EQC function of the Copernicus Climate Data Store

Climate Change

ESIP IQC monthly telco

Carlo Lacagnina, Barcelona Supercomputing Center

Climate Change

COPERNICUS CLIMATE CHANGE SERVICE (C3S)

Copernicus is the European Union's Earth observation programme coordinated and managed by the European Commission

Copernicus provides a unified system through which vast amounts of data are fed into a range of ***thematic information services,*** designed to benefit the environment, the way we live, humanitarian needs and support effective policy-making for a more sustainable future

ECMWF opernicus Europe's eyes on Earth European Commission

Climate Change

COPERNICUS CLIMATE CHANGE SERVICE (C3S)

Copernicus is the European Union's Earth observation programme coordinated and managed by the European Commission

Copernicus provides a unified system through which vast amounts of data are fed into a range of ***thematic information services,*** designed to benefit the environment, the way we live, humanitarian needs and support effective policy-making for a more sustainable future

One of the thematic services, namely the ***Copernicus Climate Change service (C3S)***, aims at providing comprehensive information about past, present and future climate to a wide range of users

Target users are policy makers, scientists and business to achieve a more suistanable future

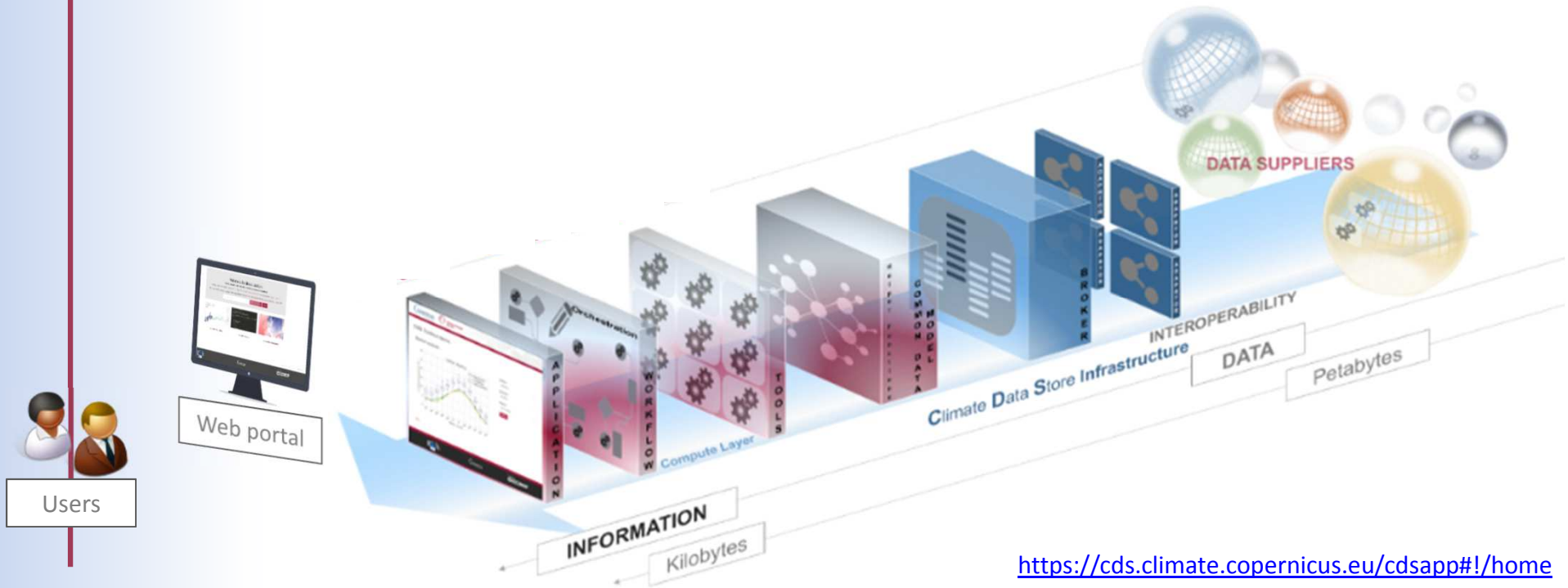
All this information is collected and made available in the ***Climate Data Store (CDS)***, the cornerstone of the C3S

ECMWF opernicus Europe's eyes on Earth European Commission

Climate Change

CLIMATE DATA STORE (CDS)

At the heart of the **C3S** infrastructure is the cloud-based ***Climate Data Store (CDS)***. It provides a single point of access to a wide range of climate datasets, namely satellite and in-situ observations, reanalyses, seasonal forecasts and climate projections



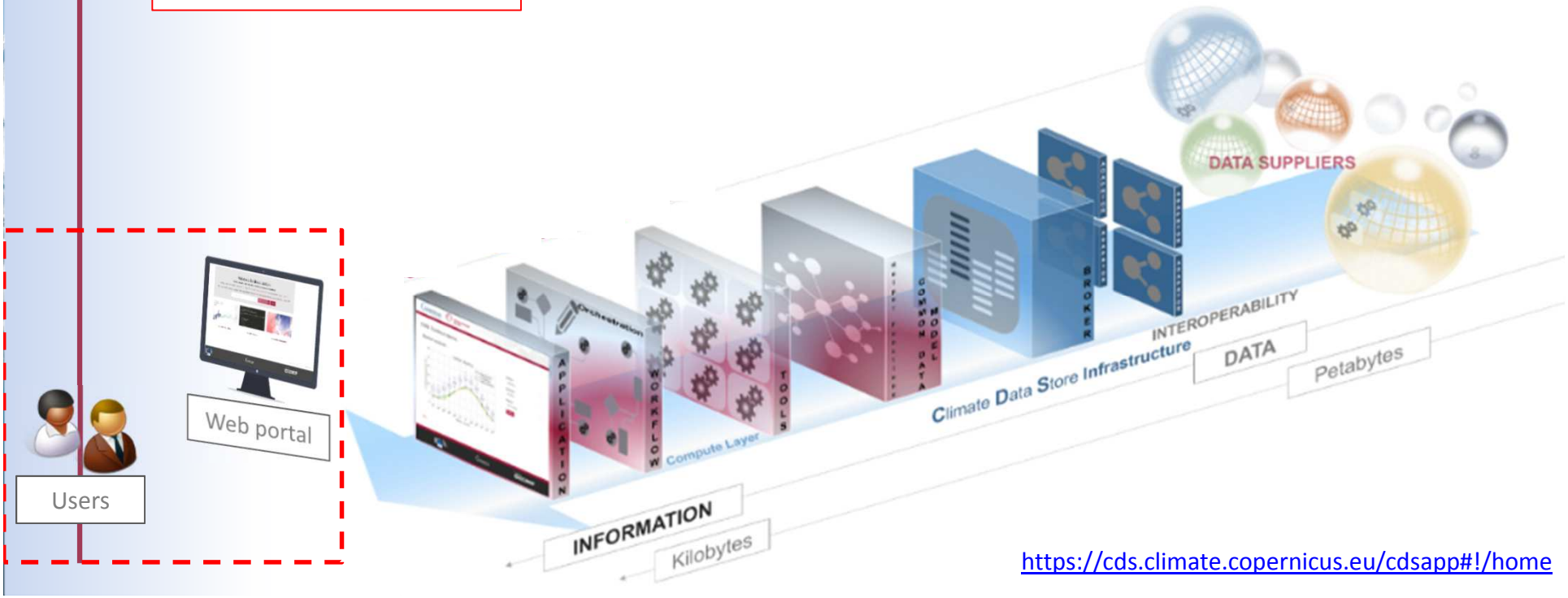
https://cds.climate.copernicus.eu/cdsapp#!/home

CLIMATE DATA STORE (CDS)

Climate Change

The ***Climate Data Store (CDS) of the C3S*** provides a single point of access to a wide range of climate datasets, namely satellite and in-situ observations, reanalyses, seasonal forecasts and climate projections

