# Distributing a huge volume of paleoclimate model data: the successful collaboration between PMIP and ESGF
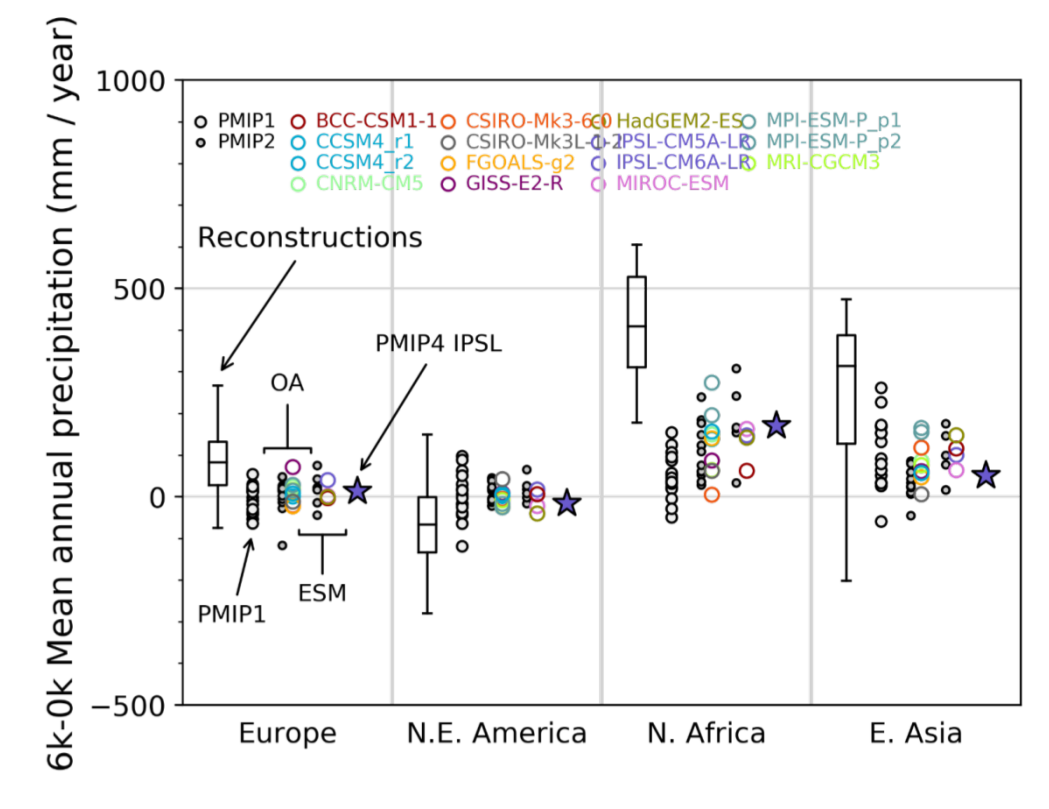
Jean-Yves **Peterschmitt**[1] , Sébastien Denvil[2] , Guillaume Levavasseur[2] , Mark Greenslade[2] , Atef Ben Nasser[2]
1) LSCE-IPSL, L'Orme des Merisiers - Bat 701, 91191 Gif-sur-Yvette, France, Jean-Yves.Peterschmitt@lsce.ipsl.fr 2) IPSL, UPMC, Paris, France

