



A joint vision for future collaboration

The EOSC as a scalable computing infrastructure for Earth Observation

Oceans observation data is now replicated and stored on a daily basis with B2SAFE

BBC joining the EOSC Digital Innovation Hub

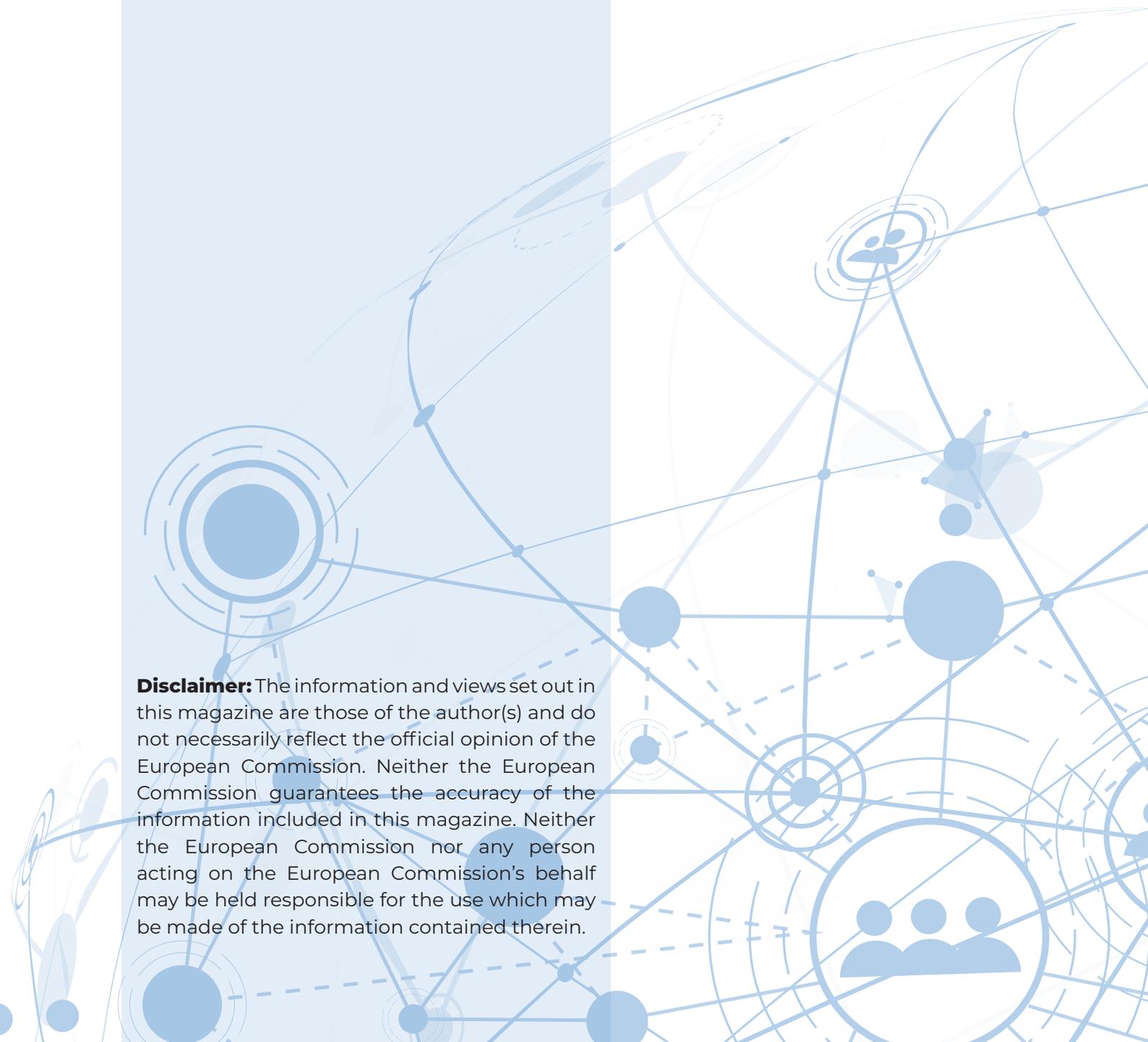
EGI Cloud Compute: success stories from the business pilots

EOSC Early Adopter Programme in full swing

## EOSC-hub Magazine

The EOSC-hub Magazine is a publication of the EOSC-hub project, edited to showcase major results and achievements of the project, collaborations ongoing with other initiatives and updates from the communities. The magazine also provides an overview of the latest highlights from the European Open Science Cloud (EOSC) landscape.