Users can explore the CDS data through the web portal



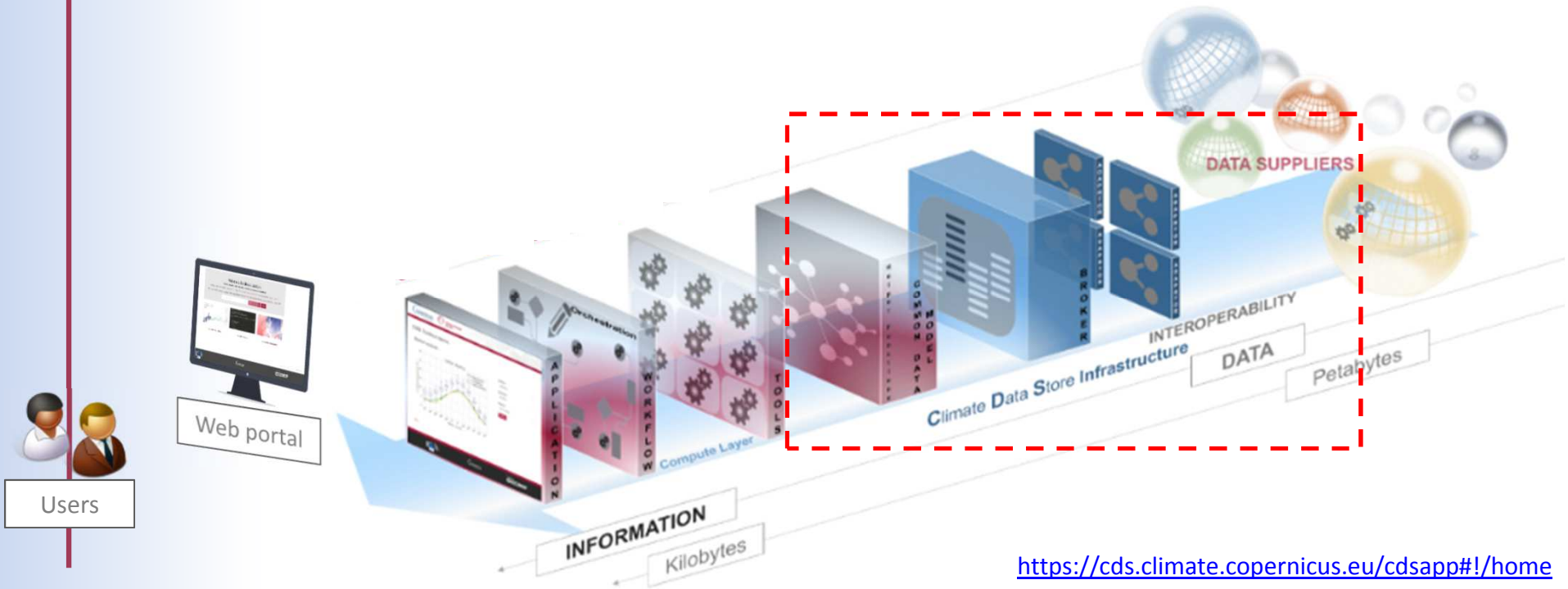
https://cds.climate.copernicus.eu/cdsapp#!/home

CLIMATE DATA STORE (CDS)

Climate Change

The ***Climate Data Store (CDS) of the C3S*** provides a single point of access to a wide range of climate datasets, namely satellite and in-situ observations, reanalyses, seasonal forecasts and climate projections

Data can be downloaded



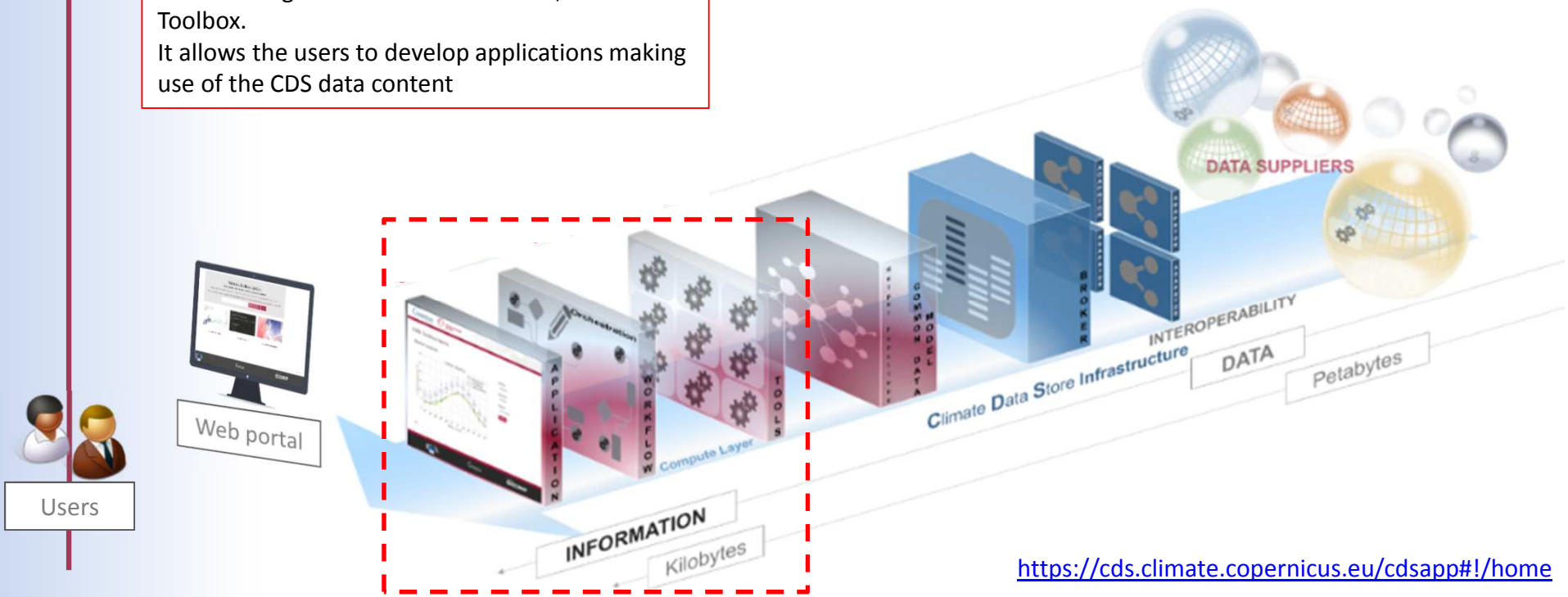
https://cds.climate.copernicus.eu/cdsapp#!/home

CLIMATE DATA STORE (CDS)

Climate Change

The ***Climate Data Store (CDS) of the C3S*** provides a single point of access to a wide range of climate datasets, namely satellite and in-situ observations, reanalyses, seasonal forecasts and climate projections

Data can be explored (e.g. visualized, subsetted) on-line using a suite of software tools, i.e. the Toolbox.
It allows the users to develop applications making use of the CDS data content



https://cds.climate.copernicus.eu/cdsapp#!/home

Climate Change

EVALUATION & QUALITY CONTROL OF THE CDS

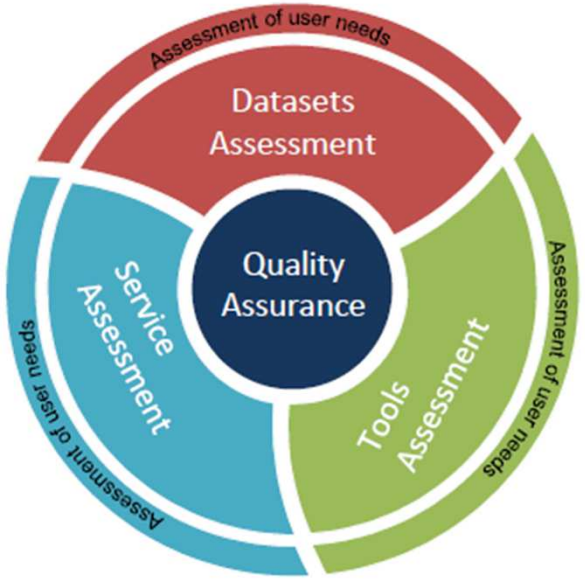
Such a complex infrastructure requires an ***Evaluation and Quality Control (EQC)*** function providing an overarching quality assurance service for the whole CDS:

CDS datasets: provide information about the technical and scientific quality and fitness-for-purpose, along with independent assessment of the datasets

CDS Toolbox: assessment of maturity and fitness for purpose of the software provided to explore the datasets

CDS service: performance assessment of the CDS infrastructure (e.g. speed, responsiveness, system availability)

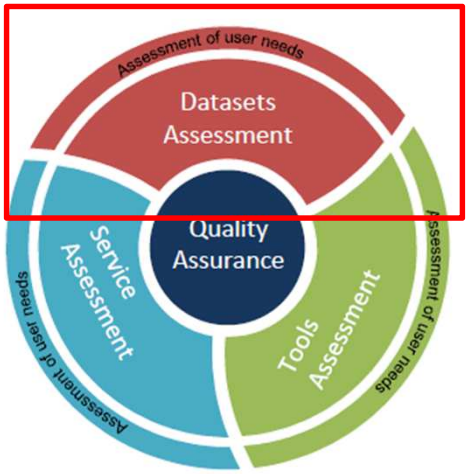
CDS users: user requirement assessment to measure users' satisfaction with the CDS. Map evolving user needs into viable user requirements to ensure a user-oriented evolution of the CDS



ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

EQC OF THE CDS DATASETS



ECMWF Copernicus Europe's eyes on Earth European Commission

EQC OF THE CDS DATASETS

Climate Change

The EQC function builds a workflow to assess the quality of the CDS datasets, leading to the publication of Quality Assurance Reports (QARs)

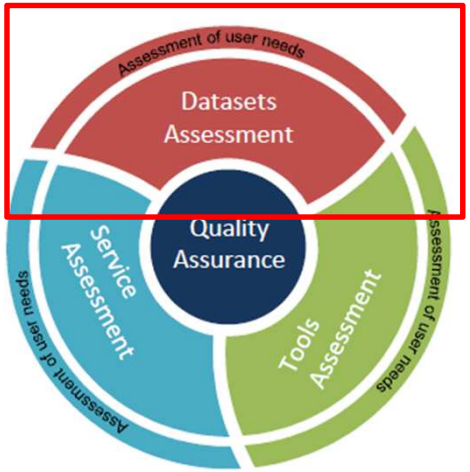
EQC update

Fast assessment
- Compliance with a set of minimum requirements

In-depth assessment
- Includes scientific assessment and maturity matrix

QAR available in the CDS
- The QAR is organized in a table integrated in the CDS web portal, giving access to detailed EQC information according to user selections

User feedback
- User needs are analysed for improvement and expansion of the CDS datasets and related EQC information



ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

EQC OF THE CDS DATASETS

The EQC information is made by dataset documentation according to provider indications and reviewed during the EQC process + an independent assessment conducted by the EQC team

Fast assessment

- Compliance with a set of minimum requirements

- ❑ Documentation: e.g. quantity name, units, format, resolution, provider, version, description of processing, uncertainty characterization
- ❑ Data checker: e.g. space/time completeness of data and metadata, physical ranges of plausability

In-depth assessment

- Includes scientific assessment and maturity matrix

- ❑ Documentation: e.g. quality flags, cloud masking, product traceability chain, validation report, inter-comparison activities
- ❑ Independent assessment: e.g. compliance with international standards, maturity matrix (whether best practises have been followed), fitness for purpose of identified use cases (dataset is robust and sufficient for the user's specific application), performance metrics

Climate Change

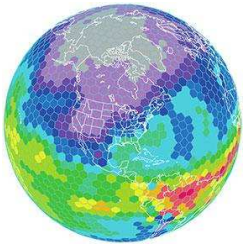
EQC OF THE CDS DATASETS

QAR available in the CDS

- The QAR is organized in a table integrated in the CDS web portal, giving access to detailed EQC information according to user selections

A challenge: the CDS datasets encompass a wide variety of data types:

- ❑ Satellite observations
- ❑ In-situ observations
- ❑ Reanalysis
- ❑ Seasonal forecasts
- ❑ Global and regional climate projections



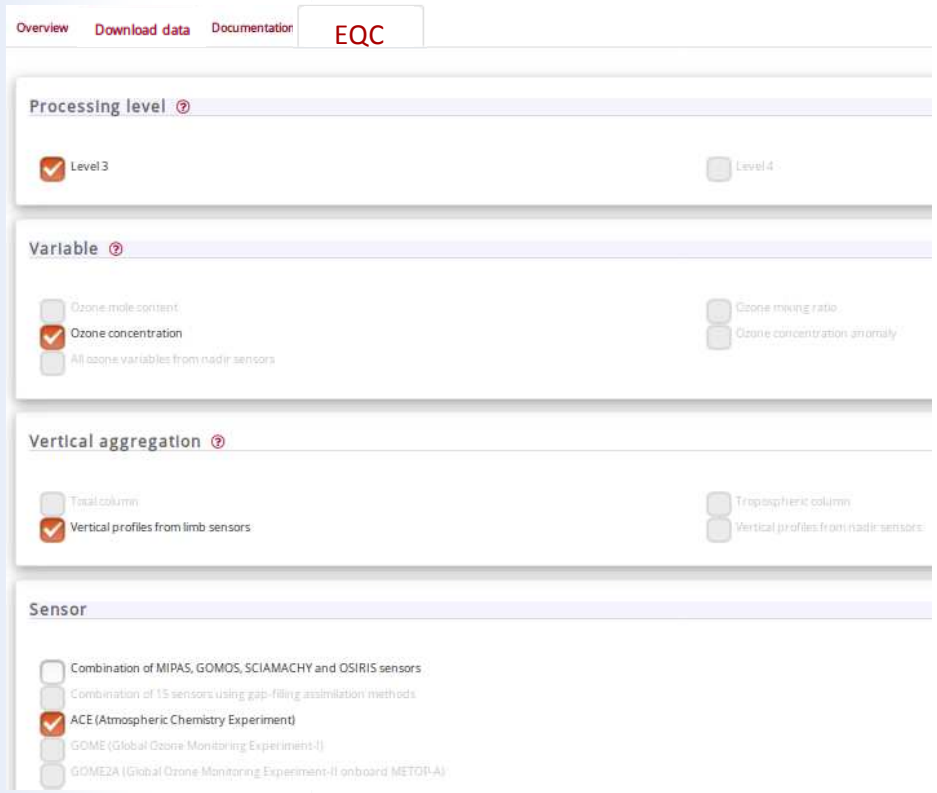
This poses challenges to provide a ***seamless and homogeneous EQC information*** for the whole CDS datasets

To overcome this issue a ***synthesis table*** is integrated in the CDS web portal

ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

CDS PORTAL



- User selects the dataset and variable of interest
- The user has then the option to consult the related EQC information, clicking on the appropriate tab
- A set of information is presented through a web page displaying a synthesis table, the webpage is created dynamically