Paleoclimate Modelling
PMIP 4
Intercomparison Project

A Brief History of PMIP Time

Phase	PMIP 1	PMIP 2	PMIP 3	PMIP 4
DB online	1996	2005	2011	**2019**
Size	1.7 Gb	482 Gb	distributed *several Tb*	distributed *LOTS of Tb ...*
Distribution	ftp server LSCE (+PCMDI)	DODS server LSCE	**CMIP5** ESGF	**CMIP6** ESGF
Number of groups/models	22	18	25	20
Number of countries	11	10	12	14
Data format & Convention	NetCDF AMIP/GDT	NetCDF CMIP+PMIP2/CF	NetCDF CMIP5/CF	NetCDF CMIP6/CF
Main experiments	0k 6k 21k	Same as PMIP 1	PMIP 2 + Last Millennium	PMIP3 + Last Interglacial + Mid Pliocene Warm Period + Last Deglaciation + DeepMIP + ...
Example IPSL atmos	lmcelmd5 LMD 5.3 64x50 x L11	IPSL-CM4-V1-MR LMDZ 96x72 x L19	IPSL-CM5A-LR 96x95 x L39	IPSL-CM6A-LR 144x143 x L79
Example NCAR atmos	ccsm3 CCSM3 128x64 x L18	CCSM CCSM 3.0 128x64 x L17	CCSM4 288x192 x L26	CESM2 288x192 x L32
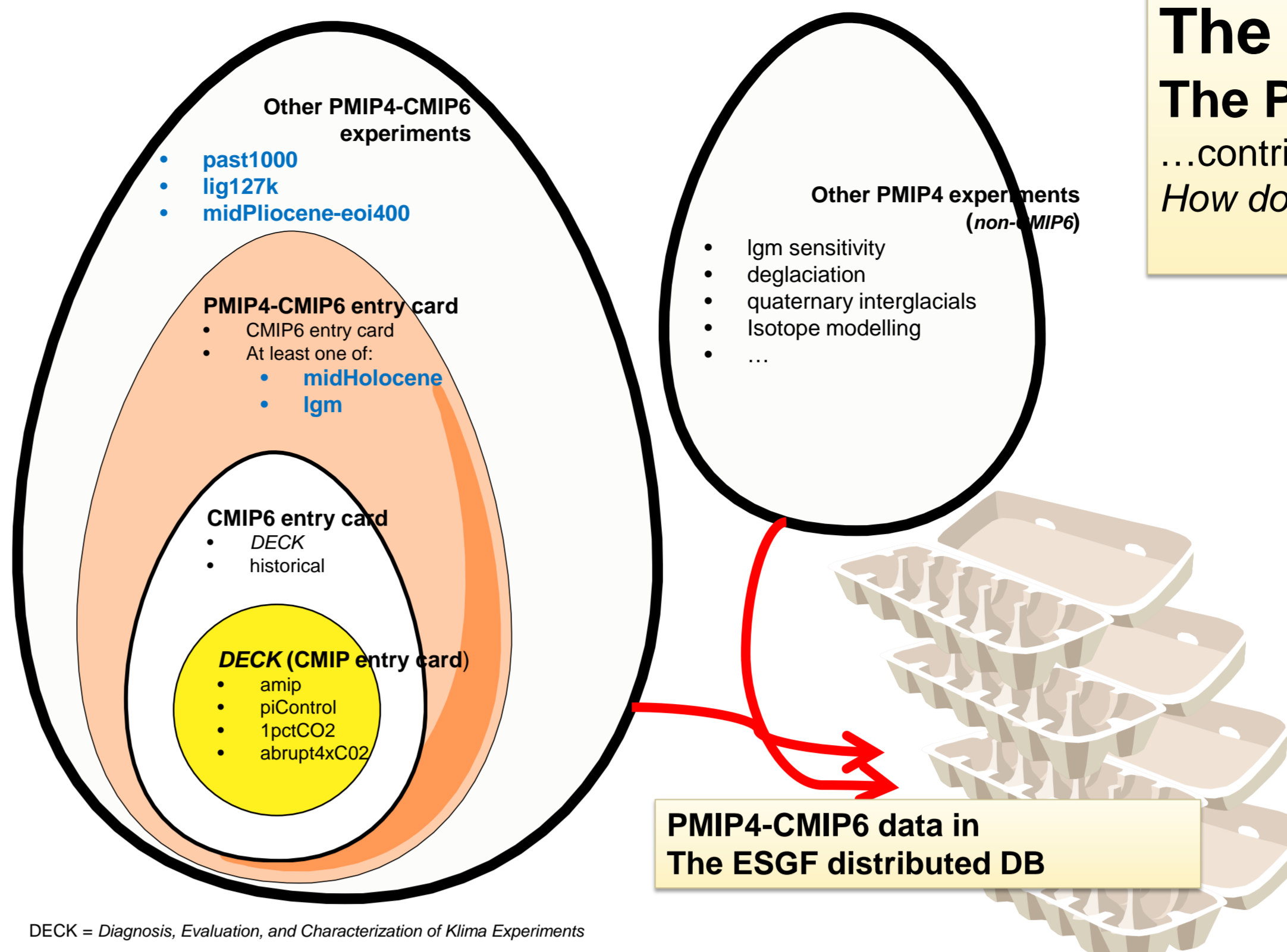
The Time (in the) Machine
The PMIP4-CMIP6 experiments…
…contribute to the CMIP6 question:
How does the Earth System respond to forcing?

