Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

CMIP standards and the need for conventions in the Copernicus Climate Data Store

CF workshop, Exeter, UK, 23/09/2019

Exeter, 23/09/2019

Pierre-Antoine Bretonnière

Outline

1. CMOR and CMIP
   a. Setting the scene
   b. Much more than “technical details”
   c. Governance
   d. CF in CMOR

2. Copernicus, C3S and C3S_512
   a. Project(s) overview
   b. Identified standards and file formats
   c. Necessity for generic, multi-dataset conventions
   d. Data checkers

What is CMOR?

- Climate Model Output Rewriter
  - C library developed by PCMDI that ingests any kind of file and convert them into NetCDF4, with project specific (CMIP5, SPECS, PRIMAVERA, CMIP6,...) requirements.
- The requirements cover many aspects such as:

time interval

file names

directory tree

units

file compression

standard name

dimensions

frequency

variable name

axis orientation

What is CMOR?

```
pbretonn@bscearth319:/esarchive/exp/ecearth/a1ua/cmorfiles/DCPP/EC-Earth-Consortium/EC-Earth3/dcppA-hindcast$ tree
.
├── r1i1p1f1
│   ├── 6hrPlev
│   │   ├── psl
│   │   │   └── gr
│   │   │       └── v20190320
│   │   │           └── psl_6hrPlev_EC-Earth3_dcppA-hindcast_s19601101-r1i1p1f1_gr_196011010900-196110312100.nc
│   │   ├── uas
│   │   │   └── gr
│   │   │       └── v20190320
│   │   │           └── uas_6hrPlev_EC-Earth3_dcppA-hindcast_s19601101-r1i1p1f1_gr_196011010900-196110312100.nc
│   │   └── vas
│   │       └── gr
│   │           └── v20190320
│   │               └── vas_6hrPlev_EC-Earth3_dcppA-hindcast_s19601101-r1i1p1f1_gr_196011010900-196110312100.nc
│   ├── AERday
│   │   └── zg500
│   │       └── gr
│   │           └── v20190320
│   │               └── zg500_AERday_EC-Earth3_dcppA-hindcast_s19601101-s1960-r1i1p1f1_gr_19601101-19611031.nc
│   ├── Amon
│   │   ├── clt
│   │   │   └── gr
│   │   │       └── v20190320
│   │   │           └── clt_Amon_EC-Earth3_dcppA-hindcast_s19601101-r1i1p1f1_gr_196011-196110.nc
│   │   ├── evspsbl
│   │   │   └── gr
│   │   │       └── v20190320
│   │   │           └── evspsbl_Amon_EC-Earth3_dcppA-hindcast_s19601101-r1i1p1f1_gr_196011-196110.nc
│   │   ├── hfls
│   │   │   └── gr
│   │   │       └── v20190320
│   │   │           └── hfls_Amon_EC-Earth3_dcppA-hindcast_s19601101-s1960-r1i1p1f1_gr_196011-196110.nc
│   │   ├── hfss
│   │   │   └── gr
│   │   │       └── v20190320
│   │   │           └── hfss_Amon_EC-Earth3_dcppA-hindcast_s19601101-r1i1p1f1_gr_196011-196110.nc
│   │   ├── hurs
│   │   │   └── gr
│   │   │       └── v20190320
│   │   │           └── hurs_Amon_EC-Earth3_dcppA-hindcast_s19601101-s1960-r1i1p1f1_gr_196011-196110.nc
│   │   ├── hursmin
│   │   │   └── gr
│   │   │       └── v20190320
│   │   │           └── hursmin_Amon_EC-Earth3_dcppA-hindcast_s19601101-r1i1p1f1_gr_196011-196110.nc
│   │   ├── huss
│   │   │   └── gr
│   │   │       └── v20190320
```

What is CMOR?