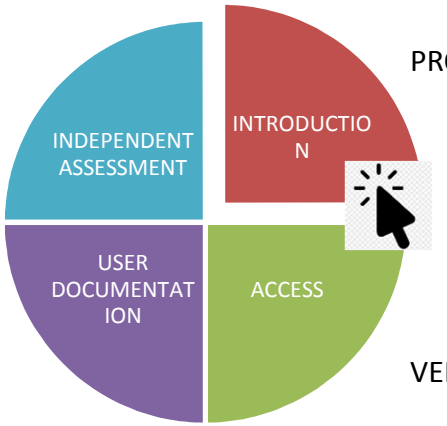
Show EQC information

ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

CDS PORTAL

Overview | Download data | Documentation | EQC



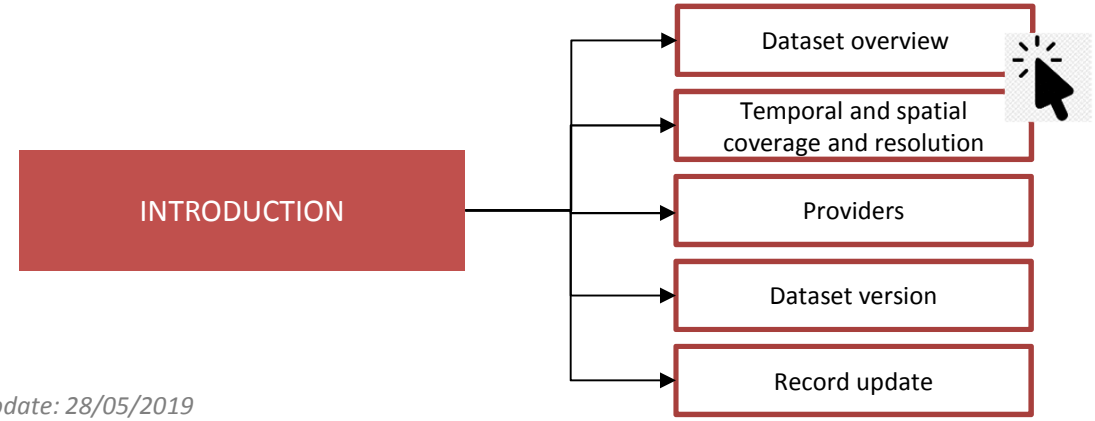
PRODUCT NAME: *Ozone monthly gridded data from 1970 to present*

VARIABLE: *Ozone concentration*

PROCESSING LEVEL: *Level 3*

SENSOR: *ace (atmospheric chemistry experiment)*

VERTICAL AGGREGATION: *vertical profiles from limb sensors*

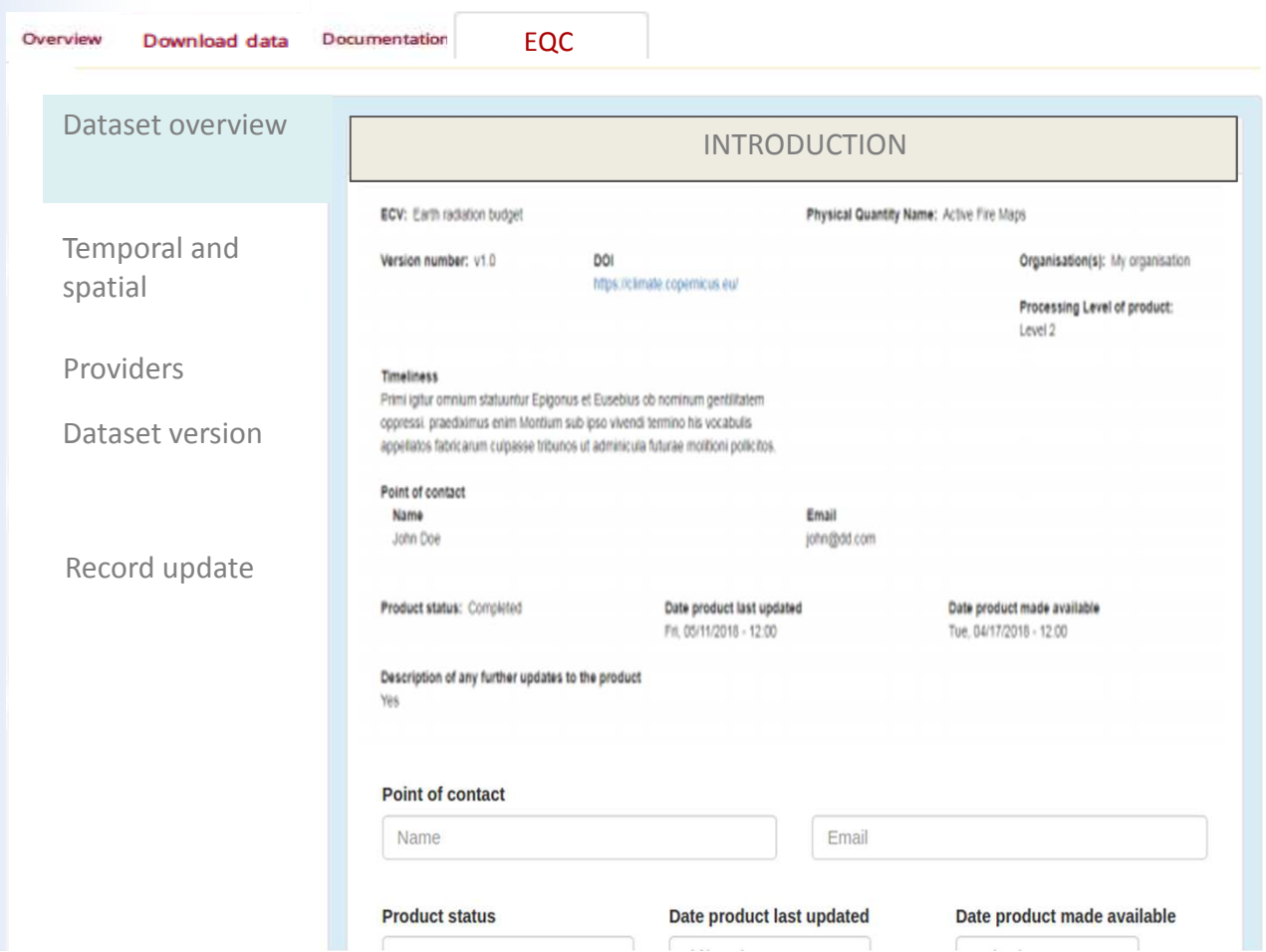


EQC last update: 28/05/2019

- The EQC information is organized and homogenized across all datasets
- The layout is agnostic of the product type selected
- The user clicks on the box of interest and will be directed to another web page (also created dynamically) displaying the specific EQC information

EQC OF THE CDS DATASETS

Climate Change

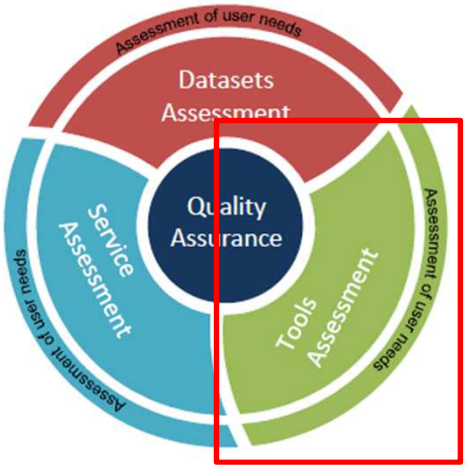


- A web page is built dynamically, showing the information stored and managed by the Content Management System (CMS).
- The user selections form the query to interrogate the SQL database managed by the CMS.