DECK = *Diagnosis, Evaluation, and Characterization of Klima Experiments*

PMIP4-CMIP6 data in The ESGF distributed DB

PMIP4 Treasure Island
The PMIP4 participants *as of July 23rd 2019*

The PMIP4-CMIP6 participants — Information as of July 23rd 2019

PMP web   PMIP4 wiki

CMIP6 Terms of Use: data should be licensed under the Creative Commons Attribution-ShareAlike 4.0 International License

	Institute institution_id	Country	0k piControl	LM past1000 (1000 years)	6k midHolocene	21k lgm	Last Interglacial lig127k	Mid Pliocene warm period midPliocene-eoi400	Other PMIP experiments	Atm i_lon x j_lat x lev	Ocn i_lon x j_lat x lev	Model id source_id
1	AWI	Germany	Done		Done	Done	Done	Done ?	LIG DeepMIP	192x96 x L47	126859 x L46	AWI-ESM-1-1-LR
			Done		Done	Done	Done	Yes		192x96 x L47	256x220 x L40	MPI-ESM1-2-LR
2	CNRM-CERFACS	France	500				Yes			[256 to 20]x128 x L91		CNRM-CM6-1
3	CAS CAS-ESM	China	No End 2019	No End 2019	No End 2019		No End 2019			256x128 x L30	362x196 x L30	CAS-ESM1-0
4	CAS CAS-FGOALS	China	Done		Yes August 2019		Yes August 2019			Cube96 x L32	360x218 x L30	FGOALS-f3-L
			Done	No end-2019	Yes August 2019	No end-2019	Yes August 2019	No end-2019		180x90 x L26	360x218 x L30	FGOAL-g3
5	CSIR-Wits-CSIRO	South Africa								Cube192 x L35	Cube384 x L35	VRESM-1-0
6	INM	Russia			Done September 2019	Done September 2019	Done September 2019	Yes September 2019	DeepMIP	180x120 x L21	360x318 x L40	INM-CM-8
7	IPSL	France	500	Done end-2019	550 2x200	No	550	200	Yes	144x143 x L79	362x332 x L75	IPSL-CM6A-LR
8	KIOST	Korea	Done		No					Cube48 x L32	360x200 x L52	KIOST-ESM
9	MIROC (University of Tokyo and JAMSTEC)	Japan	Done	Yes August 2019	Yes August 2019	Yes September 2019				256x256 x L63	MIROC-ES2L MIROC4m
10	MPI-M	Germany	Done	Yes September 2019	Yes July 2019	Yes July 2019			Yes	192x96 x L47	256x220 x L40	MPI-ESM1-2-LR
11	MRI	Japan	Done	Done	Done	Done	Done			320x160 x L80	360x364 x L61	MRI-ESM2-0
12	NASA-GISS	USA	?	No	Done	Yes	Done	50		144x90 x L40 Cube90 x L102	360x180 x L32	GISS-E2-1-G GISS-E3-G
13	NCAR	USA	?	No End 2019	Done September 2019	No End 2019	No End 2019	No September 2019		288x192 x L32	320x384 x L60	CESM2
14	NCC (BCCR)	Norway	Done	No Dec 2019	No End 2019	No End 2019	No End 2019	No End 2019	DeepMIP	144x96 x L32	360x384 x L70	NorESM2-LM
			Done		Yes July 2019	Done	Done			144x96 x L26	360x384 x L70	NorESM1-F
15	NUIST	China	Done	No	Done	Done	Done			192x96 x L47	384x362 x L46	NESM3
16	EC-Earth-Consortium (Stockholm University)	Sweden	Done		Done September 2019	Done September 2019	Done September 2019	Yes September 2019		512x256 x L91	362x292 x L75	EC-Earth3
				Yes 2019						320x160 x L62	362x292 x L75	EC-Earth3-Veg-LR
					No	No	No	No		512x256 x L91	362x292 x L75	EC-Earth3-Veg
17	MOHC UK Academic Community	UK	Done	No	Done	Done	Done	Yes	Yes	192x144x L85	360x330 x L75	UKESM1-0-LL HadGEM3-GC31-LL
18	UTAS	Australia	No October 2019	No October 2019	No October 2019				Yes	64x56 x L18	128x112 x L31	CSIRO-Mk3L-1-3
19	UofT	Canada	Done		Yes	Done		Done		288x192 x L26	384x320 x L60	UofT-CCSM4
20	VUAmsterdam	The Netherlands	Done		No End-2019	Done	No End-2019		Last Degla	64x32 x L3	120x65 x L20	iLOVECLIM1.2

