



MAGAZINE

ISSUE 3
March 2019



A large, abstract background graphic consisting of a complex network of interconnected nodes and lines, rendered in shades of blue, white, and grey, creating a sense of depth and connectivity.

EOSC-hub: one year of achievements

Five ways of working with EOSC-hub

EOSC-hub and the ESFRI cluster projects

EOSC in practice: ICOS

The two-fold benefit of collaborating with EOSC-hub: the EnviDat story

OPENCoastS and EOSC-hub

New improvements for data-driven research in the Geohazards community

HADDOCK helps scientists to look at the evolution of brain tumours

Data sharing done right



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 3 // Table of content

EOSC-hub: one year of achievements	2
Five ways of working with EOSC-hub	4
EOSC-hub and the ESFRI cluster projects	5
EOSC in practice: ICOS	6
The two-fold benefit of collaborating with EOSC-hub: the EnviDat story	8
OPENCoastS and EOSC-hub	9
New improvements for data-driven research in the Geohazards community	10
HADDOCK helps scientists to look at the evolution of brain tumours	11
Data sharing done right	12

About EOSC-hub

The 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, EUDAT CDI, 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



EOSC-hub: one year of achievements

EOSC-hub started in January 2018 with a mission to create 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.

One year on, here is a summary of the main achievements.



Optimised access to IT equipment and services

- ▶ EOSC-hub service portfolio integrating over 70 services from EGI, EUDAT, INDIGO-DataCloud and from research communities
- ▶ From Nov 2018, EOSC-hub services are discoverable and orderable via the Marketplace in the context of the EOSC portal
- ▶ July-Dec 2018: +27% registrations of service entries in the Marketplace
- ▶ EOSC-hub operates the EOSC Portal website, the Marketplace and the whitelabel EOSC AAI
- ▶ Service Providers onboarding process in place and integrated with the EOSC portal
 - ▶ 21 new services currently under integration
 - ▶ Streamlined service management
 - ▶ > 60 individuals FitSM certified

Success story:

HADDOCK is an EOSC-hub Thematic Service for structural biologists provided through the EOSC Marketplace. In the past year: +30% users, +126% simulations submitted.



Partnerships with industrial and private partners

- ▶ The EOSC Digital Innovation Hub (DIH) was successfully established in late 2018 and is now officially part of the EU Digital Innovation Hub catalogue
 - ▶ 6 business pilots up and running
- ▶ Vouchers programme for free Cloud service access now open!
- ▶ Earth Observation (EO) service providers actively engaged in the Marketplace

Success story:

Terradue, an EO private company, provides the engineering and operational support for the Geohazards Exploitation Platform (GEP), which offers a rich set of ready to use EO data processing services for the analysis and monitoring of earthquake, volcanoes and landslides. GEP is integrated with the EGI Federated Cloud and is now part of the Marketplace offering.

- ▶ 1.4 Million Cloud Compute hours
- ▶ +108% increase of use of Terradue services



Over 800 individuals trained in research and academic sectors



- ▶ 9 Data Management training events (6 with OpenAIRE-Advance)
- ▶ Other areas
- ▶ Security, AAI, Storage and Cloud Technologies, service-specific training for users
- ▶ 4 FitSM Certification Training events + 2 webinars + 1 individual FitSM process grouping webinar
- ▶ Domain-specific: CODATA, VIRGO, ENVRI, INSTRUCT and NGSchool

Success story:

OPENCoastS: first national EOSC-hub thematic service provider which joined the Marketplace, opening access to international user communities thanks to VA funding.

- ▶ 148 participants from 14 countries attended the OPENCoastS e-tutorial.

Progress with technical implementation



- ▶ 2 technical joint roadmaps developed (OpenAIRE and GÉANT) focusing on DMP, AAI, Cloud interoperability, etc. and one under development with RDA
- ▶ 4 Thematic Service Providers successfully integrated into the EOSC Marketplace exploiting the open interfaces EOSC-hub provides
- ▶ Implementation of AARC AAI Blueprint for the EOSC AAI
- ▶ Implementation of the EOSC-hub service management system compliant to the FitSM standard and to ISO service management and quality standards

Success story:

EOSC branded multi-tenant service: for federated AAI supporting all standard authentication options (eduGAIN, social media, ORCID) and group membership capabilities.