European Commission

Climate Change

EQC OF THE CDS TOOLBOX



ECMWF Copernicus Europe's eyes on Earth European Commission

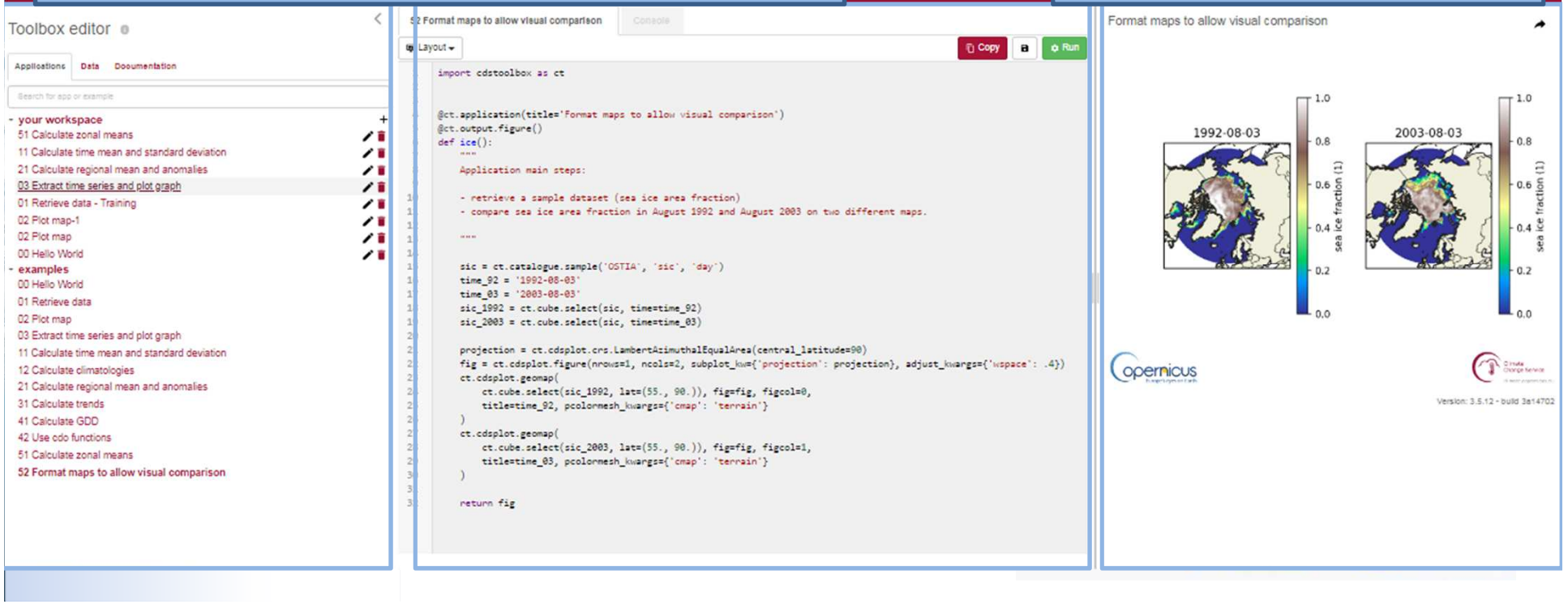
CDS TOOLBOX

Climate Change

The Toolbox is a suite of software tools, which allows the users to explore (e.g. subsetting, averaging, visualizing) the CDS products and develop their own applications based on the CDS data content

Workspace and editor

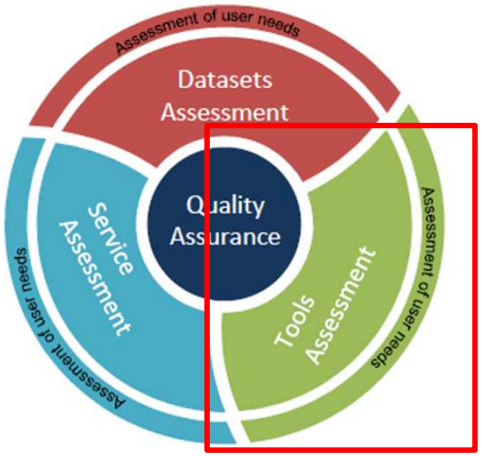
Execution status and Results



Climate Change

EQC OF THE CDS TOOLBOX

The EQC function assesses the quality of the CDS Toolbox from multiple angles: tools, workflows, Common Data Model (CDM), application editor, provenance tracking system. Here we focus on the tools only



In particular, the EQC framework of the Toolbox aims to:

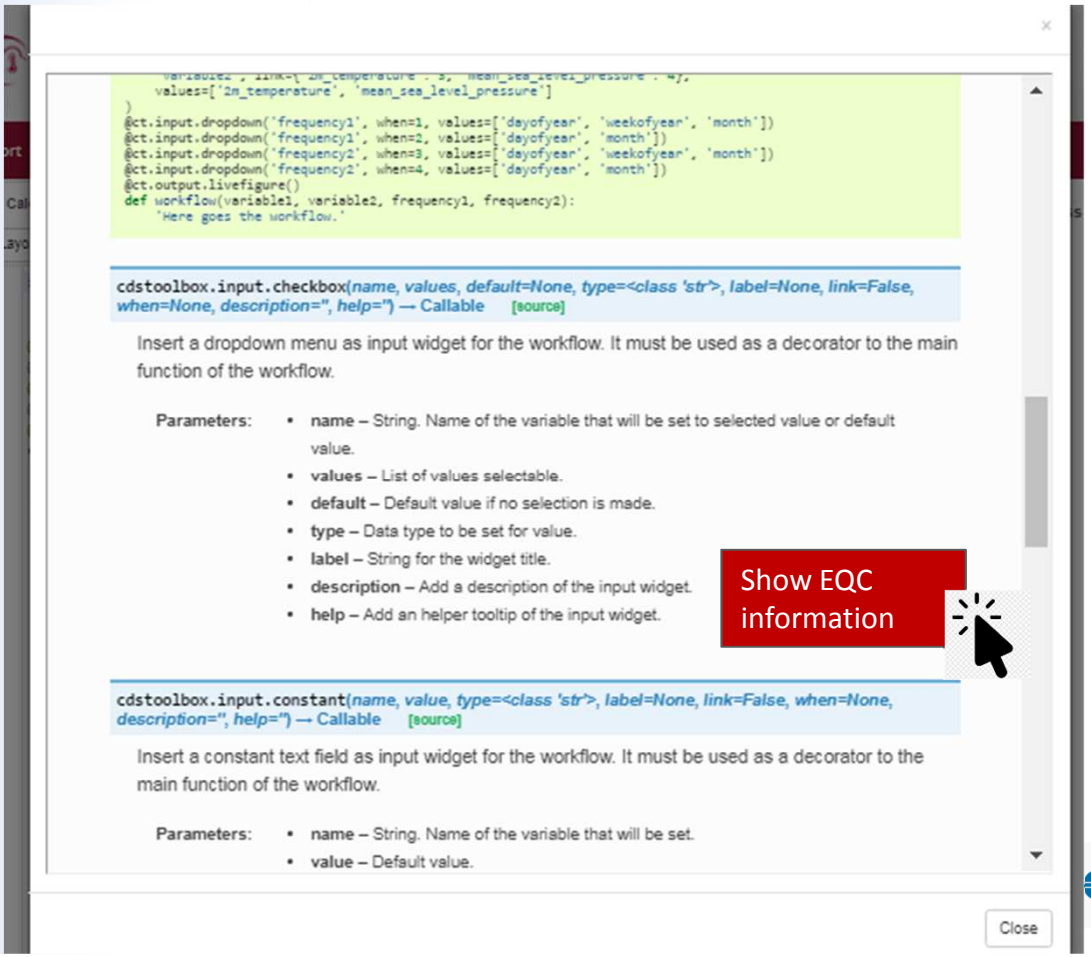
- ❑ Assess the maturity of the software tools: robust in terms of code versioning and testing and well documented
- ❑ Evaluate the fitness-for-purpose of the software through use cases identified together with the users, that checks the applicability of the Toolbox to specific operations.