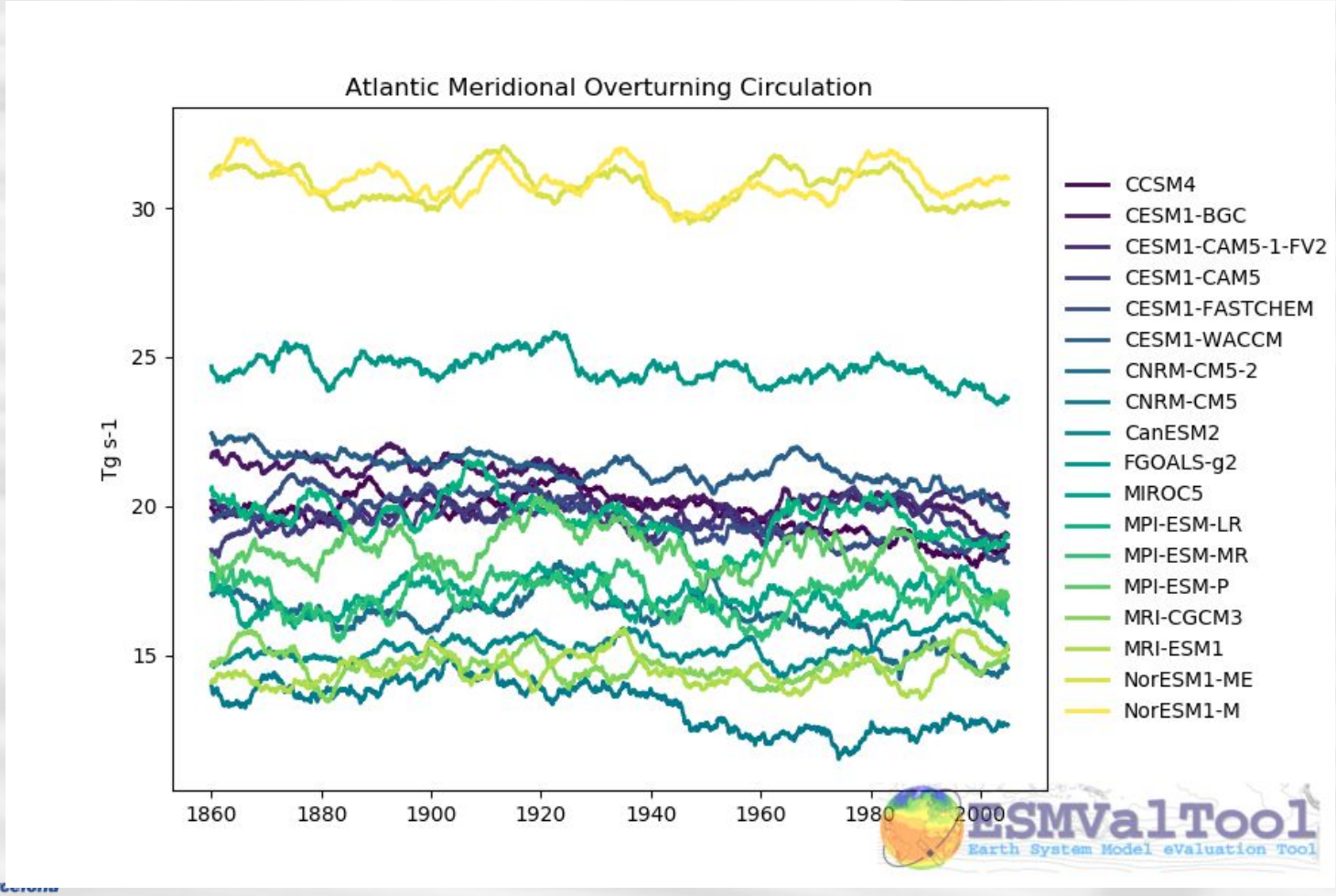
```
pbretonn@bscearth319:/esarchive/exp/ecearth/alua/cmorfiles/DCPP/EC-Earth-Consortium/EC-Earth3/dcppA-hindcast$ tree
.
├── r1i1p1f1
│   ├── 6hrPlev
│   │   ├── psl
│   │   │   netcdf tos_Omon_EC-Earth3_dcppA-hindcast_s19601101-r1i1p1f1_gr_196011-196110 {
│   │   │   dimensions:
│   │   │           time = UNLIMITED ; // (12 currently)
│   │   │           lat = 256 ;
│   │   │           lon = 512 ;
│   │   │           bnds = 2 ;
│   │   │   variables:
│   │   │           double time(time) ;
│   │   │                   time:bounds = "time_bnds" ;
│   │   │                   time:units = "days since 1960-11-01 00:00:00" ;
│   │   │                   time:calendar = "proleptic_gregorian" ;
│   │   │                   time:axis = "T" ;
│   │   │                   time:long_name = "time" ;
│   │   │                   time:standard_name = "time" ;
│   │   │           double time_bnds(time, bnds) ;
│   │   │           double lat(lat) ;
│   │   │                   lat:bounds = "lat_bnds" ;
│   │   │                   lat:units = "degrees_north" ;
│   │   │                   lat:axis = "Y" ;
│   │   │                   lat:long_name = "latitude" ;
│   │   │                   lat:standard_name = "latitude" ;
│   │   │           double lat_bnds(lat, bnds) ;
│   │   │           double lon(lon) ;
│   │   │                   lon:bounds = "lon_bnds" ;
│   │   │                   lon:units = "degrees_east" ;
│   │   │                   lon:axis = "X" ;
│   │   │                   lon:long_name = "Longitude" ;
│   │   │                   lon:standard_name = "longitude" ;
│   │   │           double lon_bnds(lon, bnds) ;
│   │   │           float tos(time, lat, lon) ;
│   │   │                   tos:standard_name = "sea_surface_temperature" ;
│   │   │                   tos:long_name = "Sea Surface Temperature" ;
│   │   │                   tos:comment = "Temperature of upper boundary of the liquid ocean, including temperatures below sea-ice and floating ice shelves." ;
│   │   │                   tos:units = "degC" ;
│   │   │                   tos:cell_methods = "area: mean where sea time: mean" ;
│   │   │                   tos:cell_measures = "area: areacello" ;
│   │   │                   tos:history = "2019-03-20T07:54:30Z altered by CMOR: Reordered dimensions, original order: lat lon time." ;
│   │   │                   tos:missing_value = 1.e+20f ;
│   │   │                   tos:_FillValue = 1.e+20f ;
│   │   │
│   │   │   // global attributes:
│   │   │                   :_NCProperties = "version=1|netcdflibversion=4.4.1.1|hdf5libversion=1.8.19" ;
│   │   │                   :Conventions = "CF-1.7 CMIP-6.2" ;
│   │   │                   :activity_id = "DCPP" ;
│   │   │                   :branch_method = "no parent" ;
│   │   │                   :branch_time = 0. ;
│   │   │                   :branch_time_in_child = 0. ;
│   │   │                   :branch_time_in_parent = 0. ;
│   │   │                   :contact = "cmip6-data@ec-earth.org" ;
│   │   │                   :creation_date = "2019-03-20T07:54:30Z" ;
│   │   ├── AERda
│   │   │   └── 
│   │   ├── Amon
│   │   │   ├── 
│   │   │   ├── 
│   │   │   ├── 
│   │   │   ├── 
│   │   │   ├── 
│   │   │   ├── 
│   │   │   ├── 
│   │   │   └── 
```