Completed integration activities: Interconnection of EGI Check-in with EUDAT B2ACCESS; Online Certification Authority (Rcauth) migration and support for high availability.

Extended user base



- ▶ Expanded access to services outside the traditional user base via the EOSC portal
- ▶ 8 Competence Centres piloting EOSC-hub Services: ICOS, ELIXIR, Radio-Astronomy; Disaster Mitigation, EPOS-Orfeus, CLARIN, GEO-DAB
- ▶ A stakeholder database with over 50 potential leads
- ▶ 5 new funded ESFRI-cluster projects working in collaboration with EOSC-hub
- ▶ An online community of around 2000 members (Twitter, LinkedIn, SlideShare)

Success story:

EGI Federation: increase in the use of computing resources:

- ▶ 4.4 billion CPU core wall time delivered in 2018
- ▶ > 1 million computing cores for the first time in the EGI history
- ▶ 356 PB disk & 380 PB tape storage

EUDAT CDI :

- ▶ Over 700K datasets discoverable (repositories from 16 communities harvested, 10k datasets in B2SHARE)
- ▶ 90M data objects registered within the CDI (+15% in 2018)
- ▶ 1600 users registered on B2ACCESS (+33% in 2018)



Five ways of working with EOSC-hub

Gergely Sipos writes about EOSC-hub collaborative framework for EOSC implementation

The EOSC-hub Marketplace (marketplace.eosc-portal.eu) plays a central role in implementing the EOSC vision, where researchers can order the services, tools and computing resources they need for their work. EOSC-hub offers five engagement channels for stakeholders to benefit from EOSC and the EOSC-hub Marketplace:

Use EOSC-hub services

The EOSC-hub service catalogue is available through the EOSC-hub Marketplace and includes services provided by the EGI Federation, EUDAT, the INDIGO-DataCloud project and services provided by research communities and research infrastructures.

The EOSC-hub services cover Compute & Storage, Data Management, Processing & Analysis, Security & Operations, Sharing & Discovery, Training & Support.

The Engagement Team offers advice on which services to use, and on how to combine multiple services for maximum value.

Interested? Go to the EOSC-hub Marketplace marketplace.eosc-portal.eu

Join as provider

The EOSC-hub Marketplace is open for service providers willing to develop EOSC into a rich environment.

By contributing services and resources to the EOSC-hub Marketplace, providers can reach a wider user base and promote service adoption outside traditional user groups. The Marketplace is also a free online platform to manage service requests, interact with users and provide support for them, including negotiation of service levels (SLA).

Interested? Apply via the webform: eosc-portal.eu/join-provider

Training and user support

EOSC-hub runs a training programme on topics ranging from data management issues to integration support and discipline-specific tutorials.

Training events and materials are catalogued by EOSC-hub on the project website (eosc-hub.eu/training-material) and the project promotes online and face-to-face courses.

Interested? Browse the Training Materials eosc-hub.eu/training-material

Digital Innovation Hub

EOSC-hub has set up the EOSC Digital Innovation Hub (DIH) as a mechanism for private companies to collaborate with public sector institutions in order to access technical services, research data, and human capital.

The DIH is open for new partners to join either as service, resource or technology providers or as new business pilots. There is also a voucher scheme to support testing.

Interested? Fill the form here: bit.ly/2SMVFQd

Networking

EOSC-hub is keen to build solid networks to support the national e-Infrastructures (NGIs) that represent the backbone of the EGI Federation, OpenAIRE national nodes (NOADs), EUDAT CDI nodes and GÉANT's national members (NRENs).

EOSC-hub is open for partnership with additional stakeholders to reach and support new types of users and providers.