Software quality assessment is based on the internationally-recognized standard ISO/IEC 9126 and extensions (e.g. ISO/IEC 25010:2011)

ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

EQC OF THE CDS TOOLBOX

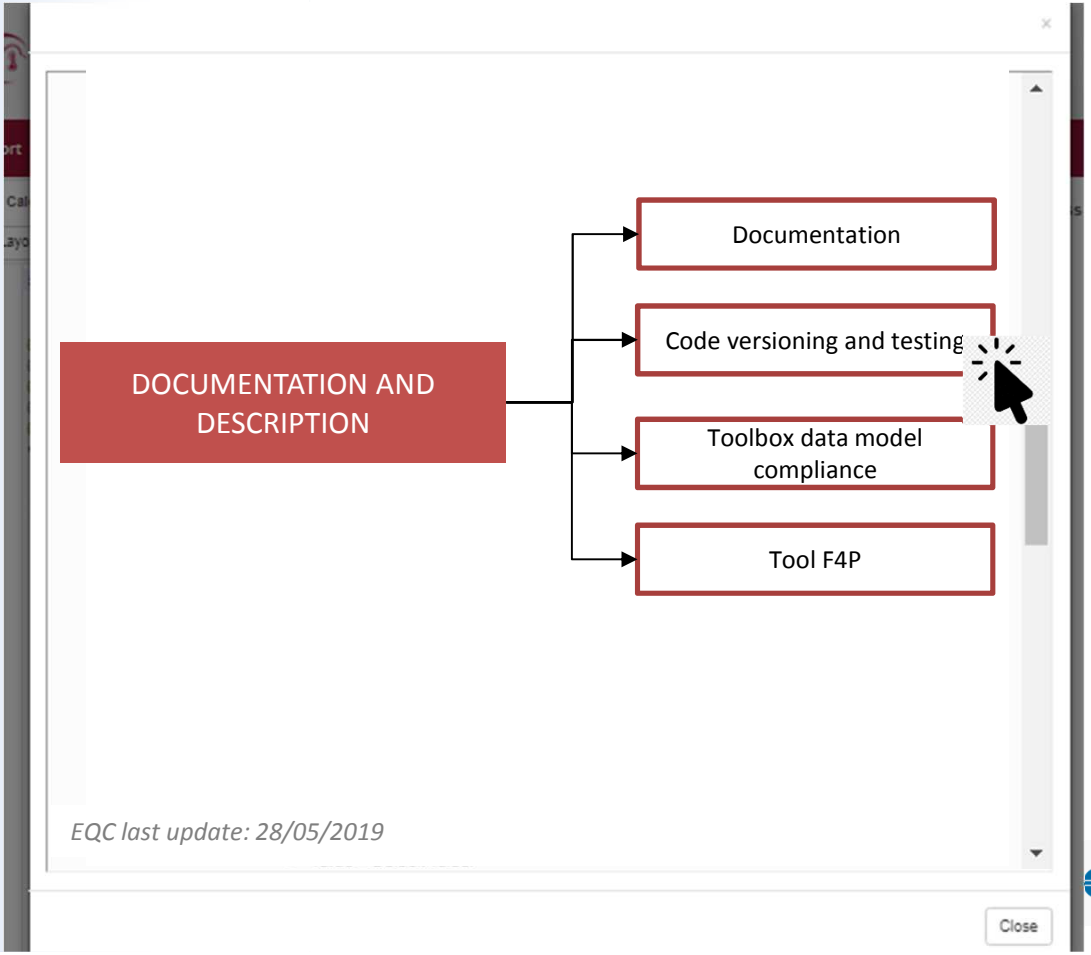


- The user clicks one of the tools, a web-page opens describing breafly the tool function
- The user is offered the possibility to access the details of the related EQC information

ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

EQC OF THE CDS TOOLBOX



- A web page is built dynamically, showing the EQC information organized and homogenized in a synthesis table
- The user clicks on the box of interest and will be directed to another web page (also created dynamically) displaying the specific EQC information

ECMWF Copernicus Europe's eyes on Earth European Commission

EQC OF THE CDS TOOLBOX

Climate Change

- Documentation
- Code versioning and testing
- Toolbox data model compliance
- Tool F4P

DOCUMENTATION AND DESCRIPTION

Is the tool container-based?	Accessibility of the source code	Source code
Yes	The source …	Github.com/ES/…

Independent Evaluation on Code Versioning and Testing and

Description of code testing with maturity matrix

Code	Basic Check	Functionality (Suitability) unit testing and integration testing	Performance through standardized tests
Information Availability (a)			
Information Appropriateness (b) (if script based)			
Information Appropriateness (c) (if container based)			
Functionality 1 (d)			
Functionality N (d)			

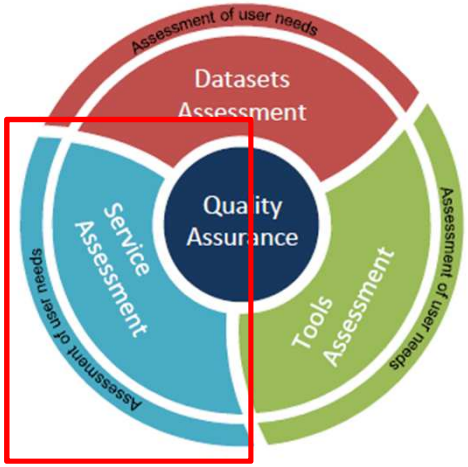
- A web page is built dynamically, showing the information stored and managed by the Content Management System (CMS).
- The user selections form the query to interrogate the SQL database managed by the CMS.

opernicus Europe's eyes on Earth

European Commission

Climate Change

EQC OF THE CDS SERVICE

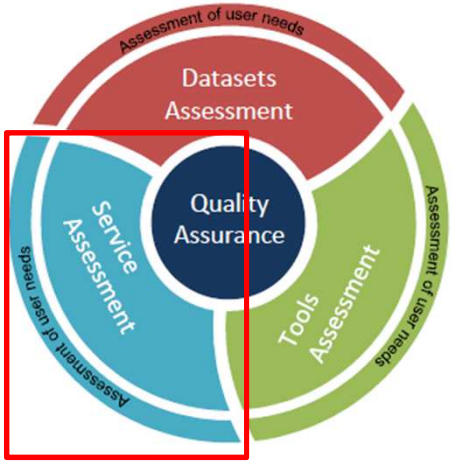


ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

EQC OF THE CDS SERVICE

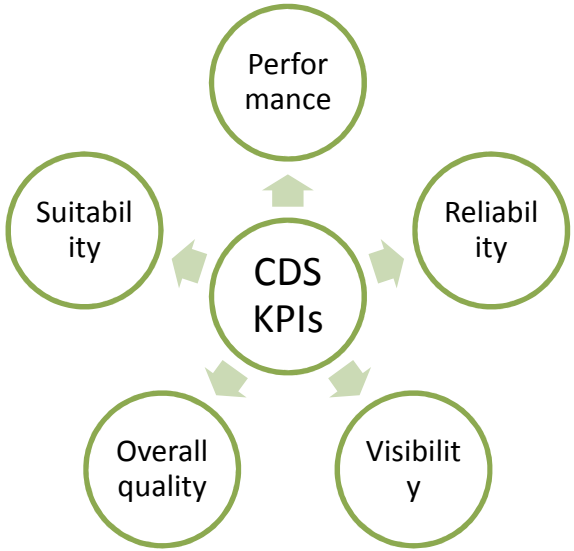
The EQC function measures and reports the technical quality of the CDS service (e.g. system availability, response time)



In particular, monitoring of the CDS infrastructure is based on:

- ❑ A set of Key Performance Indicators (KPIs)
- ❑ On-line rating widgets to monitor user satisfaction
- ❑ Development of a web dashboard hosting the KPIs and widgets statistics

The KPIs have been inspired by the internationally-recognized standard ISO/IEC 25010 and 25011