**Disclaimer:** The information and views set out in this magazine are those of the author(s) and do not necessarily reflect the official opinion of the European Commission. Neither the European Commission guarantees the accuracy of the information included in this magazine. Neither the European Commission nor any person acting on the European Commission's behalf may be held responsible for the use which may be made of the information contained therein.



# Issue 4 // Table of contents

---

A joint vision for future collaboration _____	3
The EOSC as a scalable computing infrastructure for Earth Observation _____	4
Oceans observation data is now replicated and stored on a daily basis with B2SAFE _____	6
BBC joining the EOSC Digital Innovation Hub _____	8
EGI Cloud Compute: success stories from the business pilots _____	9
EOSC Early Adopter Programme in full swing _____	10

## About EOSC-hub

**T**he EOSC-hub project brings together multiple service providers to create the Hub: a single contact point for European researchers and innovators to discover, access, use and reuse a broad spectrum of resources for advanced data-driven research.

For researchers, this will mean a broader access to services supporting their scientific discovery and collaboration across disciplinary and geographical boundaries.

The project mobilises providers from the EGI Federation, EUDATCDI, INDIGO-DataCloud and other major European research infrastructures

to deliver a common catalogue of research data, services and software for research.

EOSC-hub is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement 777536.

*Start: January 2018*

*End: December 2020*

*Total budget: €33 million*

*100 Partners*





# A joint vision for future collaboration

## *Tiina Kupila-Rantala presents the white paper outlining the common vision for EOSC*

I have the pleasure to announce the publication of the White Paper “Common Vision, Service Provision, and Role in the EOSC Governance”, prepared in conjunction with our colleagues from OpenAIRE-Advance. The white paper is an output of our Joint Activity Plan and advances the strategic alignment of the two projects.

The work kick-started in November last year and involved fifteen experts representing EOSC-hub and OpenAIRE-Advance projects as well as their underlying e-Infrastructures, the EGI Federation, EUDAT CDI, INDIGO-DataCloud, and OpenAIRE.

### **Priorities**

The white paper highlights a number of priorities that need to be tackled to ensure an effective implementation of the EOSC vision. In our view, the wider EOSC community should focus on:

- Increasing the accessibility and usability and capacities of all facilities contributing to the EOSC landscape (for example, the EOSC Portal and its components the Marketplace and the Catalogue).
- Defining an operational budget in EOSC to cover: access-enabling services and processes, and, increase capacity and/or new capabilities complementing what is already funded at national- and European-level.
- Ensuring the commitment of long-term funding for the sustainability of the EOSC landscape, independent from short-term projects.
- Defining a clear contact point for each e-infrastructure for strategic steering purposes.

### **Actions**

Moving theory into practice, the white paper suggests a number of concrete actions for EOSC-hub and OpenAIRE-Advance, closely related to the priorities listed above:

- Provide input to an EOSC charter on how to increase accessibility, efficiency, effectiveness,
- transparency and simplicity of EOSC services for each individual researcher.
- Provide strategic advice to the EOSC



Governance with the support of their national constituents from the academic and research community in the form of a research and academic infrastructure situation report.

- Provide business plans for EOSC services operations.
- Leverage self-organization principles to form a strategic point of contact to infrastructures for the EOSC Governance Board to interact with.