Interested?
Contact us via
contact@mailman.eosc-hub.eu.



EOSC-hub and the ESFRI cluster projects

Research Infrastructures have strong links with research communities and projects, manage significant data volumes and develop innovative data analytics tools, ensuring effective research data exploitation.

Five ESFRI cluster projects have been launched early this year, providing a gathering point for various ESFRI projects and landmarks to connect to the EOSC.

ENVRI-FAIR for environmental research

envri.eu/envri-fair/

ENVRI-FAIR will implement the ENVRI-hub - a virtual, federated machine-to-machine interface to access environmental data and services provided by the contributing RIs.

The complete set of thematic data services and tools will be incorporated into the EOSC service catalogue, through the EOSC-hub Marketplace.

PaNOSC for multidisciplinary scientific analysis

panosc.eu

PaNOSC will help the Photon and Neutron ESFRIs to adopt and implement data management, simulation and analysis services, and to make their open data available to the EOSC.

It will work closely with EOSC-hub partners to integrate general-purpose distributed computing and data management solutions and promote its products through the EOSC Portal.

ESCAPE for astronomy and particle physics

www.escape2020.eu

ESCAPE brings together ESFRI facilities of astronomy, astroparticle & particle physics into a single EU collaborative cluster. Plus, it will create a cross-border & multi-disciplinary environment that will benefit EOSC thanks to the management of extremely large data

volumes at the multi-exabyte level.

It will also support “scientific software” as a major component of RI data to be preserved and exposed in EOSC through dedicated catalogues.

SSHOC for social sciences and humanities

www.sshopencloud.eu

SSHOC aims to provide an open cloud for social sciences and humanities where data, tools, and training are available and accessible for users. This open cloud aims to be a part of the EOSC.

The consortium covers the whole data cycle, from data creation and curation, to optimal data reuse, and can address training and advocacy to increase actual reuse of data.

EOSC-Life for life sciences

EOSC-life.eu

EOSC-Life brings together biological and medical RIs to create an open collaborative space for digital biology.

It aims to publish FAIR life science data resources for cloud use creating an ecosystem of innovative tools in EOSC and enabling groundbreaking data-driven research in Europe by connecting life scientists to EOSC.

EOSC-hub's Role

A fundamental goal of these INFRA-EOSC clusters is to reuse externally supported technical solutions and services (for example those provided via EOSC-hub), and to have their own services and data in the EOSC Portal. Representatives from these cluster projects are in the EOSC-hub Strategy Board advising the project on exploitation and service deployment plans. On the grassroots level, coordinated user support, training and outreach activities are also foreseen.

For more information, visit:

bit.ly/2SQUYFs



EOSC in practice: ICOS

Alex Vermeulen on how ICOS is contributing to the science of climate change

What are the main research questions of your community?

ICOS is tied to one of the largest challenges of our time - climate change. Climate change is caused by greenhouse gas emissions related to human activities, such as traffic and industrial production. ICOS observes the gases, their sources, transport, sinks, and finally their balance, and produces high-quality, standardised data on them. The ICOS observations are shared freely to be used, for example, in scientific and policy relevant assessments. Even these results can be published and communicated through ICOS.

In the Paris Agreement, most of the world committed to reduce the human caused greenhouse gas emissions to zero by 2050. However, there are still many uncertainties in emission levels. The changing climate itself also strongly influences the natural uptake and emissions of greenhouse gases. This needs to be taken into account when politicians set the emission reduction goals and paths.

How is ICOS organised?