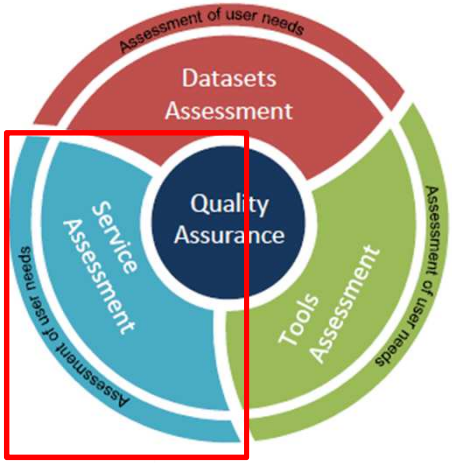


ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

EQC OF THE CDS SERVICE

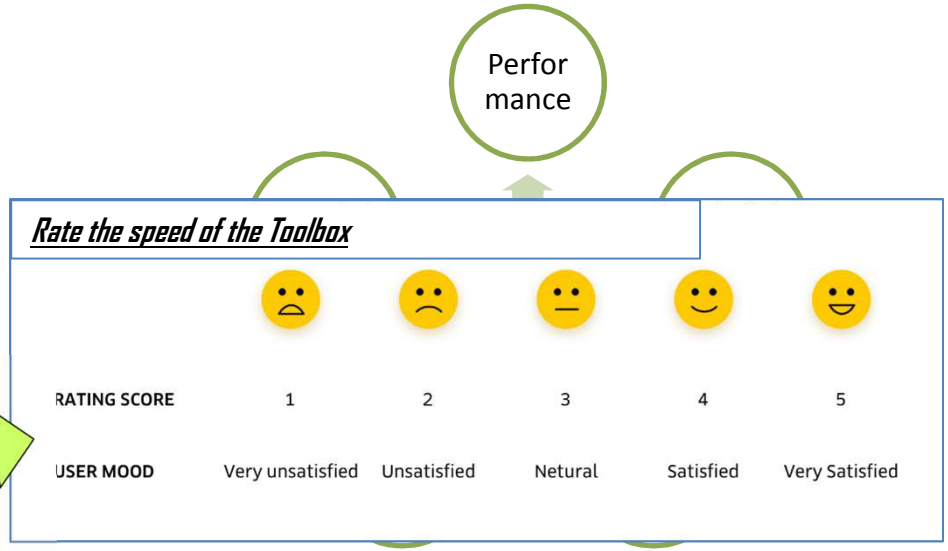
The EQC function measures and reports the technical quality of the CDS service (e.g. system availability, response time)



The KPIs have been inspired by the internationally-recognized standard ISO/IEC 25010 and 25011

In particular, monitoring of the CDS infrastructure is based on:

- ❑ A set of Key Performance Indicators (KPIs)
- ❑ On-line rating widgets to monitor user satisfaction
- ❑ Development of a web dashboard hosting the KPIs and widgets statistics

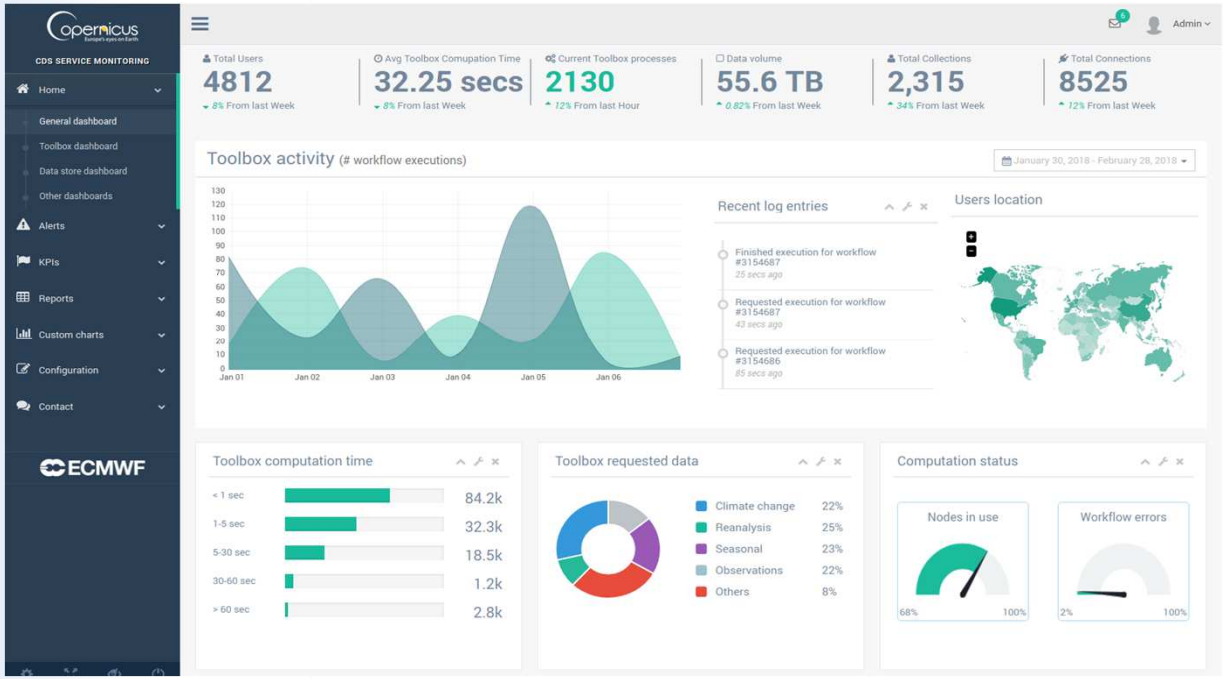


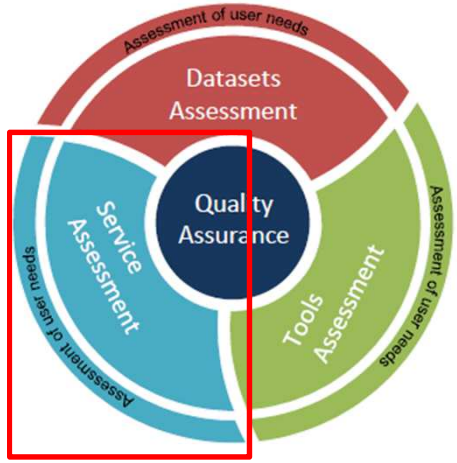
ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

