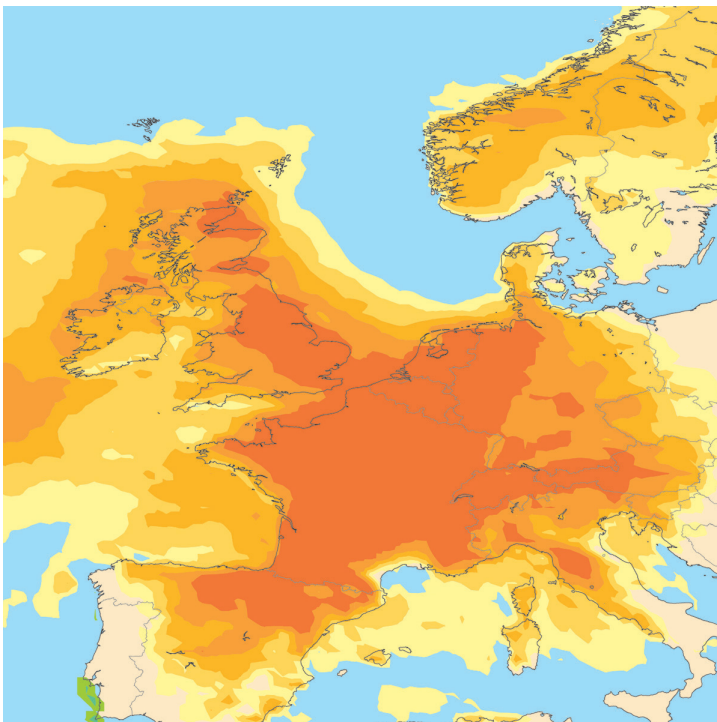




COMPUTING

WEkEO DIAS moves towards operational release



This article appeared in the Computing section of ECMWF Newsletter No. 161 – Autumn 2019, pp. 38–40.

WEKEO DIAS moves towards operational release

Baudouin Raoult, Lawrence Albinson, Ricardo Correa (all ECMWF), Guillermo Grau, Laia Romero (both Mercator Ocean), Michael Schick (EUMETSAT)

As part of the EU's Copernicus Earth observation programme, ECMWF, EUMETSAT and Mercator Ocean have joined forces to implement a Data and Information Access Service (DIAS) platform called WEKEO (Box A). The European Commission's vision is to facilitate the use of the vast amount of Copernicus Earth observation (EO) data by a wide range of users, including institutional bodies, the private sector, scientists and civil society. The idea is also for WEKEO to enable third parties to develop their own value-added products and services. ECMWF, EUMETSAT and Mercator Ocean are key players in the Copernicus programme: ECMWF operates the Copernicus Climate Change Service (C3S) and the Copernicus Atmosphere Monitoring Service (CAMS); EUMETSAT provides data services from the Copernicus Sentinel satellites it operates (Sentinel-3 Marine Mission, Sentinel-4, Sentinel-5 and Sentinel-6), complemented by EUMETSAT data and data from third parties, with a focus on the atmosphere, the ocean and climate monitoring; and Mercator Ocean operates the Copernicus Marine Environment Monitoring Service (CMEMS). WEKEO will work hand in hand with existing services for the provision of Copernicus data, such as the C3S Climate Data Store (Box B). Following a proof-of-concept phase, an operational version of WEKEO is currently being built. This includes a new web portal to be released by the end of this year.

What's in a name? A

The name WEKEO, pronounced 'wikio', alludes to the idea of a collaborative platform, as in Wikipedia, and comprises three distinct elements:

- 'WE', as in the first person plural pronoun, refers to the three centres involved (EUMETSAT, ECMWF and Mercator Ocean) together with all WEKEO users
- 'k' stands for 'knowledge'
- 'EO' stands for 'Earth Observation' and 'Environment Observatory'

WEKEO and the Climate Data Store B

ECMWF is already operating the Climate Data Store (CDS), a cloud-based service for the provision of C3S data and tools and, in the future, CAMS data and tools. The CDS and WEKEO are complementary since they serve different purposes.

Cloud services are traditionally split into three categories: infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS). Although the CDS uses cloud technologies for its implementation, users are unaware of this, as are users of Google Maps, for example; the CDS and its toolbox are therefore considered primarily as software

as a service (SaaS). WEKEO, on the other hand, will give users access to computing resources (virtual machines, virtual disks, virtual networks, etc.) and therefore primarily provides infrastructure as a service (IaaS). WEKEO users will be able to run their own software and models on that infrastructure. The CDS, on the other hand, offers functionality tailored to the processing of C3S and CAMS data through a dedicated toolbox implementing authoritative algorithms. The need for synergies between the two systems is driving the development of both platforms.

WEKEO’s offering

WEKEO consists of cloud-based services that make all Copernicus data and information available to its users. It also provides tools and processing capabilities for data and information processing, including big-data analysis tools. Data is spread across a wide geographical area, and the data volume is very large – certainly too large to replicate to many locations on a continuous basis. Accordingly, WEKEO hides the physical location of the data.