Why CMOR?

- CMOR is much more than the library itself...
- Better metadata ⟹ improved trust
- Enhanced “formatting” ⟹ more “FAIR”

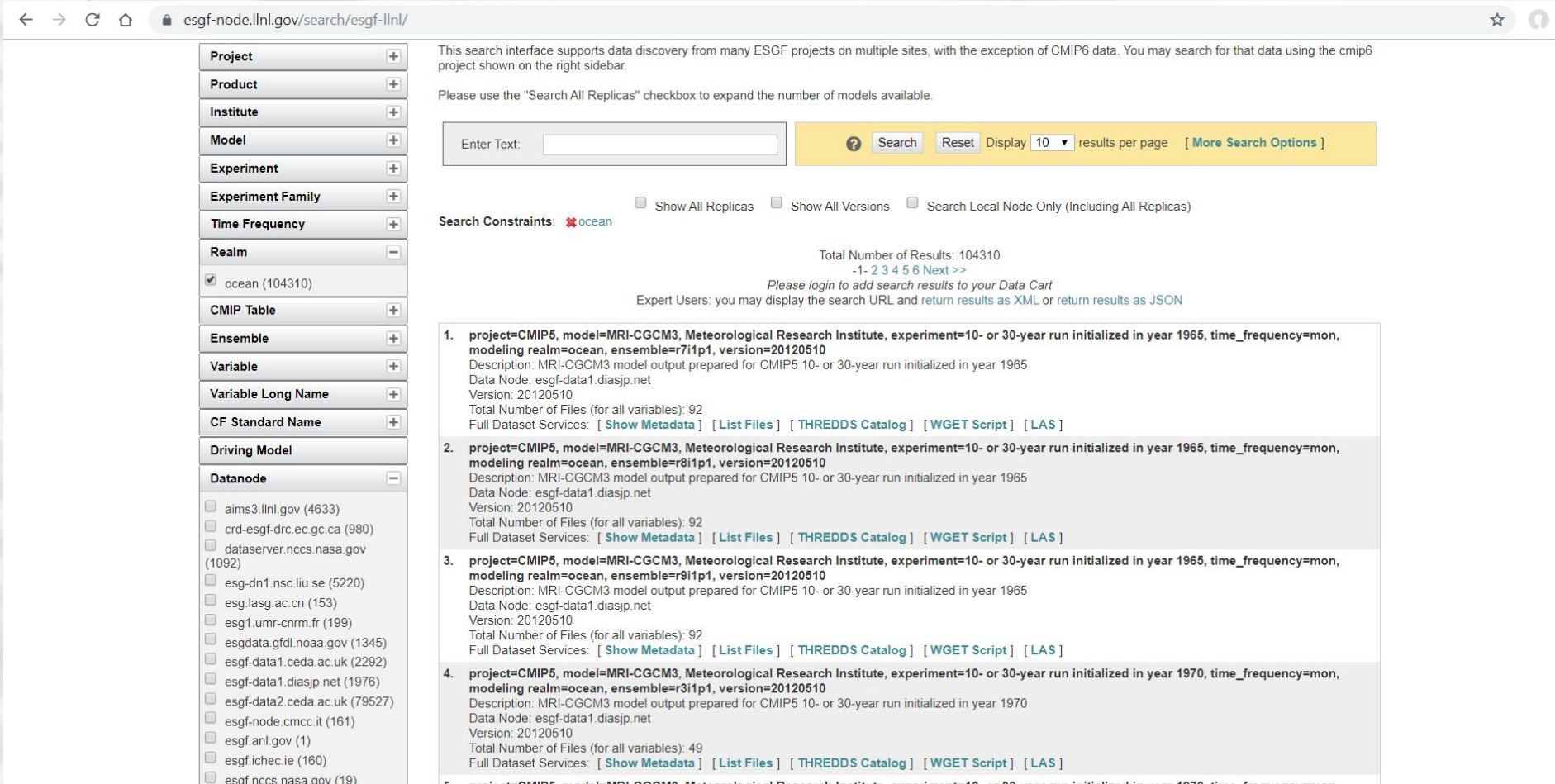
Why CMOR?