### **What's next?**

Within this framework, EOSC-hub and OpenAIRE-Advance can kick-start the implementation of the recommendations of the white paper by:

- Adding the identified actions to the EOSC-hub and OpenAIRE-Advance Joint Activity Plan.
- Further formalize and encourage the cooperation of our constituent legal entities to address issues beyond the EC projects.
- Jointly seek dialogue with the EC and the EOSC Governance.

**Tiina Kupila-Rantala** is Deputy Managing Director at CSC IT Centre for Science, and lead the Strategy Work Package (WP2) of EOSC-hub.



# The EOSC as a scalable computing infrastructure for Earth Observation

## *How EOSC could open the way for Earth Observation innovation*

The massive streams of high-resolution Earth Observation (EO) data derived from the EU Copernicus Sentinel sensors, have established Europe as the predominant spatial data provider for environmental monitoring applications. This data is made available under open license with an unprecedented frequency and spatial extent.

In principle, these data sources can inspire a wide range of science and monitoring applications from regional to continental scales. In practice, innovation is mostly happening outside Europe by large US IT companies. This leads to the unfortunate situation where EO science user communities need to rely on non-European platform suppliers for the Big Data Analytics they need to scale high-volume use of data streams.

There is world-class expertise in EO analytics in Europe. But we are missing a solution to provide core cloud services coupled to an online long-term data archive of Sentinel.

Europe needs:

- Easily accessible European computing environments to allow scaling and sharing of (Sentinel) data among a large community of users.
- A European platform, similar to the Google Earth Engine, to make large-scale storage accessible through sophisticated indexing and caching solutions with an advanced application programming interface (API).

A key requirement is a core computing and storage architecture based on principles tailored to handle very large data sets and fast user query response.

### **EOSC as a solution**

The European Open Science Cloud has the potential to become a viable European



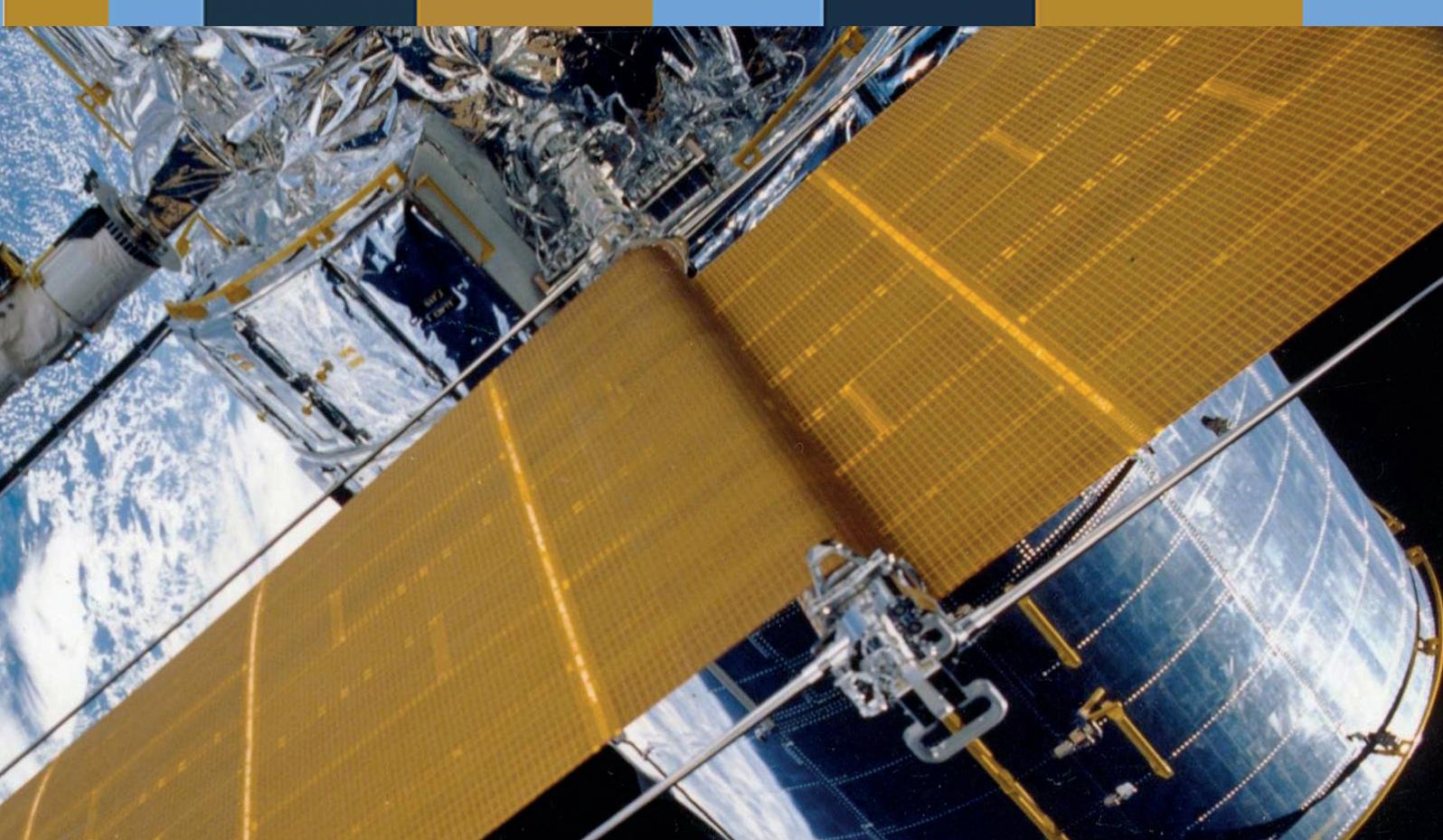
alternative to Google Earth Engine for the scientific EO community. The federated resources provided through the EOSC-hub project can become the storage and computing infrastructure necessary to enable full scaling across Copernicus data inputs. As a platform, it should lead to a collation of the many European initiatives in EO software, by establishing a common interface to massively parallel server workflow handling applied to an optimally indexed data storage format.