EQC OF THE CDS SERVICE

Preliminary version of the dashboard hosting the KPIs and widgets statistics

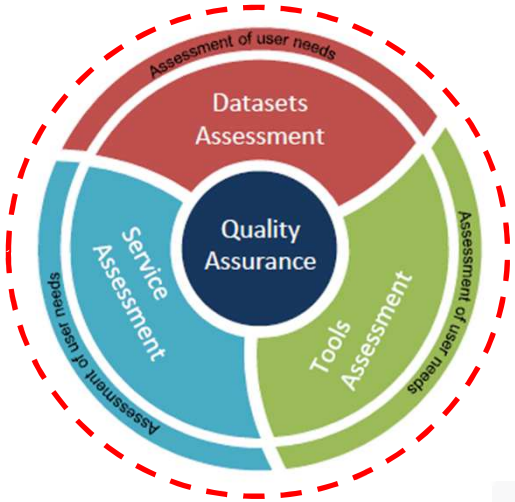




ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

USER NEEDS

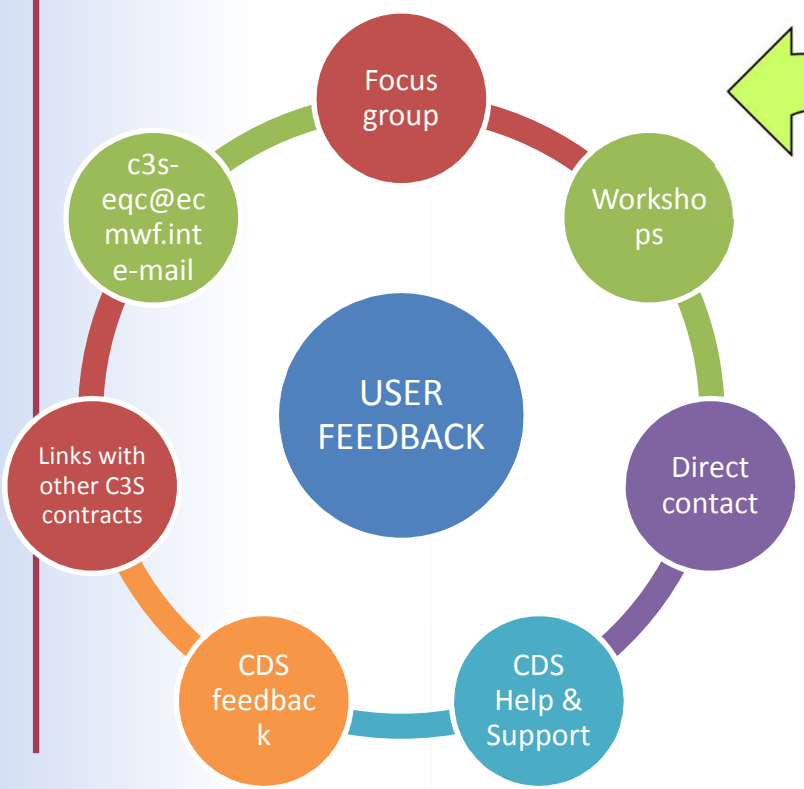


ECMWF Copernicus Europe's eyes on Earth European Commission

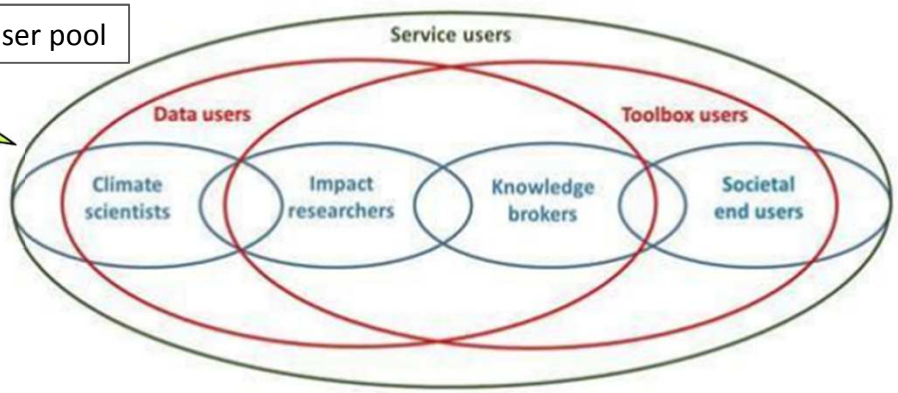
Climate Change

USER NEEDS

As the C3S is a user-driven climate service, user feedback is central to develop recommendations for improvement and expansion of the CDS data, Toolbox and service



User pool



Harvesting user requirements in a database

URDB

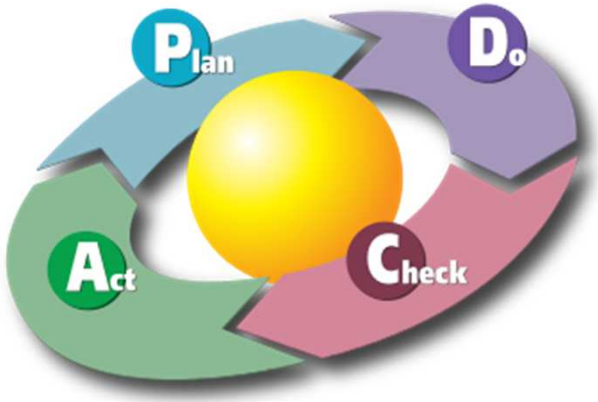
Analysed in a report for further development of the CDS

URAD

Climate Change

MANY COMMONALITIES WITH ESIP IQC

- ❑ Bring together different communities to assess aspects of quality of Earth science data
- ❑ Establish baseline of standards for data quality for adoption by data providers
- ❑ Promoting data management standards and best practices for ensuring high quality datasets for use in climate policy and services
- ❑ Data quality harmonization
- ❑ Users guidance and data conventions are fundamental to the development of the CDS
- ❑ Presenting quality information to users: meet user's needs and avoid data are misused.
- ❑ Maturity matrix and promotion of an holistic approach
- ❑ Sharing experience and lessons learnt about long-term management, preservation and curation of Earth system data



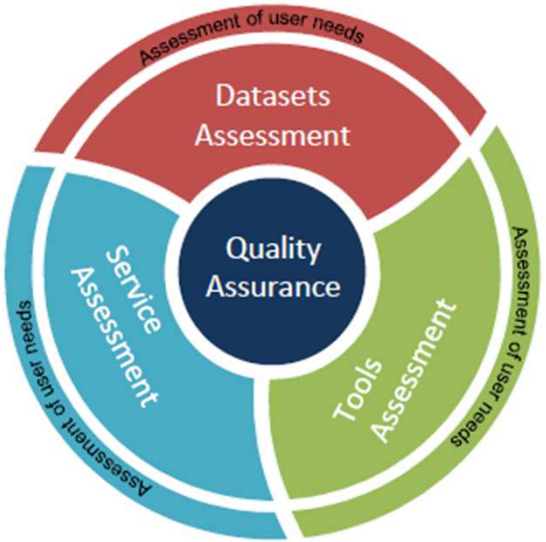
Deming cycle as suggested in Peng et al. (2018)

ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

CONCLUSIONS

- ❑ The EQC function of the CDS provides an **overarching quality assurance service for the whole CDS**: dataset, Toolbox, infrastructure, users
- ❑ The CDS contains a **wide variety of datasets**: satellite and in-situ observation, reanalysis, seasonal forecasts, global and regional climate projections
- ❑ A suite of software tools is available to explore and visualize datasets (**Toolbox**)
- ❑ Users can fully understand status and purpose of data product, with **all relevant information in one place**, based on the matters the user deems most important
- ❑ The **homogenization of the EQC information** across all datasets allows to directly compare several different products
- ❑ The EQC function provides **guidance to data producers** on the information they need to deliver and to be compliant
- ❑ **Users are central** for the development and expansion of the CDS
- ❑ **ESIP IQC** seems an appropriate place for the discussions on EQC aspects.



ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

SOME OPEN QUESTIONS…

- ❑ What type of information of the datasets are the users interested in?
- ❑ Missing a list of minimum requirements to assure a minimum quality of the datasets.
- ❑ What is the level of granularity expected by users?
- ❑ What is the frequency of updating EQC information? What can be automatized?
- ❑ How to devise a scoring system informing about the level of dataset information available?
- ❑ Lack of ISO standards for climate data and metadata quality standards.
- ❑ Are there any maturity levels defined for an EQC process?
- ❑ Are there maturity matrix for modelling output datasets?
- ❑ Are there international standards similar to GCOS that could be applied to modelling datasets?
- ❑ Gaps in standardisation of terminology, processes and practices due to a lack of an appropriate regulatory framework

ECMWF Copernicus Europe's eyes on Earth European Commission

Climate Change

THANK YOU

C3S_512 is a service contract funded by Copernicus Programmes and operated by ECMWF

carlo.lacagnina@bsc.es

Copernicus Europe's eyes on Earth | European Commission | ECMWF

Climate Change

ACKOWLEDGMENTS

The research leading to these results is part of the Copernicus Climate Change Service (C3S) (Framework Agreement C3S_512), a programme being implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Commission

ECMWF Copernicus Europe's eyes on Earth European Commission