- help data sharing for better science



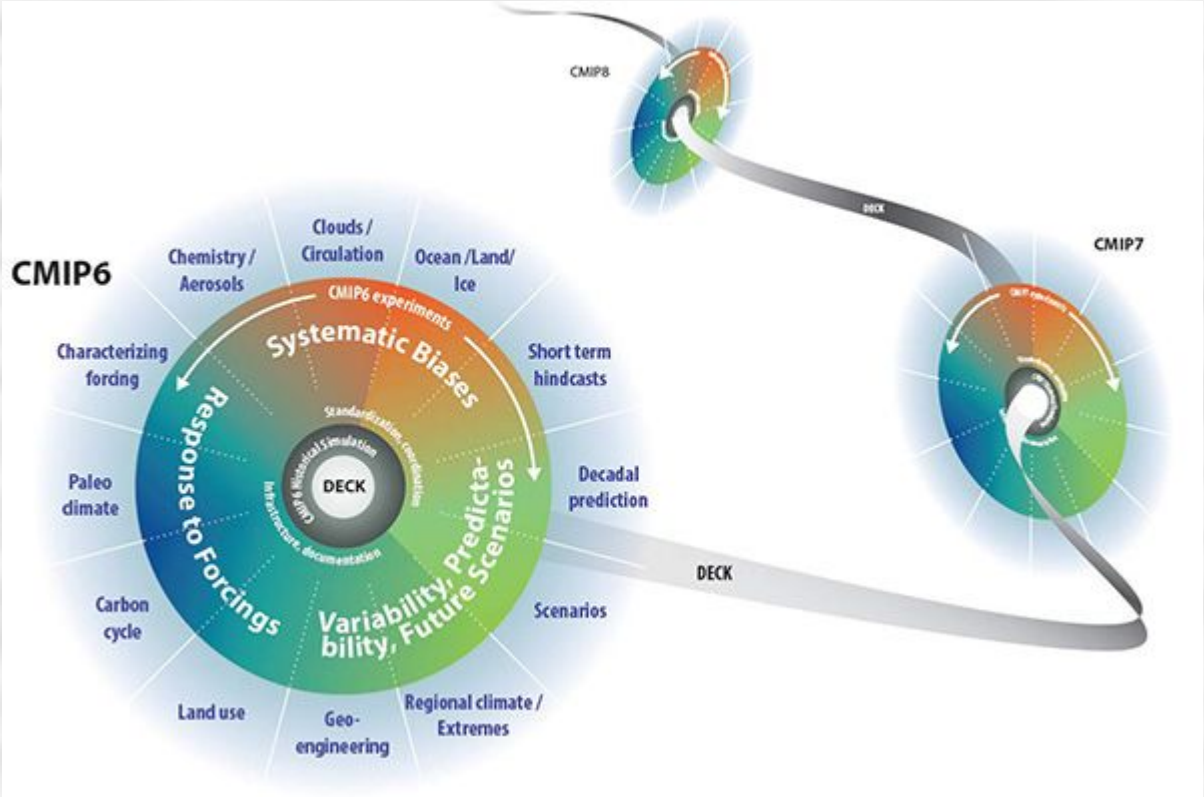
Why CMOR?

- **E**arth **S**ystem **G**rid **F**ederation: an example of data sharing platform where CMOR is the basis for data indexing and search



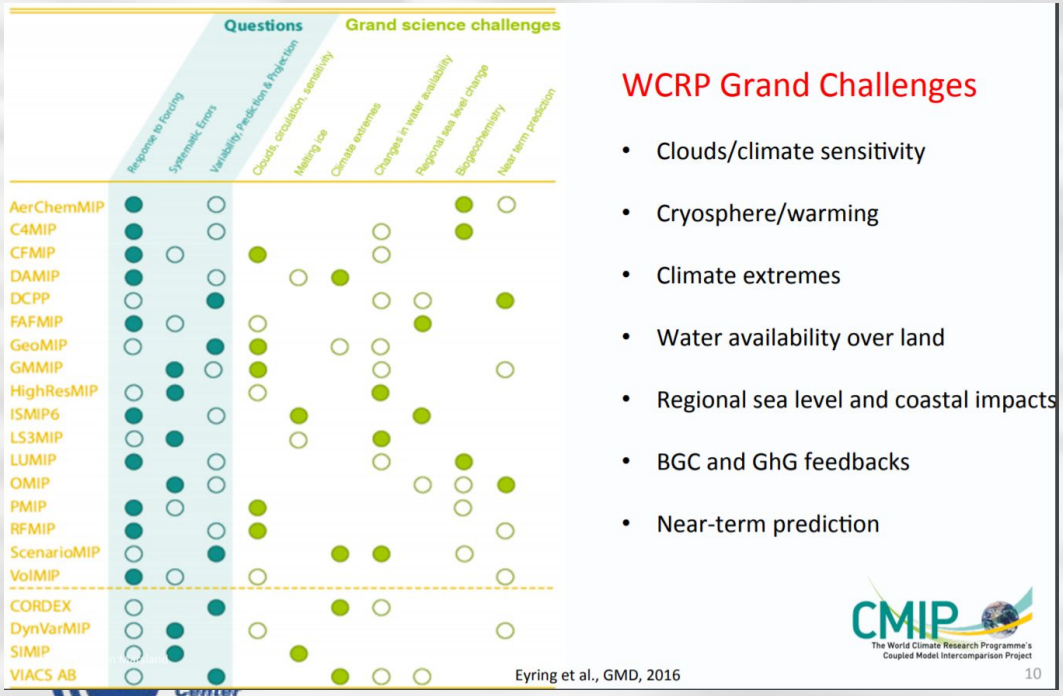
CMOR in CMIP

- CMOR(1) with CMIP3
- CMOR(2) with CMIP5
- CMOR(3) now
- Growing need to describe more experiments, models, processes,...



