# ESA CCI Portal Requirements from CLIPC

Martin Juckes, Jan. 29th, 2014.

## Background

The FP7 project Climate Information Portal for Copernicus (CLIPC) is developing a pre-operational data service infrastructure for the Copernicus Climate Service.

The vision is of a federated service infrastructure, with data held in distributed data stores. Efficient operation of such a service will depend on maintenance of appropriate standard. A file format specification for the ESA CCI data has already been developed[1].

The CLIPC and the Copernicus Climate Service will make use of high level gridded products from the ESA Climate Change Initiative. In some cases additional products (e.g. swathe data) may be needed , and the requirements expressed below apply to such products.

We assume that the "CCI Portal" will comprise both a web portal and an underlying data infrastructure.

## File Format

Compliance with the CCI data format requirements should be verified.

- Data formatting errors which may compromise user access services should be fixed.
- Where data formatting errors which are considered sub-critical, data may be published with annotation of those errors, subject to agreement with appropriate oversight body.

## Data Discovery Services

Data discovery services should support an interface that allows remote clients to submit searches through a programmatic interface.[2] The vocabularies used to organise the data (e.g. satellite names, processing methods) should be clearly defined in both human and machine readable formats.

## Access Control

The data should be accessible for users using their Copernicus climate service accounts, and preferably for users registered with other ESGF service providers.

## User support

A single help desk must provide a point of access to support for all data products, with queries submitted by email or web form. The help desk should use a triaging system to distribute queries to appropriate experts and monitor responses to ensure that these are delivered in an appropriate time frame.

## Curation

The service must ensure that all high level data products are backed up with a robust recovery mechanism.

The project must prepare a plan for the long term curation of the high level data in an accessible archive, e.g. through transition to a relevant archive with appropriate expertise and long term funding.

---

1 http://46.137.76.174/sites/default/files/CCI_Guidelines_Iss4.2_May2013_1.pdf

2 "programmatic interface": an interface allowing direct use by 3rd party software, i.e. an Application Programming Interface (API) allowing direct

## Processing Services

A range of processing options should be available to allow users and remote clients to extract information without necessarily downloading complete files. e.g. subsets, maxima and minima (also of subsets), other simple statistics and statistics of differences between variables. These services should be accessible through a programmatic interface[2].

## Virtual Data Lab (desirable)

A cloud environment in which users can be provided with direct access to computational resources, e.g. through “Platform as a Service (PaaS)” or “Infrastructure as a Service (IaaS)”, with fast access to the data archive over local networks and significant cache space for storage of intermediate products.

## Requirements capture

The CCI portal operators should interact with the CLIPC project and with the Copernicus Climate Service Operator, once the latter is appointed, to ensure that the range of products and services is appropriate.