

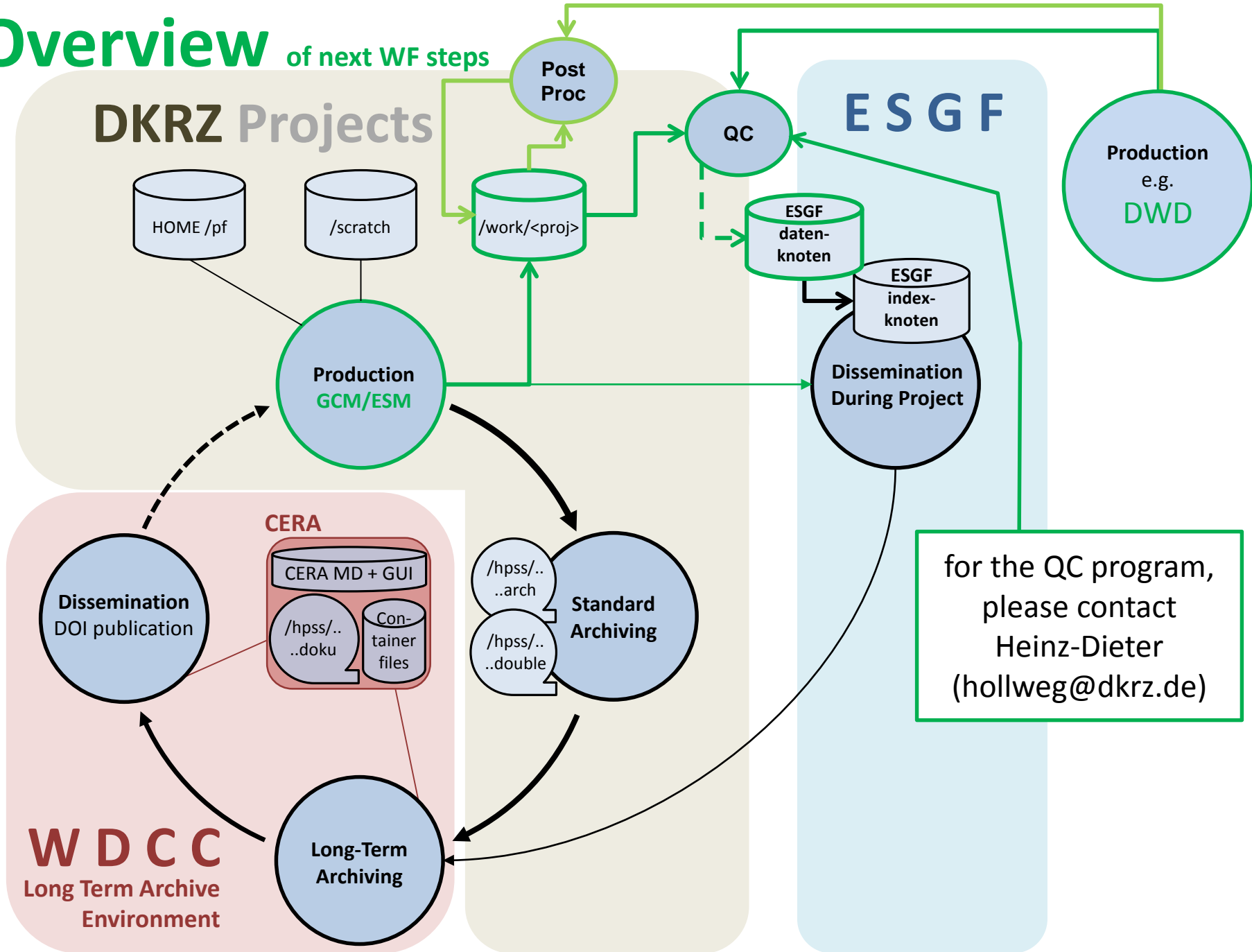


Best Practice bei der Datenerzeugung für die Earth System Grid Federation (ESGF)