In the crowded arena of EO platforms, which includes generic clouds, thematic/regional EO frameworks and wide-scoped EO platforms like WEKEO and the four DIASes operated by the European Space Agency (ESA), WEKEO puts the emphasis on:

- providing harmonised, performant access to the full range of Copernicus data and Copernicus services
- ensuring up-to-date data by accessing its original source and avoiding replication
- offering a broad spectrum of processing tools running on modern infrastructure
- focusing on an excellent user experience, including first-class user support drawing on the years of experience accumulated by the three WEKEO partners.

Most importantly, WEKEO differentiates itself from other platforms by being built around the idea of federation: data federation, with new data providers being able to join; infrastructure federation, with contributions by the three partners, industrial contractors and other interested parties; and user support federation, with distributed experts in IT and EO products and services.

WEKEO currently offers two packages designed to appeal to a wide range of users. The Essential Plan provides free and open access to all of its data holdings, as well as Jupyter Notebooks. The Advanced Plan additionally includes cloud-based processing and tools, for a flat rate depending on the allocated resources and with an initial free trial period.

From proof-of-concept to operations

The concepts behind the WEKEO platform have been demonstrated during a proof-of-concept phase (known as WEKEO-V0). During this pre-operational period, which began in June 2018, organisations and individuals were invited to try out WEKEO. Valuable feedback has been received from many of these pilot users and has been taken into consideration for new releases of the system. Gauging the demand for such services and getting to know potential users better has also been an important part of the process. Moving forward, the WEKEO partners are currently building the operational version (WEKEO-V1) of the platform. To this end, they have put out to tender the various components of the platform in line with each partner’s respective responsibilities, as illustrated in Figure 1.

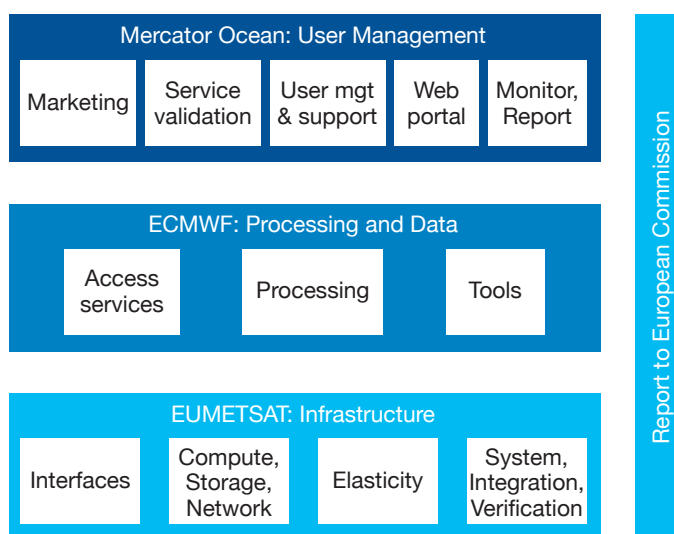


Figure 1 Distribution of responsibilities between the WEKEO partners.

ECMWF is responsible for the procurement of the software to implement data access services, processing and tools. An invitation to tender was published in July 2018. In November that year, the contract was awarded to a consortium comprising GMV Aerospace and Defence, MEEEO, The Server Labs, B-Open Solutions and Sinergise, which delivered the required software components in June 2019.

The consortium demonstrated to ECMWF a Harmonised Data Access (HDA) API, allowing registered users to access all available data in a transparent and consistent way, independent of the underlying access methods and physical data location. They also presented the creation of virtual machines based on images (i.e. templates) that were preloaded with scientific software (data analytics, plotting, etc.). Users are able to install other packages on top of that. The virtual machines also contained the necessary software to use the HDA interface. Furthermore, they showed that users could use predefined 'blueprints' to create, with a few clicks, clusters of inter-connected virtual machines to process big data, using the cluster computing framework Apache Spark, as well as a machine learning cluster based on Google's TensorFlow software. Using Jupyter Notebook, a demonstration was made of training the cluster using limited examples of data.

For Apache Spark, a WEkEO extension has been built that will make the ingestion of big datasets into the cluster much more user-friendly. The new component extends the standard Spark Data Source API to make it possible to download data directly into Spark via the WEkEO HDA interface. A number of satellite image visualisation tools will be available to users either as part of their virtual machine images, such as SNAP from ESA, or as part of additional services, such as Sentinel Hub from Sinergise.

Outlook

All delivered software has now been handed over for operational deployment to the industrial partner responsible for the WEkEO Services and Operations contract, managed by EUMETSAT. In parallel, several projects related to user management are at various stages of execution by Mercator Ocean. Work is already under way on a new WEkEO web portal which integrates all products and services, and whose first release is expected by the end of this year. A centralised monitoring and reporting facility is also envisaged, as well as a revamped user support service. With a solid team across ECMWF, EUMETSAT and Mercator Ocean and their respective industrial partners, a well-differentiated technical and user-oriented approach, and the valuable experience gained from WEkEO-V0, WEkEO is shaping up to become a valuable platform for the EO community in the near future. For more information, visit: <https://wekeo.eu>.

© Copyright 2019

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England

The content of this Newsletter is available for use under a Creative Commons Attribution-Non-Commercial-No-Derivatives-4.0-Unported Licence. See the terms at <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error or omission or for loss or damage arising from its use.