Components for the client API to define workflow graphs can be adopted from existing open source frameworks (e.g. Jupyter notebooks, python and node.js). An interface with European open datasets would demonstrate immediately the advantage of the EOSC infrastructure in practical EO science applications. This impact can be enhanced by creating an open science data sharing environment. Facilitating executable data analysis “papers” that use EOSC as the common platform would boost reproducibility and scaling of EO science results.

### **Copernicus sustainability**

The capability of commercial providers such as Google and Amazon in the big data analytics domain pose a serious risk to the continuation of the successful Copernicus programme,





aggravated by the fact that there are no concrete plans for Europe to maintain a full online archive of Sentinel data.

Here there is an opportunity to leverage the existing EOSC infrastructure and capture European EO expertise around it. ESA is already deleting the oldest Sentinel data holdings from the online archive, and while the deleted data can still be retrieved from 'cold' storage (e.g. tape archives), this limits applications requiring long time series. To the best of our knowledge there is no European initiative planning to host the long-term Sentinel data archive.

The EOSC could therefore be used to:

- Host the online archive of all Sentinel data.
- Create an optimized IT infrastructure environment for the EO community

Any action taken in this regard should be in collaboration with and build on existing experiences, such as the Copernicus DIAS and other projects.

## How it could be set up

So how could we accomplish this?

- First, assess the feasibility of establishing EOSC as the host for the permanent archive of all Copernicus Sentinel data, knowing that the data volume in 2018 was 9,7 Petabytes, with annual growth of 6 Petabytes per year.

- Second, working closely with ESA, adopt a phased approach towards hosting the full archive. For example:

- Start with most important geographical areas
- Prioritise the S1 and S2 sensors
- Structure data transfer to prioritise European uptake
- The gradual phasing in can be accompanied by a gradual reduction of the ESA hosted archive
- In parallel, establish alternatives to S3 storage of complete image files to facilitate fast multi-scale data collation and interactive analytics and visualization. This activity would greatly benefit from experience in other EOSC big data analytics domains.