Legend — Running status: expected completion date — Not (running) | Yes | Done | Won't do
Available in CMIP6 DB (nb ens & nb years in CMIP6 DB) | Available in PMIP4 DB (nb years in PMIP4 DB)

CMIP6 Controlled Vocabulary lists: Institute | Model | CV @ github
CMIP6 | CMIP6 Participation Guidance for Modelers | CMIP6 Data Request
ESGF | CMOR3 | ESGF Node @ IPSL

Great Expectations
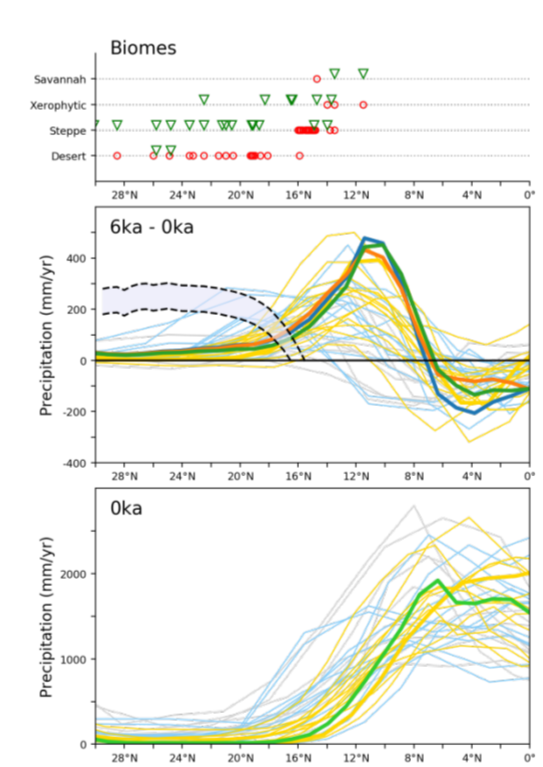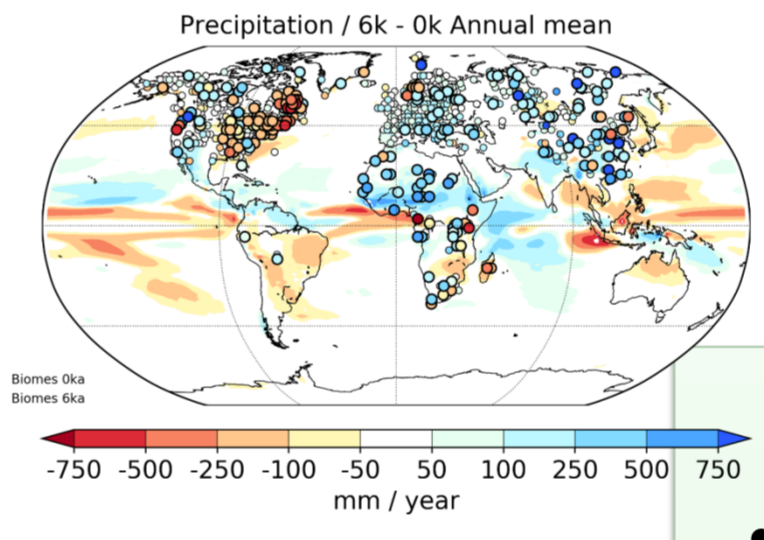### ESGF and the MIP data distribution

