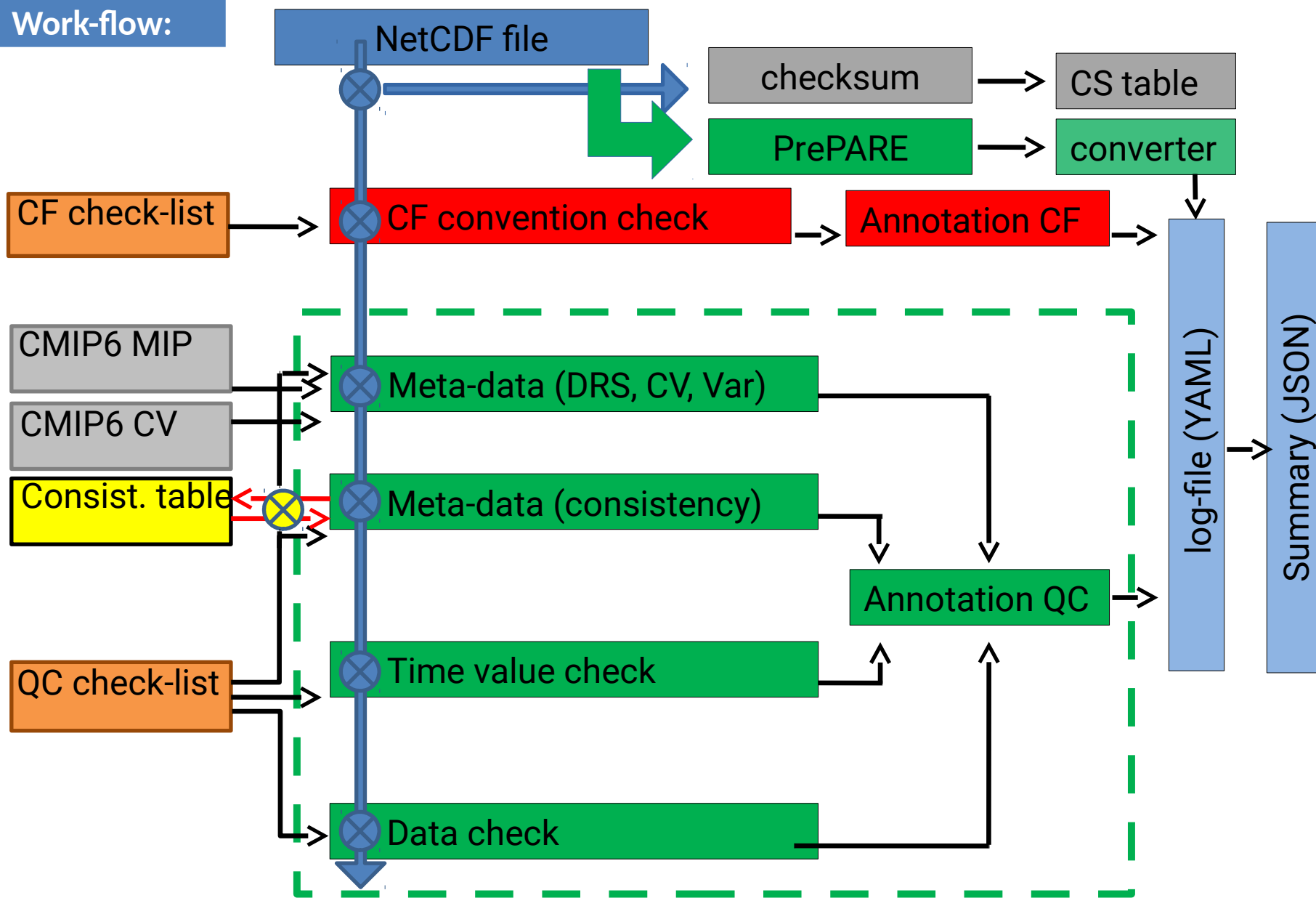


# QA-DKRZ

## Heinz-Dieter Hollweg

DKRZ, [hollweg@dkrz.de](mailto:hollweg@dkrz.de)

# Work-flow:



## Changes

- Separation of installation / operation:  
operation: bash → python; installer: bash; core: C++
- Freeze installation (multiple users)
- User-defined table-modifications (*PROJECT\_AS*)
- Log-file: shorter caption, additional info
- Summary: same annotations for different variables are grouped (*stand-alone python*)
- QA for a stream of files (in preparation)
- Shipping to a computer without internet (in prep.)
- Updated: <https://readthedocs.org/projects/qa-dkrz>

# Structure of QA-Results: Files and Directories

## **check\_logs** (root-directory)

### **log-files** (**files**: DRS-based name.log, YAML)

entry for each checked file; possibly with annotations.  
Output of PrePARE is converted and incorporated.

### **Period** (**files**: DRS-based-name.period, YAML)

time range of atomic variables. If too short, then marked.