## Conclusion

The federated nature of EOSC makes it a prominent candidate to serve the long term data storage and analytics challenges of the EU's Copernicus Sentinel program. By leveraging its expertise in other Big Data Analytics domains, it can extend its scope to serve the European Science Earth Observation community.

This article was prepared with contributions from EuroGEOSS and the EC Joint Research Centre



# Oceans observation data is now replicated and stored on a daily basis with B2SAFE

---



There is an increasing concern about global change and its regional impact. For example, sea level is rising at an accelerating rate of 3 mm/year, Arctic sea ice cover is shrinking and high latitude areas are warming rapidly<sup>1</sup>. These effects are caused by a mix of long-term climate change and natural variability. Lack of sustained observations of the atmosphere, oceans and land has hindered the development and validation of climate models.

The international Argo programme was initiated in 1999 as a pilot project endorsed by the Climate Research Program of the World Meteorological Organization, the Global Ocean Observing System (GOOS), and the Intergovernmental Oceanographic Commission (IOC) to set up the first-ever global in-situ ocean observing network in the history of oceanography thanks to a global array of more than 3500 instruments reporting subsurface ocean properties to a wide range of users via satellite transmission links to data centres.

This increasing volume of data being produced, curated, and made available by Argo requires services that are able to optimize the delivery and staging of data for researchers and other users of scientific data. Specialized data services for managing data life cycle, for creating and delivering data products, and for customized data processing and analysis all play a crucial role in how Argo serves its communities, and many of these activities are time-critical needing to be carried out frequently within specific time windows.

The EOSC-hub project is helping the Euro-Argo community, the European chapter of the Argo framework, in managing this huge amount of real-time data via the EOSC-hub Marine Research Competence Center.

EOSC-hub is supporting Euro-Argo in developing a Data Analytics Platform to support data

---

<sup>1</sup> [https://www.ipcc.ch/site/assets/uploads/2018/02/AR5\\_SYR\\_FINAL\\_SPM.pdf](https://www.ipcc.ch/site/assets/uploads/2018/02/AR5_SYR_FINAL_SPM.pdf)



storage and processing. Data are staged in/out the EUDAT infrastructure using the B2STAGE service, the EUDAT service meant to transfer research data sets between data resources and computing workspaces, and processed through the use of EGI Cloud resources in the form of containers.

A web application to enable users to configure data subsetting or automated subscription to visualize and stage the data from B2SAFE (the EUDAT data replication and long term preservation service) to and from the data analytics platform has been recently developed, allowing a daily synchronisation.

This is a great achievement for the Euro-Argo community as all new data sets received from regional centers are pipelined through using configurable processes ranging from processing to long term preservation. Thanks to this new Data Analytics platform, researchers can concentrate their efforts on analysing data rather than taking care of their underneath management.

The Euro-Argo / EOSC-hub collaboration is part of the activities of the Marine Competence Centre.



# BBC joining the EOSC Digital Innovation Hub

The BBC Research & Development presented their EOSC-hub use case during the 2nd EOSC-hub week in Prague, April 2019. Luka Murn, R&D Engineer at BBC R&D, introduced the use case related to his projects, Video Coding and JOLT.

The team within BBC R&D which mainly focuses on Video Coding Research also delivers innovative results on multiple aspects of video technology. The goal of their work is to support the delivery of high-quality content to all BBC audiences. In addition to this work, the team is also looking into various innovative video solutions to ensure that BBC is future-proof in terms of providing new services.

Since their work needs a lot of computing resources, the BBC R&D's team found support from the Polish Poznan Supercomputing and Network Center (PSNC), to access the needed infrastructure available via EOSC-hub and be part of the EOSC Digital Innovation hub.

Luka said about the collaboration: "The use of large-scale processing resources has the

capability to transform how BBC obtains, produces and delivers content in challenging scenarios. A move away from expensive bespoke broadcast specific facilities and hardware to more commoditized scalable-cloud based resources will enable BBC R&D to further scale its research activities and provide innovative media technologies."