How can we successfully **share TERABYTES of model data** ? Use **Standards**!

- PMIP 3 & 4 data files are hosted in the **Coupled Model Intercomparison Project** (**CMIP 5 & 6**) **distributed database**
- The **standards** and tools for handling the data are defined and maintained by the **Earth System Grid Federation** (**ESGF**).
  - Climate and Forecast (CF) convention
  - Data Reference Syntax (DRS)
  - Controlled Vocabulary (CV)
  - Data Request (DR)
- **Standards** make it possible to efficiently and reliably exchange and use data
  - Standard NetCDF file format: self-documented binary file format that can be easily handled by programs
  - Standard variable names and metadata (units, axes, description, …)
  - 1 variable/file, and standard file names
  - Standard documentation
- Data files can be searched on any **ESGF portal**, and will be downloaded from the **ESGF Data Nodes** (DN) hosting the data
  - DN are maintained my major participating modelling groups or countries

It's alive!
Frankenstein movie, 1931

PMIP data on ESGF, 2019

Please combine the different models correctly to avoid creating a monster!

## Sense and Sensitivity
The PMIP working groups

- **PMIP data distribution**
  M. Kageyama
- **Past2Future: insights from a constantly varying past**
  J. Hargreaves, C. Brierley, P. Braconnot
- **Past2K**
  J. Jungclaus
- **PMIP-DATA**
  P. Bartlein, S. Harrison, P. Bartlein, M. Cruciix, A. Paul
- **Data assimilation**
  M. Crucifix
- **PlioMIP**
  A. Haywood, H. Dowsett
- **Quaternary Interglacials**
  B. Otto-Bliesner, D. Lunt
- **Pre-Pliocene climates**
  D. Lunt, B. Otto-Bliesner
- **Ice sheet uncertainties**
  A. Abe-Ouchi
- **Isotope modelling**
  A. LeGrande
- **Last Deglaciation**
  R. Ivanović, L. Gregoire, D. Roche

Scientists in Wonderland
How can we help you write your own book ?

Typical MIP data usage workflow
(Let us know what you needs)