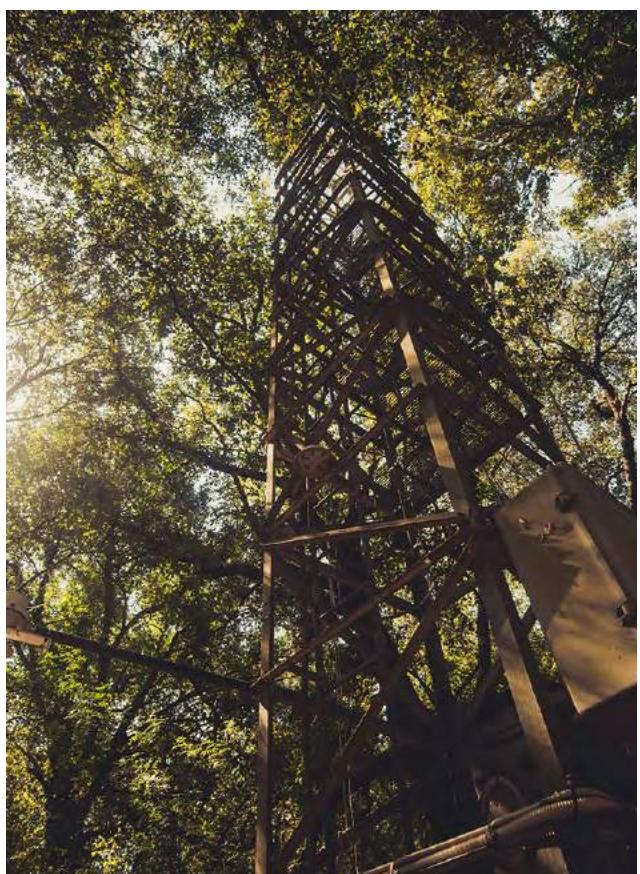
ICOS involves more than 70 institutes from all over Europe and about 500 scientists and technicians in the field, laboratories and other facilities. We have more than 130 stations distributed throughout the twelve Member and Observer countries, and are working hard to add new states to the network, in order to further improve the spatial coverage. ICOS has been established as an ERIC, European Research Infrastructure System, since November 2015. The Thematic Centres, laboratories, Carbon Portal and Head Office are all located in different member countries, and operated jointly by several members.

What are the services ICOS provides?

ICOS provides quality-controlled high-precision time series of greenhouse gas atmospheric concentrations and exchange fluxes in all three domains: Atmosphere, Ecosystem and Ocean.

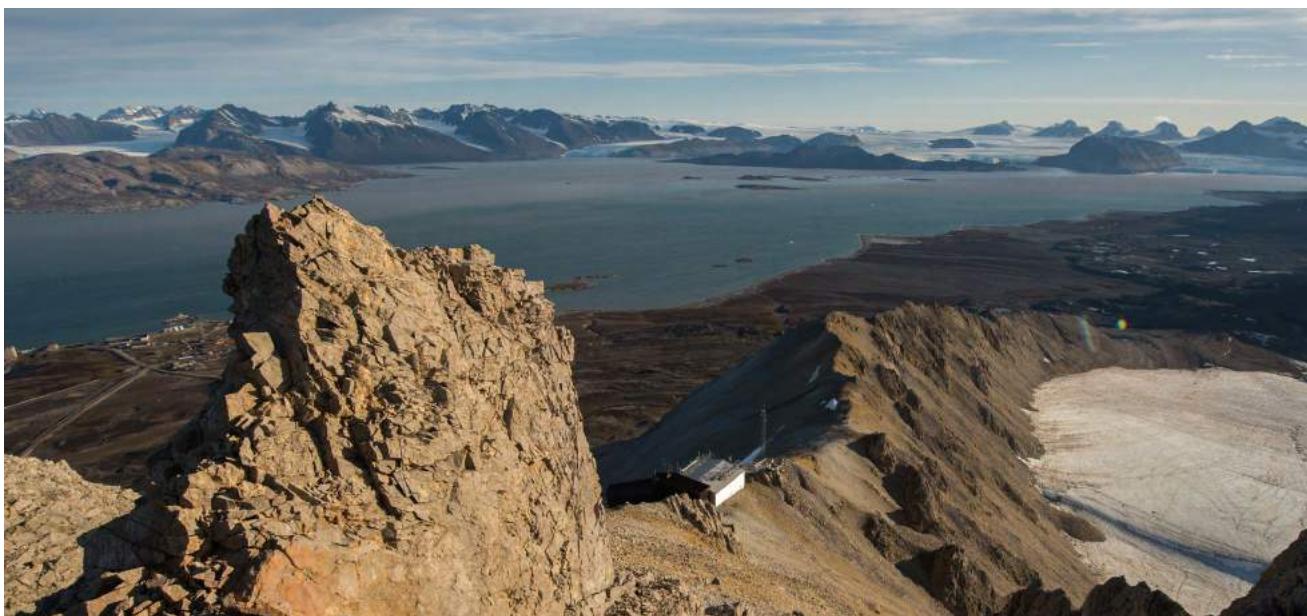
ICOS also delivers near real-time information with a delay of maximum 24 hours of the same variables. And finally we store and make available all ICOS data, both raw data and relevant metadata, through the faceted data portal. We strive for maximum transparency and granularity, so each data object has a persistent identifier. All metadata and data are linked through a linked open data approach, and users can also query our (meta)data.