The EOSC DIH is the place to go for SMEs and large businesses to engage with the European Open Science Cloud and get access to technical services and research data.

We are pleased and excited for a fruitful collaboration with the BBC R&D team.

**Luka Murn** is R&D Engineer at BBC R&D.

More on BBC R&D: [bbc.co.uk/rd](http://bbc.co.uk/rd)

More on the EOSC DIH: [eosc-hub.eu/digital-innovation-hub](http://eosc-hub.eu/digital-innovation-hub)



# EGI Cloud Compute: success stories from the business pilots

## *Virtual machines on-demand with complete control over computing resources*

**E**GI Cloud Compute gives researchers to deploy and scale virtual machines on-demand. It offers guaranteed computational resources in a secure and isolated environment with standard API access, without the overhead of managing physical servers.

The service is part of the EOSC-hub catalogue and was first launched in 2014 and has since supported high-impact research and business use cases. Here are some recent examples from the EOSC-hub Digital Innovation Hub.

### **App to monitor athletic performance**

As a EOSC-hub business pilot, Moxoff developed VAMOS – a web-application where authenticated users can analyze and monitor performance of athletic gestures.

They deployed two virtual machines using the EGI Cloud Compute service: one for the front-end of VAMOS with 4 CPUs, and another with more resources for the back-end with 12 CPUs and 1 GPU.

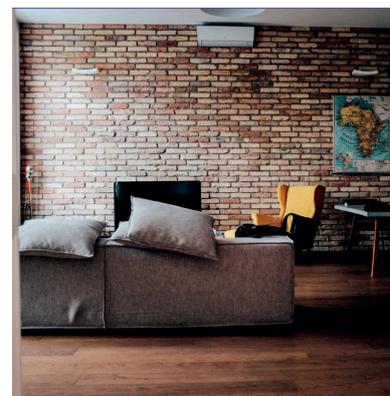
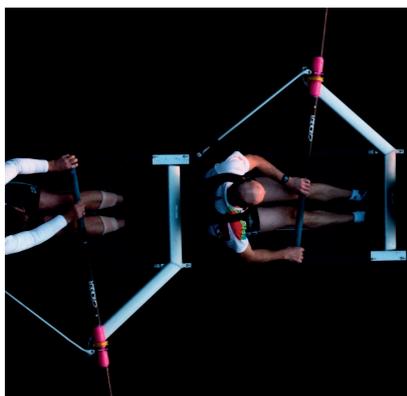
VAMOS allows coaches and athletes to save time and increase their efficiency as the data is processed by advanced algorithm and methods (such as functional data analysis) automatically, extracting all the KPIs they need into standard reports.

### **Improved analytics for the furniture industry**

AIDIMME designed and deployed DataFurn – a platform to provide intelligence and insight into furniture industry trends.

DataFurn was developed as an EOSC-hub pilot and used ten EGI Cloud Compute virtual machines with 80 CPUs to deploy its architecture. In six months of operations, DataFurn consumed the equivalent of 230 compute-months (~170,000 CPU hours).

This cloud computing power allowed DataFurn to experiment on different resource-intensive algorithms for analyzing and indexing the related content that has been curated for manufacturing SMEs.



### **Data cloud services to manage harmful algae blooms**

Harmful Algal Blooms happen when toxic microalgae proliferate beyond control and take over coasts, rivers and lakes with costly impacts, for example red tides.

Ecohydros developed CyberHAB: a versatile platform powered by cloud computing able to combine large volumes of data for the management of harmful algal blooms.

CyberHAB uses Jupyter Notebooks as an interface and EGI Cloud Compute service for the computing back-end. They also deploy the INDICO-DataCloud IM PaaS Orchestrator.

EGI Cloud Compute is available through the EOSC Marketplace and is provided by the 22 cloud data centres of the EGI Federation.



# EOSC Early Adopter Programme in full swing

**F**resh from concluding its first call, the EOSC Early Adopter Programme, run by the EOSC-hub project, is gearing up to launch its second call by mid-October 2019.