### **Annotations** (**files**: unique DRS-based-name.json, JSON)

extracted from a log-file → summary.

## Log-file (YAML)

*Preamble ...*

*qa-revision: master-eee3335|CF\_STD\_NAME:v47,19September2017|CMIP6\_CVs:master-10da227|CMIP6\_MIP\_tables.xlsx:2017-10-04T18:58:03|CMOR:cmor-3.2.7-np112py27\_0|cmip6-cmor-tables:master-fd640d8*

*items:*

- *date:* 2017-11-30T13:05:50

*file:* tas\_Amon\_1pctCO2\_MPI-ESM-LR\_r1i1p1f2\_gn\_200601-210012.nc

*data\_path:* /hdh/hdh/test

*result\_path:* /tmp/QA\_Results/test

*conclusion:* CF: FAIL, CV: disabled, DATA: disabled, DRS: FAIL, TIME: PASS

*events:*

- *event:*

*caption:* DRS CV filename: check failed.

*impact:* L2

*tag:* 1\_2

*text:*

- Found source\_id=<1pctCO2>, expected <MPI-ESM-LR>

*... to be continued*

## Log-file (YAML)

... *continued*

- event:

**caption:** CMOR Error: The attribute <Conventions> could not be validated.

**text:**

- The current input value is <CF-1.7> which is not valid Valid values must match the
- regular expression: [<^CF-1.7 CMIP-6.2\ ( UGRID-1.0\)\{0,\}\$> ...] Check your Control
- Vocabulary file </hdh/hdh/QA-DKRZ/test/test.CMIP6/tables/cmip6-cmor-  
tables/Tables/CMIP6\_CV.json>.

- event:

**caption:** CMOR Error: The attribute <data\_specs\_version> could not be validated.

**text:**

- The current input value is <1.0.0> which is not valid Valid values must match the
- regular expression: [<^[[[:digit:]]\{2,2\}\. [[[:digit:]]\{2,2\}\. [[[:digit:]]\{2,2\}\$>
- ...] Check your Control Vocabulary file
- </hdh/hdh/QA-DKRZ/test/test.CMIP6/tables/cmip6-cmor-

tables/Tables/CMIP6\_CV.json>.

**status:** 1

## Summary (JSON)

```
{
  "QA_conclusion": "FAIL",
  "project": "CORDEX",
  "data_path": [ "/work/kd0956/CORDEX/data/cordex/output" ],
  "selection": [ "EAS-44/CLMcom/MPI-M-MPI-ESM-LR./*/rcp45" ],
  "DRS_0": "cordex",
  "DRS_1": "output",
  "DRS_2": "EAS-44",
  "DRS_3": "CLMcom",
  "DRS_4": "MPI-M-MPI-ESM-LR",
  "DRS_5": "rcp45",
  "DRS_6": "r1i1p1",
  "DRS_7": "CLMcom-CCLM5-0-2",
  "DRS_8": "v1",
  "DRS_9": "*",
  "DRS_10": "*",
  "DRS_11": "v20170914",
```

*to be continued ...*

## Summary (JSON)

... continued

"annotation":

```
[
  {
    "DRS_9": [ "mon", "day" ],
    "DRS_10": [ "mrros", "va850" ],
    "caption": "Variable <DRS_10>: Suspected outlier for record-wide maxima.",
    "severity": "L1"
  },
  {
    "DRS_9": [ "mon" ],
    "DRS_10": [ "rsus", "clt", "mrro", "pr" ],
    "caption": "Variable <DRS_10>: Suspected outlier for record-wide minima.",
    "severity": "L1"
  }
]
```



## QA-DKRZ Status

		CMIP5	CORDEX	CMIP6	Comment	
Conv	CF	v1.4	v1.4	v1.7	www.cfconventions.org	
	UGRID	-	-	v1.0	ugrid-conventions.github.io	
DRS	(Path)					
	(File)					
CV		1)			1) CMOR guide → machine read.	
Var. Requir.				2) frozen	2) CMIP6_MIP_tables.xlsx	
Consistency					atom. variable	experiment
Time					continuous	non-continuous
Data					NaN, Inf, replications, outlier	
CMOR		-	-	PrePARE	http://cmor.llnl.gov	
WPS						
OpenDAP						

## Scaling

- Meta-data checks: to the **number of files**
- Files related to CF Conventions Chapter 9  
(*e.g. simulation of satellite tracks or atmospheric curtains*):  
to the **volume of data**

## NetCDF file: open/close

- a) temporal alignment check of atomic variables (QA-DKRZ)
- b) file check (by QA-DKRZ C++ executable)
- c) PrePARE
- d) PrePARE calling the CF Checker ?

An annotation was produced for a File:

- A) **What to do with annotations?**
- B) **How to inform the ESGF end-user?**
- C) **When to reject a file?**
- D) **When to reject an experiment?**

**Who decides?**