We also provide users with the opportunity to collaboratively get access to ICOS and ancillary data in so called virtual research environments, for example making use of Jupyter notebooks. Through this they get direct access to the data without the



ICOS Castelporziano Ecosystem Station is located 25km North from Rome. Picture copyright ICOS, image by Konsta Punkka





ICOS Zeppelin Atmosphere Station is located in the Arctic Ocean, Svalbard archipelago, Norway.
Picture copyright ICOS, image by Konsta Punkka

need to download and locally store the data, and they can create a transparent and reproducible workflow. The end result of this work can then be published and shared through the ICOS Carbon Portal. We plan to select useful products and models to develop through this platform and turn them through co-design with the atmospheric communities into operational services that are served from the ICOS Carbon portal.

Furthermore, we will offer services that are directly usable by policy makers and the general public to evaluate the progress of emission reductions in their region. This will be done by using data assimilation techniques that use the ICOS observations and combine this with elaborate models to improve the emission estimates and decrease their uncertainties.

What are the computational challenges?

The basic observational data that ICOS provides, including the raw data, is not very large, however it won't fit the average single server. The ancillary data needed to run the models that are required to interpret the data are, however, several orders of magnitude larger. For example, data assimilation models require millions of CPU compute hours, large memory resources and data storage for every calculation. And the relatively simple Footprint Tool that we run, needs several terabytes of spinning disk storage.

How do you address these challenges?

We already make use of EOSC-hub compute and

storage services and we are working together with the project, as a Competence Centre, to optimize our workflows. For example:

In the Footprint Tool case, data is stored on B2SAFE and taken to/from the ICOS Carbon Portal. Users interact only with the ICOS Carbon Portal, which instantiates virtual machines (VMs) in the EGI Federated Cloud.

For the Ecosystem GHG Flux case, ICOS will build on a current pilot under development in the framework of ENVRIplus. This is currently using the gCube Data Analytics platform, combining resources from D4Science and EGI to orchestrate parallelised High-Throughput Compute processes.

How do you imagine your field in ten years?

To improve the quality of the data assimilation and to make full use of all the information in the ICOS observations, the evaluation models need a better resolution. An increase in resolution with a factor of 10 – which is needed to reach the kilometer scale – would increase the demand by factor 1000. So, more challenges ahead!

ICOS website: www.icos-ri.eu

ICOS data portal: www.icos-cp.eu/data

Alex Vermeulen is the Director of the open data portal of ICOS, i.e. Integrated Carbon Observation System.



The two-fold benefit of collaborating with EOSC-hub: the EnviDat story

Ionut Iosifescu Enescu explains why and how EnviDat is collaborating with EOSC-hub and the results achieved so far

How would you describe EnviDat to someone who is not very familiar with it?