The programme is geared towards research communities interested in exploring the state-of-art technologies and services offered by the European Open Science Cloud (EOSC) and provides a first taste of what the science cloud will offer in the future.

The Early Adopter Programme (EAP) aims to enable research communities to scale up their own in-house infrastructure and to access a richer set of resources. Through the EOSC-hub project and the EAP partners, successful early adopters will be provided with expertise and resources to enable active usage of the EOSC and foster a culture of co-operation between researchers and EOSC providers.

## What's on offer

The programme offers a wide range of services (all available through the EOSC Marketplace) for data discovery and reuse, data processing and analysis, data management, curation and preservation, and access, deposition and sharing, federated authentication and authorization.

The services and resources in scope for the programme are those currently provided by the EOSC-hub project and the Early Adopter Programme partners, namely OCRE, OpenAIRE and GÉANT.

## Next call

Upon launch of the 2nd call, applicants should submit an online form describing their requirements and the scope and scale of services they require. The 2nd call for the EAP is scheduled for mid-October.

Successful applicants would have demonstrated a well-defined scientific challenge that can be addressed by the use of multiple services from EOSC-hub and its partners. Proposals that can demonstrate scientific excellence and societal impact will have an advantage.

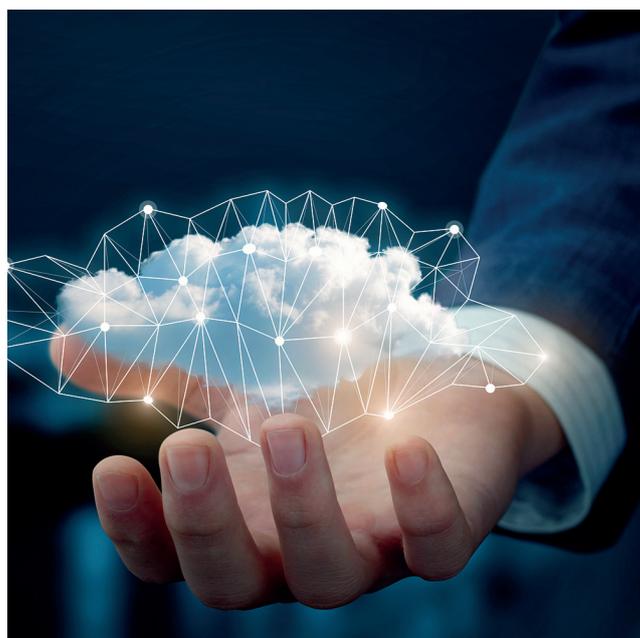
## MoU with OCRE makes commercial services available in EAP

During the 2nd round of the Early Adopter Programme, commercial IaaS cloud resources from providers in the GÉANT IaaS framework will be made available through OCRE. This is part of a wider collaboration between the EOSC-hub and OCRE that will be cemented in a Memorandum of Understanding between the two projects.

OCRE aims to accelerate the adoption of digital services in the European research community, by bringing together cloud providers, Earth Observation (EO) organisations and the research and education community, through ready-to-use service agreements and €9.5 million in adoption funding. Resulting services from a tender it will run in 2019 will be included in EOSC.

### Learn more about OCRE:

[www.ocre-project.eu](http://www.ocre-project.eu)



**For more information about the Early Adopter Programme, visit**  
[eosc-hub.eu/eosc-early-adopter-programme](http://eosc-hub.eu/eosc-early-adopter-programme)





**EOSC-hub**

# **EOSC-hub Week 2020**

**18-20 May 2020**

**ZKM - Center for Art and Media,  
Karlsruhe, Germany**

**#SAVETHEDATE**





 [eosc-hub.eu](https://eosc-hub.eu)

 [@EOSC\\_eu](https://twitter.com/EOSC_eu)

 [/company/eosc-hub](https://www.linkedin.com/company/eosc-hub)

 [info@eosc-hub.eu](mailto:info@eosc-hub.eu)



EOSC-hub receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777536.

