A Registry for Data Specifications in Earth System Modelling

Martin Juckes[1]

[1]Science and Technology Facilities Council, Oxfordshire, UK

Correspondence to: Martin Juckes (martin.juckes@stfc.ac.uk)

Abstract. Establishing a registry to hold specifications of data variables and file formats for Earth System Models. Based on principles established in the CMIP6 Data Request, designed to support a wider community.

1 Introduction

2 Context

The CMIP6 Data Request specifies thousands of variables which are requested to support analysis of CMIP6 experiments.

The specifcations are provided in a technical document together with supporting software and web interfaces.

The level of standardisation of data products within CMIP enables relatively easy use of the data by a broad community.

This paper describes the creation of a sustainable registry, which is designed to enable the registration of technical specifications of variables by a broader community and also to streamline the process for CMIP.

The framework for registry organisational design set out in the ISO 19135 standard for geospatial registries is adopted. This framework facilitates a clear separation of roles enabling efficient operation of the registry to support both the diverse requirements of the scientific domain and the diverse requirements of the science programme managers and governance bodies.

2.1 Challenges

The principle challenge comes from the need to support projects involving consortia of consortia (CCPs). Large consortium projects may already be challenging when it is necessary to align practises of teams from many institutions, and the teams in question may have many competing priorities outside the consortium. With a CCP, such as CMIP, the problems become far more challenging. Consortia may go through a long process to reach an agreement on some topic, only to find that an independent consortium has adopted a conflicting approach.

3 Approach

3.1 Who is being supported

The registry is primarily targetted at the Consortia of Consortia challenge, dealing with projects such as CMIP and CORDEX.

The **domain** supported by the registry is research in the climate of the Earth system. Within this domain, there are communities establised around programmes such as CMIP and CORDEX.

Within these communities there are multiple consortia, such as the MIPs in CMIP or the regional coordination groups in CORDEX.

The key point about these communities, which is essential for the functioning of the registry, is the existence of a governance body which sets framework requirements for the information to be stored.

3.2 What is being managed

A registry, following ISO19135, is considered as an organisation hosting registers. Each register is a collection of registered items of information, submitted by the scientific community. The register functions as a shared shared reference resource supporting the aims identified by the governance group associated with the register.

The ability of the registry to support a range of registers with different governance groups is of critical importance. This ability follows naturally from the ISO standard.

Different registers may have different submission criteria. Some may contain items which register information specific to the group submitting the information. In this case, the ap-

proval process need only look at the technical compliance of the information provided.

In other cases, the items are intended to be a shared resource. When a new item is proposed, there needs to be a community wide discussion to ensure that any addition is supported across the community.

3.3 The Governance Framework

The

3.4 The Operational Framework

The operation of the registry is designed to provide a flexible platform to maintain and develop the registers.

4 Summary and Outlook

4.1 Accessing the Registry

The current version of the DREQ is available from the project website: w3id.org/cmip6dr under the MIT License (BSD). It is provided a versioned XML document, which can be used directly or programmatically (both command line tools and a python library are provided). The exact version of the DREQ discussed in this paper (01.00.31) is available as a package from the Python Software Foundation at pypi.org/project/dreqPy/1.0.31/.

4.2 Challenges Arising

Acknowledgements. M.N. Juckes is funded by the UK National Centre for Atmospheric Science and by EU H2020 project IS-ENES3 (824084).

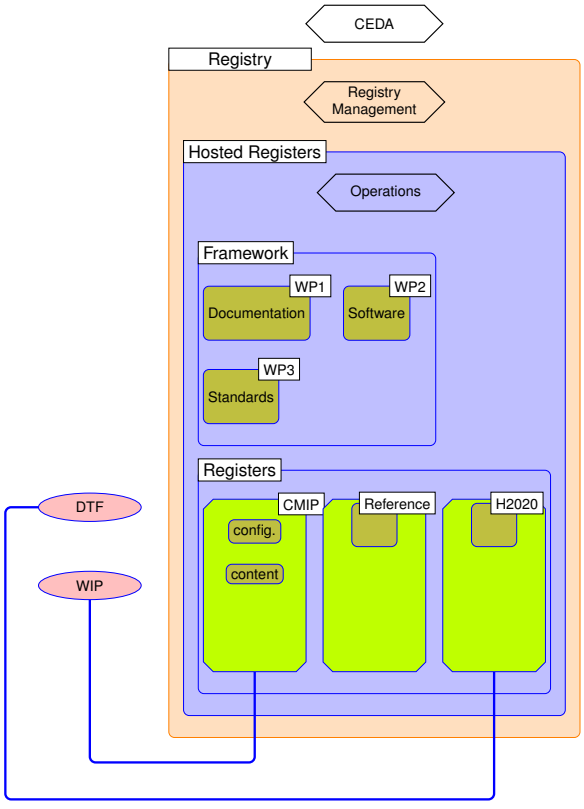

Figure 1. Organisational chart for the registry.