EnviDat (envidat.ch) is the environmental data portal of the Swiss Federal Institute for Forest, Snow and Landscape Research WSL (wsl.ch). EnviDat provides a range of services in the area of research data management with particular focus on (1) data publication with provision of metadata and Document Object Identifiers (DOIs), (2) maintenance of an efficient data repository for validated, quality-controlled and properly documented, thus reusable data sets and (3) supporting and providing options to researchers for sharing the curated data sets. As a repository, EnviDat actively implements the FAIR (Findability, Accessibility, Interoperability and Reusability) principles by ensuring that the published research outputs have licenses that are as open as possible, and as protected as necessary.

Why EnviDat approached EOSC-hub?

EnviDat aims to disseminate its data sets as broadly as possible in order to foster international research cooperation in the field of environmental science and contribute to the ongoing cultural evolution in research towards openness, shared data and opportunities for collaboration. Becoming an EOSC-hub integrated thematic service provider means leveraging on an important platform for making our datasets discoverable.

How is EOSC-hub supporting EnviDat?

The collaboration with EOSC-hub is two-fold:

- i) Using EOSC-hub services to make the EnviDat records visible for EOSC/EOSC-hub users, thus enabling an easier discovery of valuable environmental data sets that are owned by WSL. Currently 163 EnviDat datasets are already discoverable via B2FIND (bit.ly/2EPjOAu).
- ii) Promoting the EnviDat portal in the EOSC-hub Marketplace and the EOSC Portal to increase our visibility as a portal. Furthermore, we hope that



this step would increase our recognition as a professional institutional repository and portal at Swiss and European levels.

What are your first impressions?

EnviDat is delighted with the professionalism of the EOSC-hub team and the fruitful integration process, especially regarding the technical interoperability. Furthermore, the feedback related to the integration in EOSC Portal, motivated us to start the process for further professionalizing the EnviDat portal on several fronts.

What are the plans for the near future?

We welcome the exchange of know-how and best practices with any members of EOSC-hub community that are interested in our initiative. We are looking at the future developments of EOSC to see how EnviDat can better align to it.

Ionut Iosifescu Enescu is Technical Coordinator of EnviDat at the Swiss Federal Research Institute WSL.
envidat@wsl.ch



OPENCoastS and EOSC-hub

*Anabela Oliveira
introduces the service and
how EOSC-hub supports it*

What is OPENCoasts?

The OPENCoastS service builds on-demand circulation forecast systems for user-selected sections of the coast and maintains them running operationally for a user-defined time frame. This daily service generates forecasts of water levels and 2D velocities (and wave parameters in the near future) over the spatial region of interest for periods of 48 hours, based on numerical simulations of all relevant physical processes. OPENCoastS was developed by LNEC, LIP, CNRS/University of La Rochelle and University of Cantabria.

Forecast systems are fundamental components of emergency response and routine management of coastal regions. They provide coastal managers with accurate and timely predictions on water conditions (e.g., water levels and velocities, wave characteristics), and support multiple uses such as navigation, water monitoring, port operations, dredging works and construction activities on the coast.

Forecast systems can also help emergency services and civil protection authorities by anticipating natural disasters (e.g. floods, storm surges), predicting the impacts of human-made coastal disasters (e.g. oil and chemical spills) and supporting in search and rescue operations.

For over 10 years, LNEC has been deploying WIFF – a Water Information Forecast Framework, applicable from the ocean to hydrographic basins, including the urban interface.

However, good forecast systems require not only a strong knowledge of coastal processes but also of information technology, along with access to significant computational and storage resources. The setup and maintenance of these systems constitutes a major difficulty for many coastal research groups. Before the EOSC-hub project, services for generic deployment of forecast frameworks at user-specified locations were not available.



OPENCoastS and EOSC-hub

The OPENCoastS service takes advantage of two e-infrastructures for computational and storage resources: the National Advanced Computing Infrastructure – INCD (integrated in the National Roadmap for Infrastructures of the Foundation for Science and Technology of Portugal) and IFCA (Institute of Physics of Cantabria, Spain). OPENCoastS is supported by the EGI computational resources, through the EOSC-hub project, being available as one of its thematic services.