- Converting data into a CMIP5-like standard -

Jörg Wegner, Stephanie Legutke, DKRZ/DM, 7 Mai 2014, DKRZ Hamburg

Overview of next WF steps



ESGF: Conditions for Data Admission

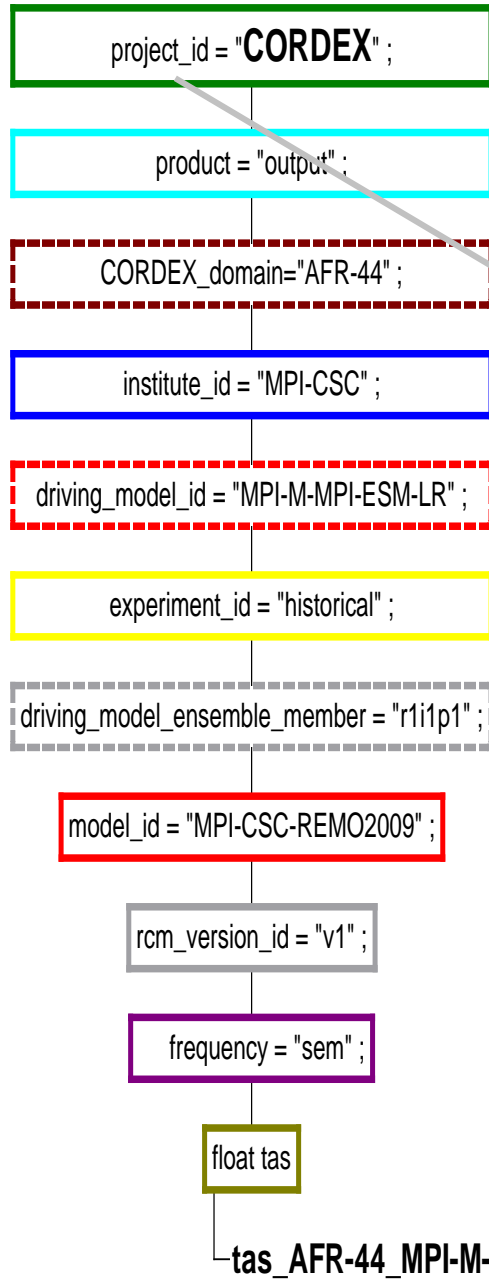
- **ESGF provides a GUI where data can be browsed, searched, and downloaded through an internet browser**
 - => the data must**
- **be generated for a model intercomparison project (e.g. CMIP5, CORDEX, PMIPn, obs4MIPs, ana4MIPs, ...)**
- **belong to the agreed-on set of project variables**
- **be provided with the project agreed-on meta data in order to allow for common search criteria**

What has to be specified ?

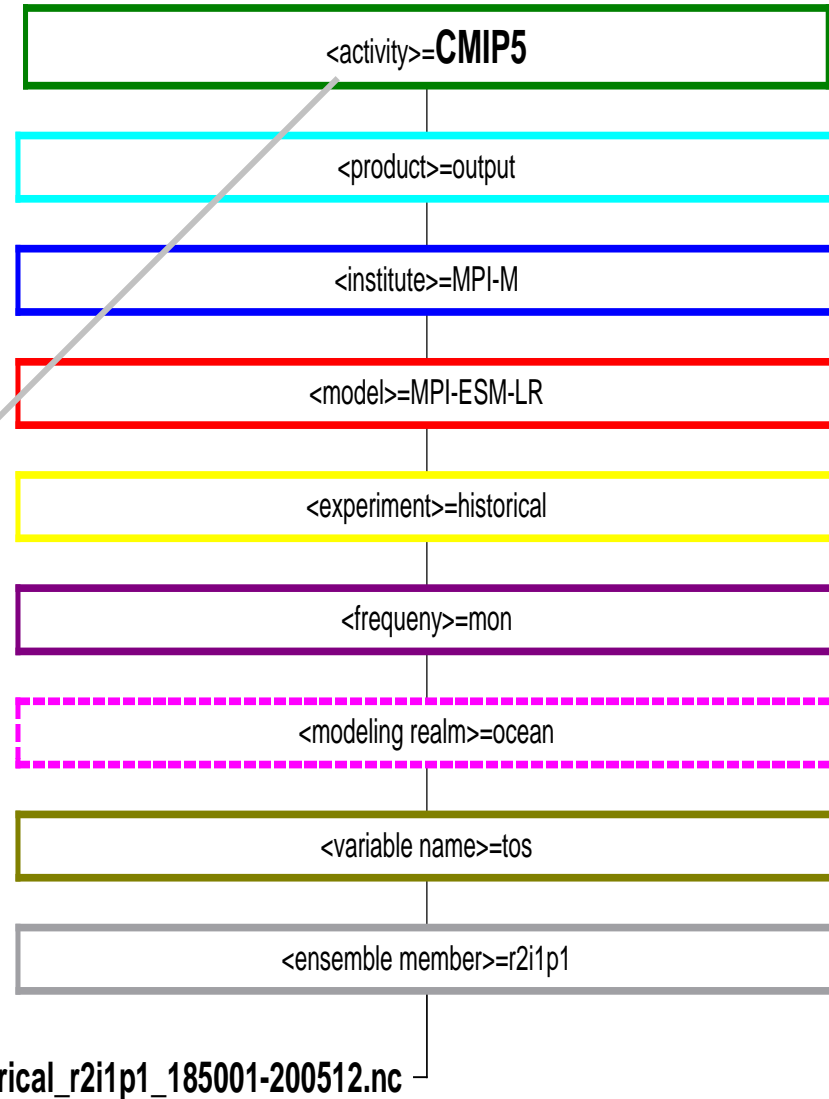
1. file format (e.g. NetCDFn, compression, single variable files, CF-1.n convention, ...)
2. file names implying the content
3. directory structure for appropriate file grouping
4. meta data (for processing tools, non-expert users)
5. controlled vocabulary (CV) for accurate search