CMOR in CMIP

- Only for CMIP6, there are already more than 23 MIPs, with different attributes and scientific needs



I. Overview CMIP6-Endorsed MIPs

1. AerChemMIP- Aerosols and Chemistry Model Intercomparison Project
2. C^{4}MIP - Coupled Climate Carbon Cycle Model Intercomparison Project
3. CDRMIP - The Carbon Dioxide Removal Model Intercomparison Project
4. CFMIP - Cloud Feedback Model Intercomparison Project
5. DAMIP - Detection and Attribution Model Intercomparison Project
6. DCPP - Decadal Climate Prediction Project
7. FAFMIP - Flux-Anomaly-Forced Model Intercomparison Project
8. GeoMIP - Geoengineering Model Intercomparison Project
9. GMMIP - Global Monsoons Model Intercomparison Project
10. HighResMIP - High-Resolution Model Intercomparison Project
11. ISMIP6 - Ice Sheet Model Intercomparison Project for CMIP6
12. LS3MIP - Land Surface, Snow and Soil Moisture
13. LUMIP - Land-Use Model Intercomparison Project
14. OMIP - Ocean Model Intercomparison Project
15. PAMIP – Polar Amplification Model Intercomparison Project
16. PMIP - Palaeoclimate Modelling Intercomparison Project
17. RFMIP - Radiative Forcing Model Intercomparison Project
18. ScenarioMIP - Scenario Model Intercomparison Project
19. VolMIP - Volcanic Forcings Model Intercomparison Project
20. CORDEX - Coordinated Regional Climate Downscaling Experiment
21. DynVarMIP - Dynamics and Variability Model Intercomparison Project
22. SIMIP - Sea Ice Model Intercomparison Project
23. VIACS AB - Vulnerability, Impacts, Adaptation and Climate Services Advisory Board

CMOR beyond CMIP

- Other types of projects need different conventions to represent different scientific concepts:
  - observations (obs4MIP)
  - reanalysis (ana4MIPs)
  - seasonal forecast (SPECS): double time axis

CF inside CMOR

CMIP =

a. experiment design (physical processes, initialization, forcings,...)
b. + coupled Earth System Models for each
c. + physical variable representation

CF is here to describe the third part: agreement on precise definition of what the representation of a physical quantity/variable by a numerical model, and is a piece of all the CMOR “convention”


CMIP
The World Climate Research Programme's
Coupled Model Intercomparison Project


opernicus
Europe's eyes on Earth

Copernicus, C3S, and C3S_512

- Copernicus is the **European Union's Earth Observation Programme**, looking at our planet and its environment for the ultimate benefit of all European citizens. It offers information services based on satellite Earth Observation and in situ (non-space) data.

- The Programme is coordinated and managed by the European Commission. It is implemented in partnership with the **Member States**, the European Space Agency (**ESA**), the European Organisation for the Exploitation of Meteorological Satellites (**EUMETSAT**), the European Centre for Medium-Range Weather Forecasts (**ECMWF**), EU Agencies and Mercator Océan.

- Vast amounts of global data from satellites and from ground-based, airborne and seaborne measurement systems are being used to provide information to help service providers, public authorities and other international organisations improve the quality of life for the citizens of Europe. The information services provided are freely and openly accessible to its users.

Copernicus, C3S, and C3S_512

Copernicus, C3S, and C3S_512

Copernicus, C3S, and C3S_512

climate.copernicus.eu/what-we-do

What we do

Our core objective is to provide reliable access to high-quality climate data. We do this through our Climate Data Store (CDS).
We also offer tools and expert guidance that make it possible to transform the data into more visual products, such as maps and charts.

Climate datasets

The CDS provides a single point of access to a variety of climate datasets, including observations, reanalyses of past observations, seasonal forecasts and climate model projections.

Tools for using climate data

The CDS features a powerful toolbox for processing and visualising data in the cloud, so that users can develop climate information suited to their needs.

Sectoral impacts

We provide real applications of CDS data and tools that demonstrate how businesses, governments and citizens can make informed decisions on how to mitigate the effects of climate change.

Quality assurance

We provide quality assurance for all CDS data, tools and applications. We continuously engage with users and independent experts to evaluate our services and ensure that they are fit for purpose.