- Determine **which variables** you need, at **which frequency** (use the *Data Request* if you don't know what is theoretically available), from **which experiments**
- Query an **ESGF portal** to determine for which models the variables are available
  - IPSL ESGF node: https://esgf-node.ipsl.upmc.fr/
- **Download** the required data for **all** the available models
  - Use *http/globus/opendap download*, or *wget* script generated by the portal
  - Use the *synda* (http://prodiguer.github.io/synda/) tool to automatically download/update data
    - *synda* can be used by institutes to **maintain a local mirror** of the most requested variables
- Use **cdo** (Climate Data Operators) and the **nco** (netCDF operator) toolkit to preprocess/reduce the downloaded netCDF data
- **Analyze** the data:
  - Use your favorite analysis software to **interactively** analyze the data
  - Write a **script** to easily reproduce the analysis if the data files are updated, or when more models become available
  - If you are not comfortable (yet) with programming, you can use an *ipython notebook* to combine interaction+script
  - Let us know about your needs, so that we can check if they could be added to a portal similar to the climate4impacts portal (https://climate4impact.eu/)
- **Cite** the data: you are supposed to link your analyses back to the original data!
  http://cmip6cite.wdc-climate.de/

ESGF portal

References
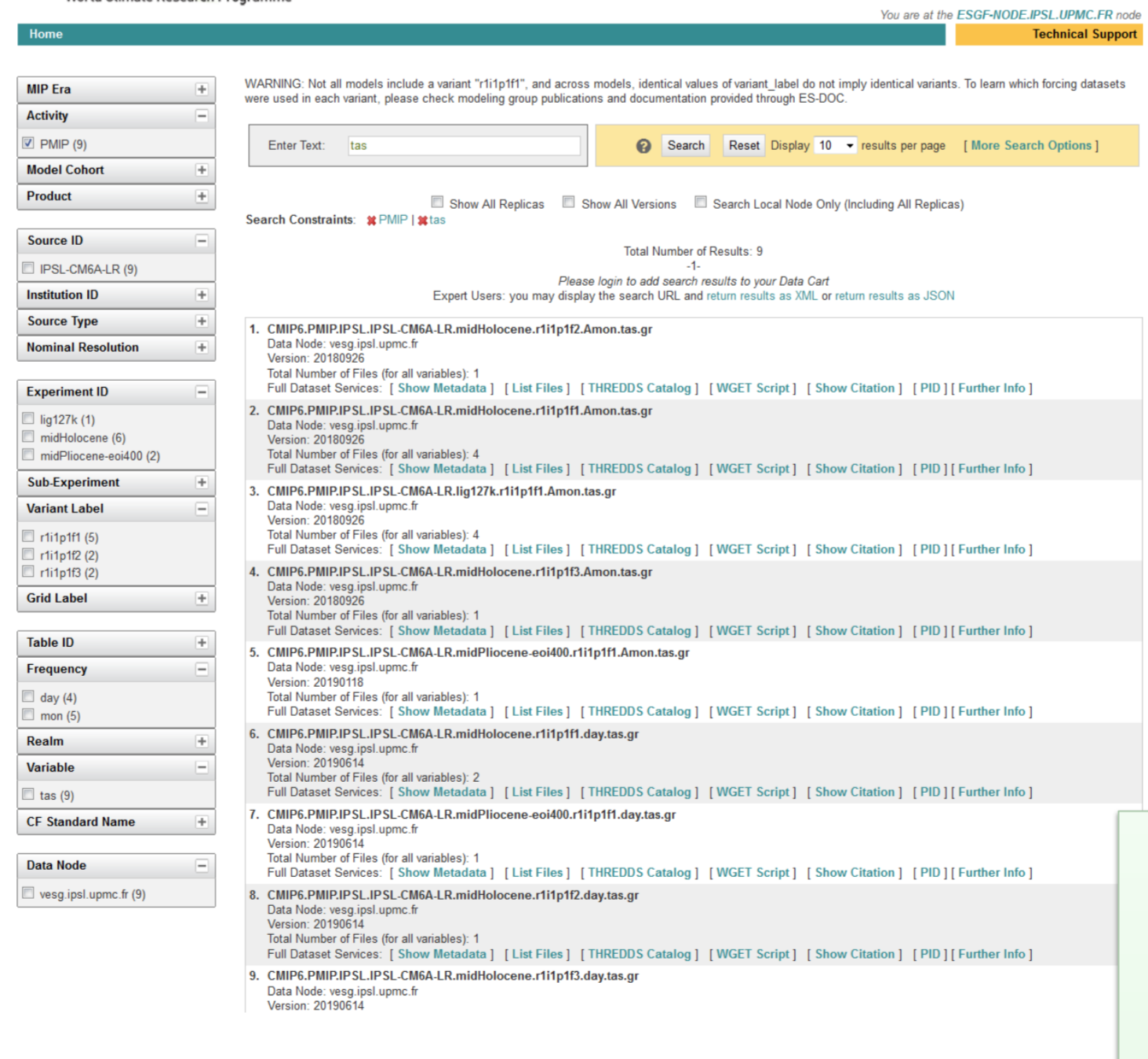
- PMIP: https://pmip.lsce.ipsl.fr/
- PMIP4: https://pmip4.lsce.ipsl.fr
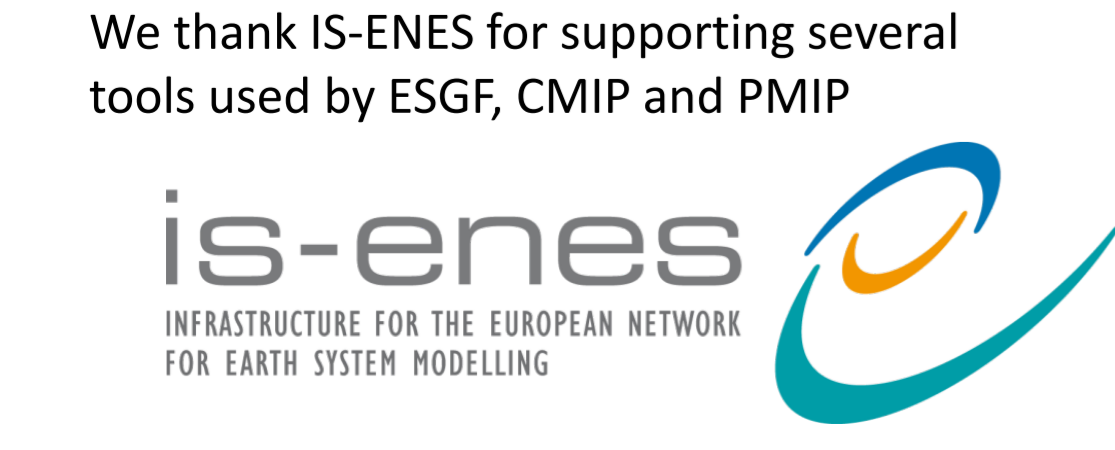- Kageyama, M. et al, *The PMIP4 contribution to CMIP6 - Part 1: Overview and over-arching analysis plan*, Geosci. Model Dev., 11, 1033-1057, https://doi.org/10.5194/gmd-11-1033-2018 , 2018
- ESGF: https://esgf.llnl.gov
- Balaji V. et al, *Requirements for a global data infrastructure in support of CMIP6*, Geosci. Model Dev., 11, 3659-3680, https://doi.org/10.5194/gmd-11-3659-2018 , 2018
- A Guide to CMIP6 Participation: https://pcmdi.llnl.gov/CMIP6/Guide/index.html
- Eyring, V. et al, *Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization*, Geosci. Model Dev., 9, 1937-1958, https://doi.org/10.5194/gmd-9-1937-2016 , 2016
- CMIP6 Data Request: https://earthsystemcog.org/projects/wip/CMIP6DataRequest
- CMIP6 CV: https://github.com/WCRP-CMIP/CMIP6_CVs
- es-doc search: https://search.es-doc.org/

We thank IS-ENES for supporting several tools used by ESGF, CMIP and PMIP

is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

Add your own PMIP literary contribution here ... and win a demo !

Lord of the tree rings

PMIP is endorsed by **PAGES** PAST GLOBAL CHANGES   **WCRP** World Climate Research Programme