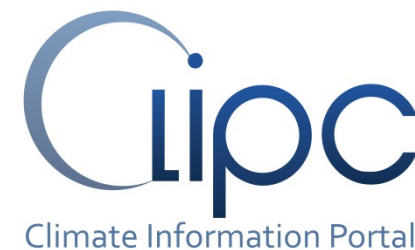




Science & Technology  
Facilities Council



# NetCDF metadata standards for climate model intercomparison

Martin Jukes

Center for Environmental Data Archival

[Martin.jukes@stfc.ac.uk](mailto:Martin.jukes@stfc.ac.uk)



# Climate Model Inter-comparisons

- Initiated by Karl Taylor (PCMDI) with the “Atmospheric model inter-comparison project” (AMIP) [1990];
- Progressed through Climate Model Inter-comparison Project, phases 1 to 5;
- Governed by WCRP/WGCM – with community discussion to resolve technical details and IPCC injecting requirements at various stages;
- Data standards considered as essential for the success of the inter-comparison projects.



# Terminology

## *Aristotle vs. IPCC*



From the perspective of Aristotle, the convective cloud includes “Earth” (the cloud condensation nuclei), “Fire” (the energy needed to generate convection) and “Water” – the IPCC community see clouds as part of the atmosphere.

*"Some of the vapour that is formed by day does not rise high because the ratio of the fire that is raising it to the water that is being raised is small" Aristotle (via Arabic (8<sup>th</sup> century) and Latin translators).*

# Terminology

## *Aristotle vs. IPCC*



The subtle differences between the **substance-based** approach of Aristotle and the **process-based** approach of the IPCC illustrates how common words can have different meanings to different groups.

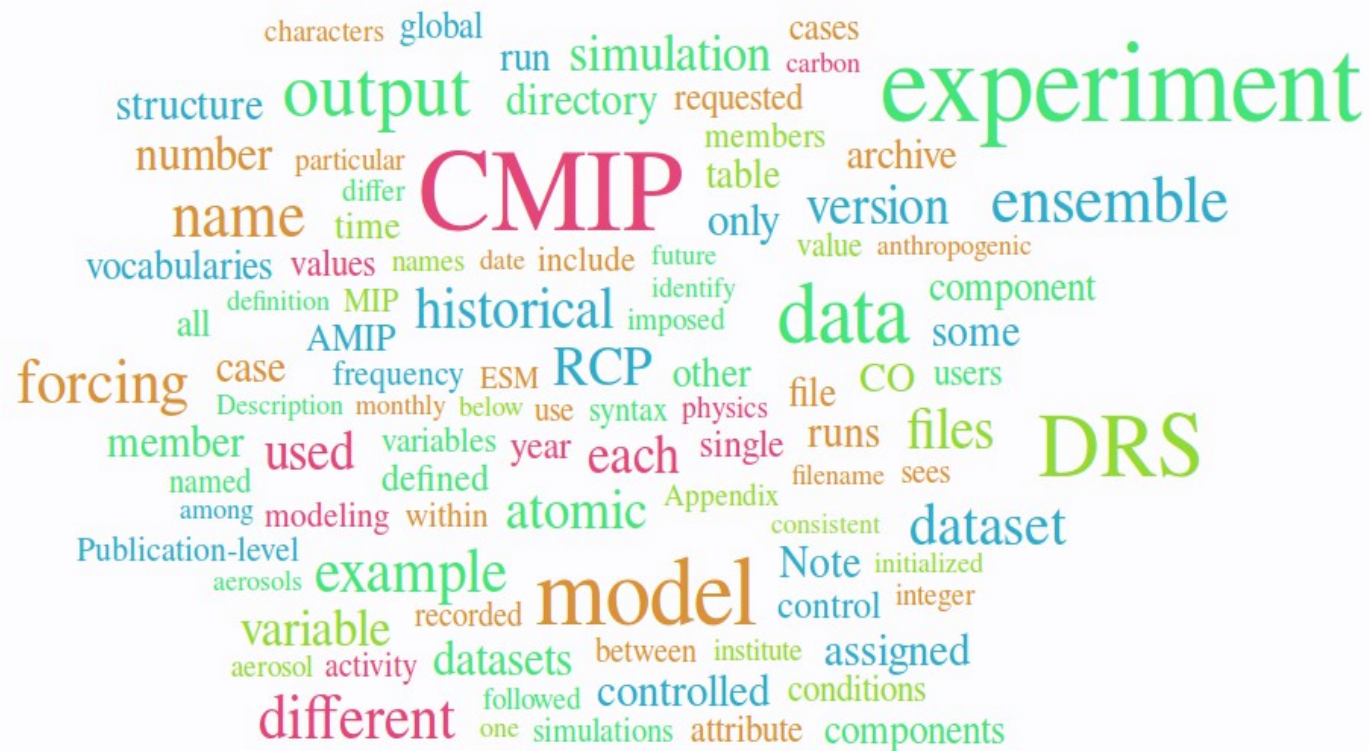


# Data standards for MIPs

- 1) File format: NetCDF
- 2) Core standards: NetCDF CF Convention
- 3) Project specifications: Model Inter-comparison Project request and requirements







## CMIP5 data specification:

- Data request;
- Output format and metadata requirements;
- Data reference syntax.

# Governance and process

MIP requirements are established through a broad community discussion – but it doesn't always include the people who have to implement the agreed rules.



Image: jerryfahrni.com



## Infrastructure Panel

In May 2014 the WCRP Working Group on Coupled Models established the WGCM Infrastructure Panel to “promote a robust and sustainable global data infrastructure in support of the scientific mission of the WGCM” by formulating “achievable goals for [a] global data infrastructure.”

Papers in preparation:

- Data citations;
- Licensing and Access Control;
- MIP checklist;
- CMIP6: projected data volumes.





# Elements of the data request: Generic issues

- Coordinates: specific variable names for all coordinates; rules on implementation of CF Convention;
- One data-variable per file;
- Ancillary information about the grid in a separate file (with an exception to CF Convention rules);
- Generic pattern for file naming;
- Generic global attributes in each file.

# Elements of the data request: for each MIP

- Technical specifications of each experiment;
- A set of tables, defining variables to be archived:  
variable names and long names, associated with a  
CF standard\_name; instantaneous or time mean  
values; treatment of missing values;
- Specific values for some coordinates;
- Vocabularies for all the terms used to define the data  
(e.g. “model”, “experiment”, “institute”). In some  
cases these aim to be consistent between MIPs.

# SWOT

## Strengths

- Inherits portability from NetCDF and robust treatment of coordinates from CF Conventions;
- Flexibility to adapt to needs of specialist groups (e.g. specific terms for “cloud forcing” analysis in CMP5).

## Opportunities

- Expansion of climate modelling community provides critical mass for maintenance of associated software libraries;
- WGCM Infrastructure Panel gives greater transparency to process of creating specifications.

## Weaknesses

- Community process for establishing standards has no clear deadlines;
- Tools to support standards are not generally available;
- Community discussions lead to vague requirements which can not be objectively verified.

## Threats

- Expansion of climate modelling community leads to a fragmentation of standards.