DRS (Data Reference Syntax)

DRS examples



NetCDF attribute
=
DRS element !



tos_Omon_MPI-ESM-LR_historical_r2i1p1_185001-200512.nc

Classes of DRS elements

Example CMIP5:

*.../'**project**'/**output**/'center acronym'/'model name'/'
experiment ID'/'frequency'/'realm'/'
variable name/'**member identifier**/'*.nc*

- The **green** DRS elements are controlled vocabulary (CV) defined by the project
- The **black** can be chosen by the data provider within the naming restrictions and coordinated with the project lead
- The **brown** elements are free to choose within general (no underscore etc.) and project restrictions (r<l>i1<m>p<n>)

Decide whether to use PCMDI CMOR2 Software

Data has to be transformed into the CMIP5 ,de facto' standard

This can be done by

- **writing the aggregated variable and the requested coordinates in the correct units and with the correct precision including all attributes to a file with the correct path and name using the NetCDF4 library**
- **passing the aggregated variable and the requested coordinates with the correct units ... including a *subset* of attributes to subroutines of the CMOR library**

Decide whether to use PCMDI CMOR2 Software

The decision on whether to use CMOR2 has to be taken ,individually' depending on whether

- CMOR MIP tables already exists
- the number of data provider is > 1 per variable
- the archive is complex
- CMIP5 and project specifications are similar

4 x yes for CMIP5, LUCID, GeoMIP => use CMOR2

2 x yes for CORDEX => ???

0 x yes for obs4MIPs => may be using a sledgehammer to crack nuts

USE CMOR2 ?

Advantage:

- **guarantees a 1.4 CF standard Netcdf file**
- **works as a first Quality Check**
- **automates lots of input and output steps**
- **creates a tracking ID each time it's called**

Disadvantage:

- **was created for CMIP5 and some functionality is not changable yet**
- **the amount of work only pays for bigger projects**

Alternative

Use cdo- or nco- operators in conjunction with ncdump and ncgen to create the netcdf file.

- ncdump & ncgen come with the NetCDF libraries**
- cdo: <https://code.zmaw.de/projects/cdo/>**
- nco: <http://nco.sourceforge.net>**