The architecture of the OPENCoastS.pt service includes:

- ▶ the user interface component, a web based portal;
- ▶ the computation component, where simulation results are generated and post-processed;
- ▶ and the archive component, responsible for preserve all relevant data.

Last year in December, LNEC and its EOSC-hub partners organised a one-day tutorial on OPENCoastS. The training was provided on site in Portugal, Spain and France, by video conferencing and by web streaming. This tutorial covered all the relevant topics to help everyone understanding and using the service. The course was a huge success and gathered 148 participants from 14 countries.

More information on the tutorial is available at opencoasts.lnec.pt

Link to the service: opencoasts.ncg.ingrid.pt

New improvements for data-driven research in the Geohazards community

Hervé Caumont talks about two GEP services recently integrated in EOSC-hub

In the past few years, the European Space Agency (ESA) ran a set of R&D activities to create an ecosystem of interconnected Thematic Exploitation Platforms (TEPs) on European footing, complementing the Earth Observation (EO) Ground segments approach. The Geohazards TEP or “GEP” is a consortium led by Terradue (Italy) and gathering six partners from Germany, Italy, Spain, France and Greece (DLR EOC, TRE Altamira, CNR-IREA, INGV, CNRS/EOST, CNRS/ENS & CRL Lab) with the aim of building the Platform and sustaining an active user community on it.

The GEP development and validation process was driven since 2015 by requirements analysis activities in consultation with users:

- ▶ A large GEP Community Survey launched in February 2016;
- ▶ Monthly “service utility” feedbacks, from registered early adopters planned to reach 100 by end 2019;
- ▶ Collaboration with application development projects.

A key design advantage of GEP is its infrastructure independence and auto-scaling capabilities. Based on the collaborative approach supported by Terradue Cloud Platform for the integration of new applications, a growing number of “Thematic Apps” is being contributed by multiple partners onto the GEP listing.

By joining the EOSC-hub “EO Pillar Services”, Terradue is demonstrating that GEP consistently delivers on reusability, vendor neutrality, open science and optimized costs of operation.

Out of these outstanding capabilities, two GEP services have been tailored for deployment and ICT operations leveraging the EOSC-hub cloud infrastructure. Both services can be accessed via the EOSC Marketplace.

GEP High-Resolution Change Monitoring for the Alpine Region

This application provides access to the information layer generated by the DLR InSAR Browse High-Resolution service, running systematically on Sentinel-1A and B data.

Every six days, it delivers co-seismic interferograms at 50m resolution and 25m pixel spacing. The data products visualization and download functions support rapid response to earthquakes occurring within the processing mask.

GEP EO Services for Earthquake Response and Landslides Analysis

The application provides access to on-demand processing services for interferogram generation, co-seismic displacement mapping, landslide rapid mapping and landslide displacement field monitoring with Sentinel-1 and Sentinel-2 data.

It supports the study of seismic hazards and the detection and measurement of terrain motion due to earthquakes and landslides.

These two GEP services have been announced at the launch event of the European Open Science Cloud (November 2018), provided by Terradue onto the EOSC Marketplace, where users will find a dedicated interface to place an order.

We are very excited to expand the reach of GEP towards the European Open Science Cloud, demonstrate federated Cloud Computing capacity, and enable collaborative work across scientists, including support for state of the art publication and sharing of research results.

More information at geohazards-tep.eu

Terradue Cloud Platform terradue.com/portal/ellip



HADDOCK helps scientists to look at the evolution of brain tumours



A brain tumor is an abnormal growth of cells inside the brain and tissues around it. Brain tumours do not spread from other parts of the body and their cause is still unknown today. Understanding the mechanism of tumour growth is important for medical treatment. There are many molecular interactions occurring on a cell membrane that can restrain expansion.

Shih-Che Sue at National Tsing Hua University, Taiwan, and Andreas Bikfalvi at Bordeaux University, formed an international team to research ways of holding back tumour growth.