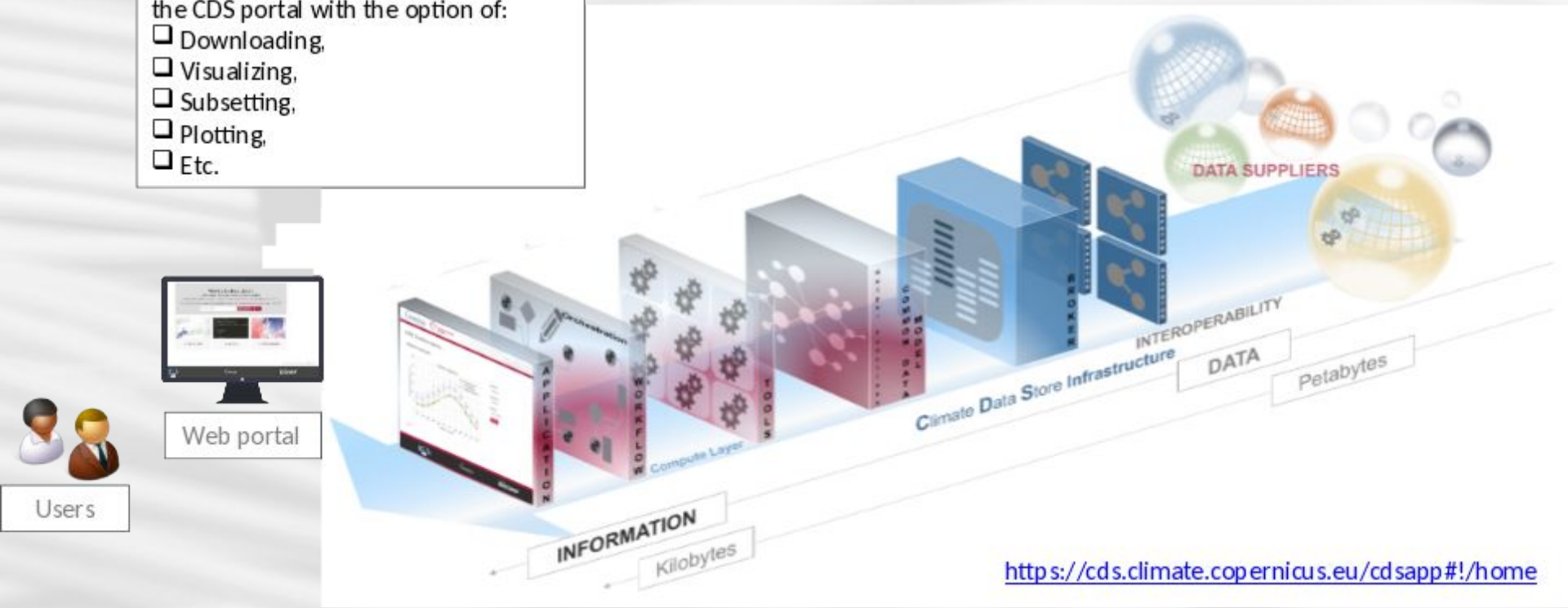
Copernicus, C3S, and C3S_512

CLIMATE DATA STORE (CDS)

At the heart of the C3S infrastructure is the ***Climate Data Store (CDS)***. It provides a single point of access to a wide range of climate datasets, namely satellite and in-situ observations, reanalyses, seasonal forecasts and climate projections

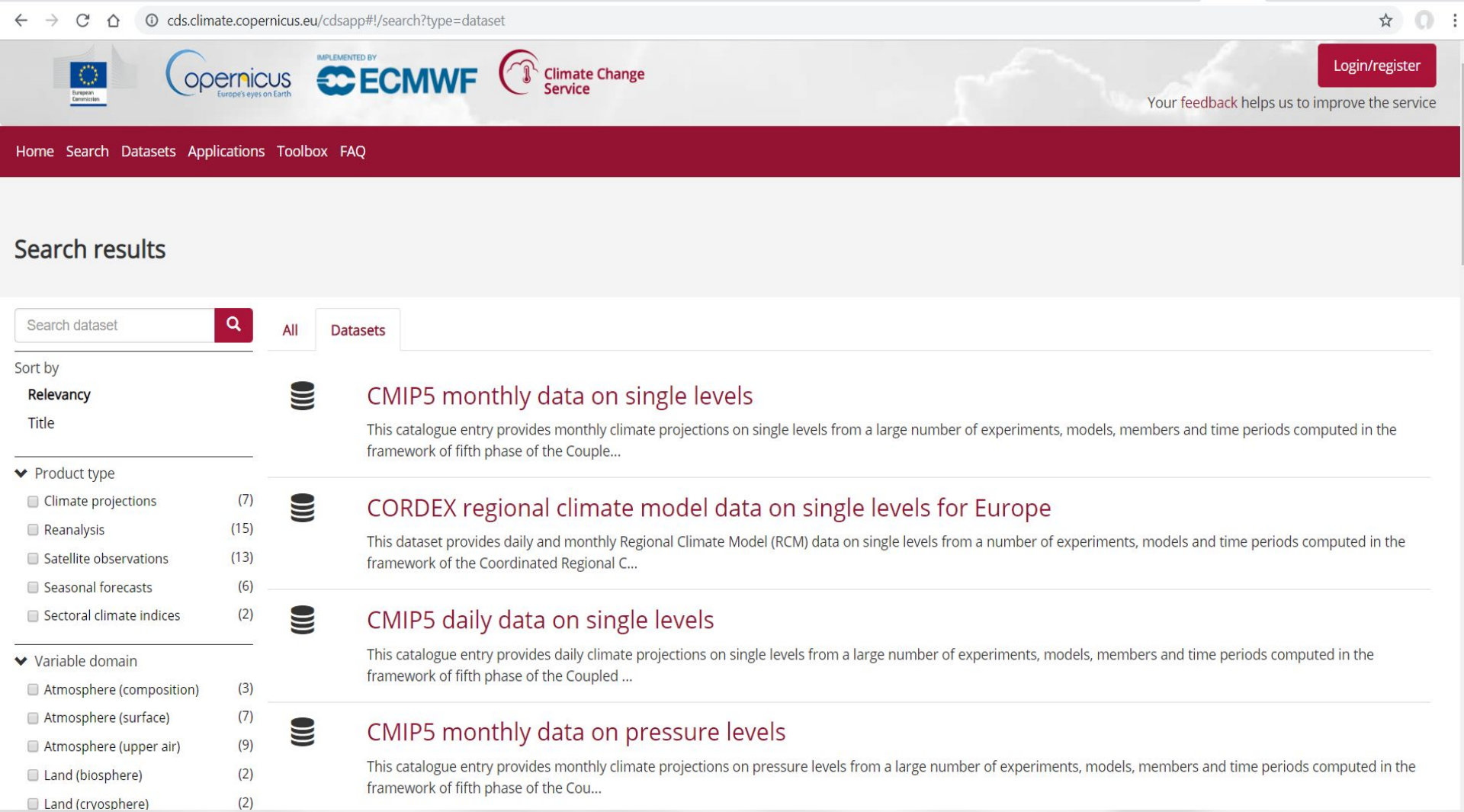
The users access the datasets through the CDS portal with the option of:

- ❑ Downloading,
- ❑ Visualizing,
- ❑ Subsetting,
- ❑ Plotting,
- ❑ Etc.



https://cds.climate.copernicus.eu/cdsapp#!/home

Copernicus, C3S, and C3S_512



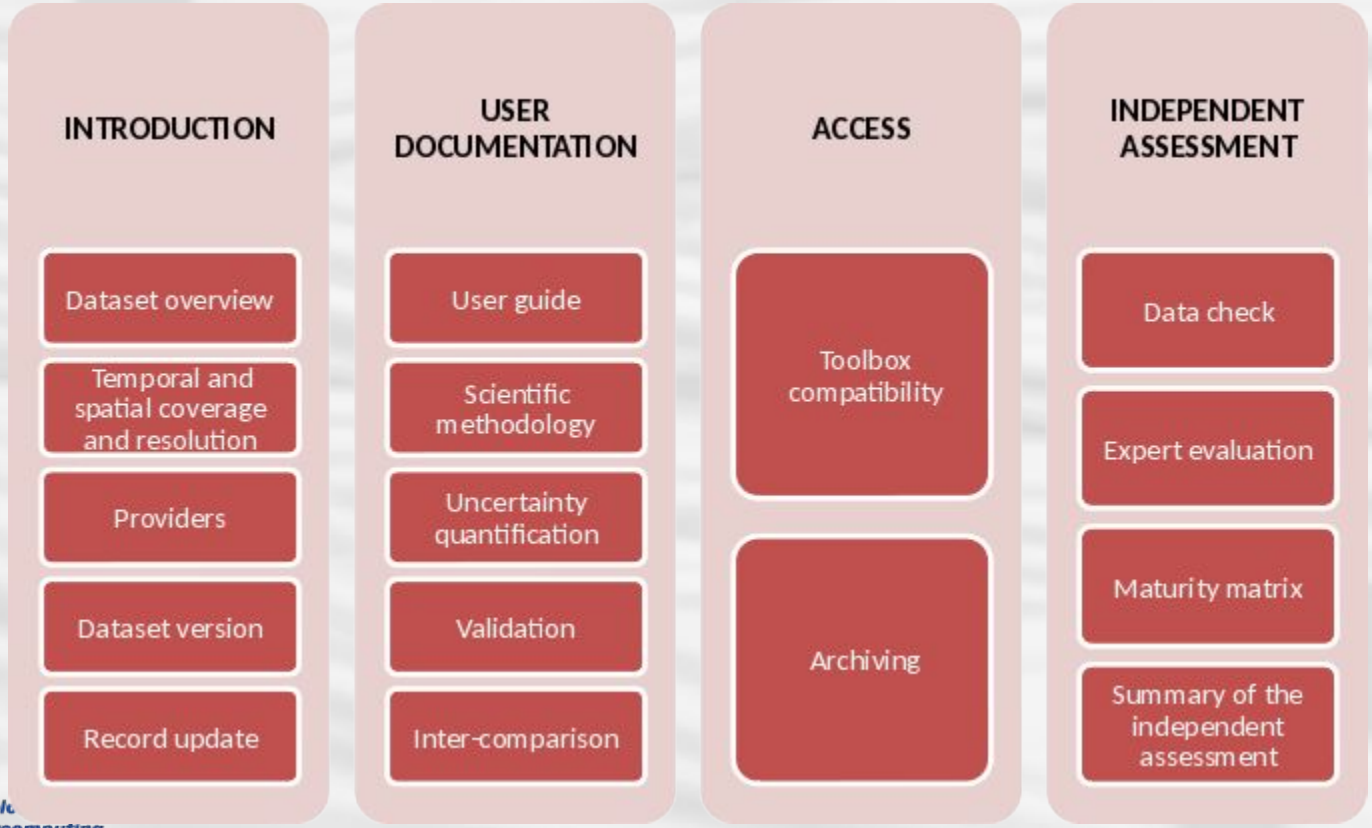
Copernicus, C3S, and C3S_512

- C3S_512 aims at providing Evaluation and Quality Control (EQC) of the Climate Data Store (CDS)