**Use NetCDF4 libraries with FORTRAN, C or Python.
(own program)**

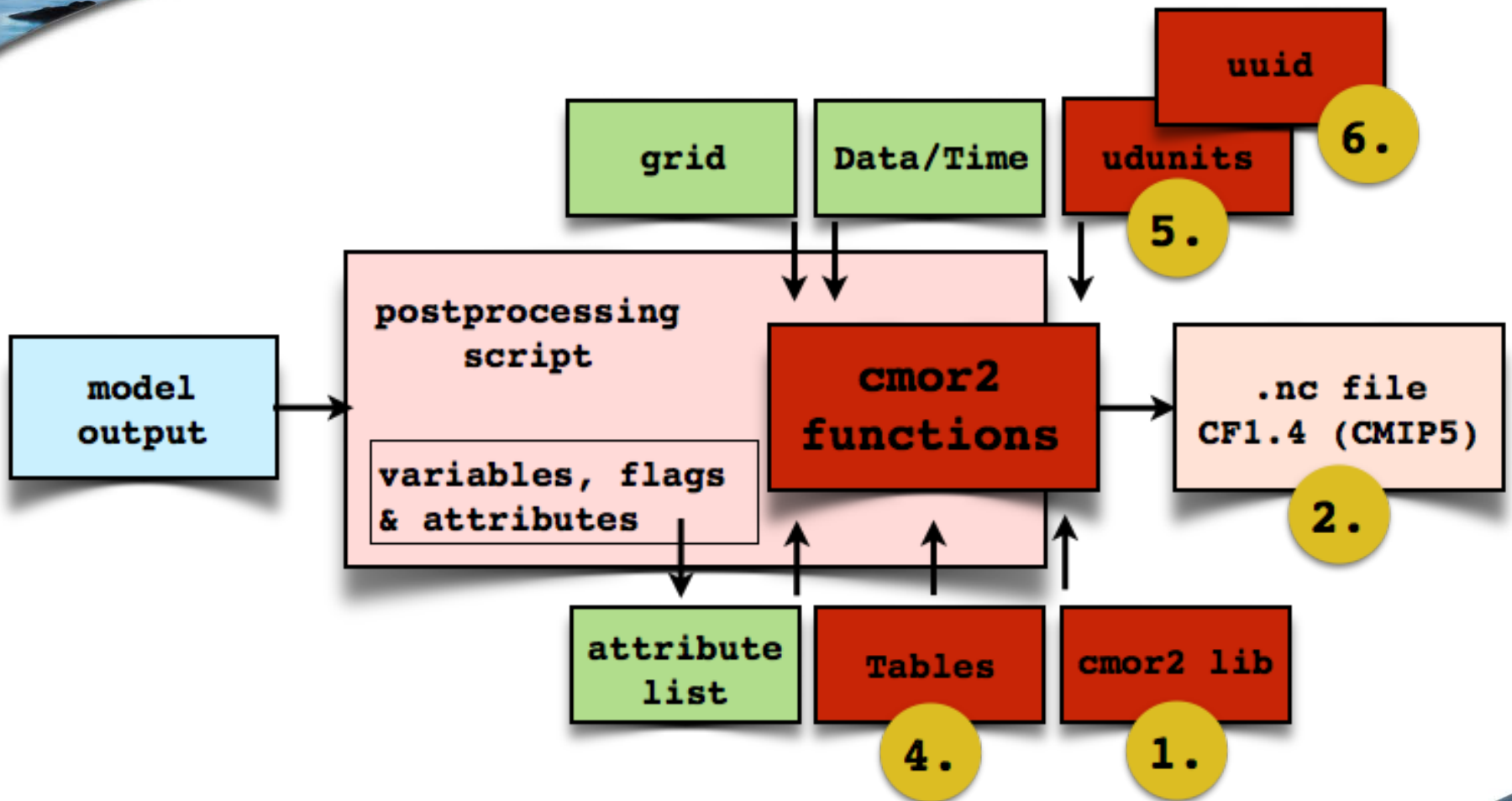
- should be installed on your computer**

CMOR2

Climate Model Output Rewriter (Version 2)

1. a software library which comprises a set of functions to produce **CF1.4** compliant NetCDF files (**CMIP5**).
(FORTRAN, C or Python)
2. based on NetCDF4 libraries
3. was designed for the **CMIP5** project.
4. a set of Tables supports cmor2 with predefined settings for dimensions, variables and attributes.
5. the udunits2 library checks the variable units.
6. the uuid library gives every file a unique identifier.

Workflow



INSTALL & USE CMOR2

1. look for or install the following libraries:

HDF5, NetCDF4, szlib, zlib, udunits, uuid

2. install CMOR2 libraries

3. write main program with CMOR functions (C, Python, FORTRAN)

4. write subroutines to feed the CMOR functions

5. use your well prepared model output for cmor2

OR

5. look for or create script to prepare model output for cmor2

6. get or create project depending experiment/axis/variable-tables

7. tune namelists

8. run script that starts cmor2

9. maybe rename files and directories to fit your requirements

OR

1-9. find someone who makes it for you

Software

CMOR - Climate Model Output Rewriter Software

<http://kitt.llnl.gov/cmor/>

Download for cmor2 via github:

```
git clone git@github.com/PCMDI/cmor.git
```

Tables provided:

```
git clone git@github.com:PCMDI/geomip-cmor-tables.git
```

```
git clone git@github.com:PCMDI/cordex-cmor-tables.git
```

```
git clone git@github.com:PCMDI/lucid-cmor-tables.git
```

```
git clone git@github.com:PCMDI/pmip3-cmor-tables.git
```

```
git clone git@github.com:PCMDI/cmip5-cmor-tables.git
```

**Thanks for listening.
Questions?**