In their recent study published in *Nature Communications*, Sue and his collaborators at Bordeaux University tested the interaction between the LRP1 molecule, a key element involved in cancer processes, and CXCR3, a membrane receptor that can regulate tumor expansion.

The team started by doing a cell experiment which reported a potential region in LRP1 recognising CXCR3. Then, Sue and his colleagues used the HADDOCK platform for molecular docking.

HADDOCK is a web portal that offers computational tools for biologists to model the structure of complexes of proteins and other biomolecules via a user-friendly interface. The computational tasks operated in the back-end of HADDOCK are powered by EGI's High-Throughput Compute service.

During the test, the team compared the derived energy of each docking model. The results showed that a domain in the LRP1 potential region – called cysteine-rich domain 26 (CR26) – had the highest possibility in interacting with CXCR3.

This recognition mode can be turned into a future therapy strategy. For example, by creating a molecular mimic based on CR26 structure, CXCR3 might be specifically recognised by the compound and CXCR3 becomes less active. Thus, the tumour expansion can be reduced.

Haddock is part of the WeNMR suite for Structural Biology.

Data sharing done right

Rob Baxter writes about EOSC-hub's recommendations to facilitate sharing and processing of sensitive data

Human health is a global issue. Disease knows no border. Today we have more digital health data than ever before, but sharing them across borders for the greater good of global research remains challenging. Public good must still be balanced against the privacy of individuals.

Making it easier to access and share research data of all kinds is the central principle of the European Open Science Cloud – and this includes health and other “sensitive” data that must only be shared under strict supervision and with the utmost care. Through EOSC-hub (and in EOSC more widely) we have an opportunity to lead the world on making sensitive data safely available for research.

In EOSC-hub we've just completed a set of draft recommendations for policies on data sharing (D2.8 First Data policy recommendations). In this we were lucky with our timing – both the EOSCpilot project and the EC Expert Group on FAIR data came out with excellent policy reports covering ethics, and the open and FAIR data agendas. We were able to take a “next-steps” approach and identify what practical actions EOSC-hub might take to advance both the FAIR agenda and the ethical sharing of sensitive research data for public benefit.

We adopted 11 key recommendations from EOSCpilot and translated them into 22 practical suggestions on data sharing across EOSC-hub, covering the spectrum of data from the open and public to the highly controlled and sensitive. These suggestions aren't final yet but are up for consideration by strategists and technical system integrators in EOSC-hub. This time next year will see the final versions published.

Right now the recommendations fall under three broad headings:

► Implement FAIR, taking a “Web first” approach to implementing the FAIR principles: data objects should be published on the Web in open, non-proprietary, machine-readable



formats, well described and referenced by resolvable persistent identifiers.

- Build technical expertise in safe data and safe settings, adopt the “Five Safes” principles of safe data, safe settings, safe projects, safe people and safe outputs, and work towards enabling continent-wide research that follows them.
- Support the wider development of ethical and information governance frameworks. In EOSC-hub we have a terrific opportunity to engage with a wide set of stakeholders, including social science and statistical data service providers, and the emerging EOSC governance function, to build a strong consensus and strong processes for cross-border research using sensitive data.

Over the coming year we'll chew these over as a project, figure out what's feasible, what we can do now and what will have to wait. Sharing sensitive data across borders properly will be difficult, but the prize is surely worth it!

Rob Baxter is Programme Manager at EPCC, the University of Edinburgh, and leads the EOSC-hub task on Data Sharing Policies





EUROPEAN OPEN SCIENCE CLOUD

The EOSC Portal

Accelerating the transition to open science and open innovation

Join as a provider now and reach new users

www.eosc-portal.eu



A large, abstract network graph serves as the background for the entire page. It consists of numerous small, semi-transparent grey dots connected by thin white lines, forming a complex web of triangles and quadrilaterals across the entire surface.

© eosc-hub.eu
Twitter: [@EOSC_eu](https://twitter.com/EOSC_eu)
in [eosc-hub-project](#)
✉ info@eosc-hub.eu



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