- "The evaluation and quality assessment will put at work the best expertise available on the **evaluation of the multi-faceted aspects of data** and product quality **based on the most complete set of standards accepted by the communities involved**. Surveys and other user engagement techniques will be used and analysed to provide a detailed mapping of the users and their needs, identifying those that should be addressed with priority. The outcome will be employed to perform a gap analysis of the current capabilities of the CDS and formulate recommendations that support the evolution of the service. "

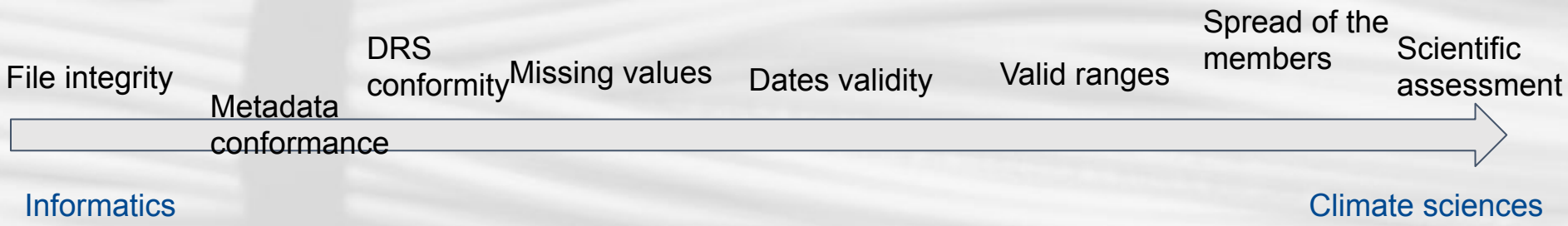
Copernicus, C3S, and C3S_512

- Led by BSC, with 6 other European partners (DWD, FMI, Predictia, CNR, WENR, MeteoFrance)
- November 2018 -> June 2021



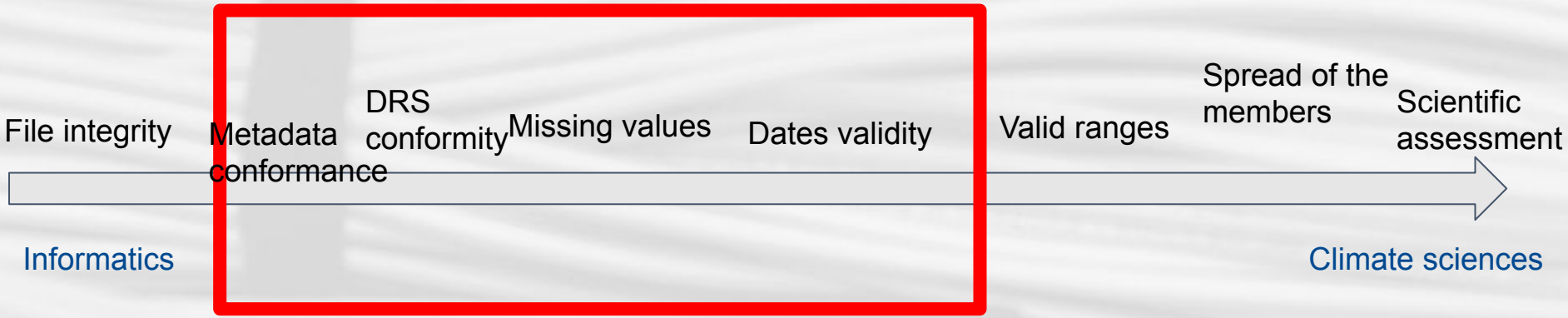
Copernicus, C3S, and C3S_512

- EQC goes from from very “informatics” check to scientific assessment of the data



Copernicus, C3S, and C3S_512

- EQC goes from from very “informatics” check to scientific assessment of the data



Copernicus, C3S, and C3S_512

- First level of check: data standard compliance
  - data come in multiple formats (grib1/2, nc3/4, csv, zip, shape...):
  - https://earth.bsc.es/gitlab/ces/c3s512-wp1-datachecker/blob/master/02_cds_inventory.md
  - deliverable on existing standards by type

Copernicus, C3S, and C3S_512

Based on these necessity, we are defining a data checker to enforce and flag the dataset, showing if they comply with the conventions recommendations.

Copernicus, C3S, and C3S_512

- Recommendations on standards:
  - CF
  - CMOR
  - SPECS/CMOR/C3S512 for seasonal
  - …

- Most of the datasets in the CDS currently don't have precise standards definition, CF is used as the default

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Thank you

pierre-antoine.bretonniere@bsc.es