct lots of details
- CMOR 3.2 released
- Summary: Information and CVs are largely in place for
  - ➔ Preparing model output for the archive
  - ➔ Checking output for compliance with CMIP6 requirement
  - ➔ Publishing output to ESGF
  - ➔ Enabling faceted search of archive

## Prospects for ongoing stable funding? **Uncertain**

---

- Development of a rational organization of experiments and model results requires a comprehensive understanding of climate models and climate science
  - Has been generously supported by DOE (largely through PCMDI)
  - Requires active leadership by climate scientists, but isn't a "research" activity and funding is difficult to defend.
- Development of metadata requirements and CV's requires input from an expanding variety of perspectives.
  - Before CMIP6, PCMDI consulted with individual experts and was largely responsible for this.
  - Now PCMDI relies on the WIP volunteers to work toward consensus
    - Process more open and inclusive
    - Process is less efficient

# Summary of resources for CVs, metadata, data requirements

---

- Document defining global attributes, the “Data Reference Syntax” (DRS), the filename template and directory structure: <https://goo.gl/Cqbd6ii> (or accessible from [https://www.earthsystemcog.org/projects/wip/position\\_papers](https://www.earthsystemcog.org/projects/wip/position_papers))
- CMIP6 reference controlled vocabularies (CVs): [https://github.com/WCRP-CMIP/CMIP6\\_CVs](https://github.com/WCRP-CMIP/CMIP6_CVs)
- Document defining search facets for CMIP6: <https://goo.gl/rAvXIB>
- What variables should be archived and what attributes should be recorded? (“Data Request” information): <https://www.earthsystemcog.org/projects/wip/CMIP6DataRequest>

# Summary of resources for CVs, metadata, data requirements

---

- **CF Conventions** <http://cfconventions.org/cf-conventions>
  - Standard names
  - Conventions
- **CMIP5 output requirements**  
[http://cmip-pcmdi.llnl.gov/cmip5/docs/CMIP5\\_output\\_metadata\\_requirements.pdf](http://cmip-pcmdi.llnl.gov/cmip5/docs/CMIP5_output_metadata_requirements.pdf)
- **CMOR3 available to meet metadata requirements**
  - Code available from <https://github.com/PCMDI/cmor>
  - Documentation available at <http://cmor.llnl.gov/>

