EOSC-synergy

Full name

European Open Science Cloud – Expanding Capacities by Building Capabilities

List of beneficiaries

Participant No	Participant organisation name	Country
1 - CSIC	Agencia Estatal Consejo Superior de Investigaciones Científicas	ES
2 - LIP	Laboratório de Instrumentação e Física Experimental de Partículas	PT
3 - KIT	Karlsruher Institut für Technologie	DE
4 - PSNC	Instytut Chemii Bioorganicznej Polskiej Akademii Nauk	PL
5 - CESNET	Cesnet Zajmove Sdruzeni Pravnickych Osob	CZ
6 - IISAS	Ustav Informatiky, Slovenska Akademia Vied	SK
7 – KNAW-DANS	Koninklijke Nederlandse Akademie van Wetenschappen	NL
8 - JISC	Jisc Lbg	UK
9 – IRD/LEGOS	Institute de Recherche pour le Development	FR
10 - UPV	Universidad Politecnica de Valencia	ES
11- LNEC	Laboratório Nacional de Engenharia Civil	PT
12 - ACK-CYFRONET	Akademia Gorniczo-Hutnicza im. Stanislawa Staszica w Krakowie	PL
13 – FCT	Fundação para a Ciência e Tecnologia de Portugal	PT
14 – EGI.eu	EGI Foundation	NL
15 - INCD	Associação INCD	PT
16 - INDRA	Indra Sistemas	ES

Table of Contents

Where to find the tables most often mentioned in the text:

Abstract

EOSC-synergy extends the EOSC coordination to nine participating countries by harmonizing policies and federating relevant national research e-Infrastructures, scientific data and thematic services, bridging the gap between national initiatives and EOSC.

The project introduces new capabilities by opening national thematic services to European access, thus expanding the EOSC offer in the Environment, Climate Change, Earth Observation and Life Sciences. This will be supported by an expansion of the capacity through the federation of compute, storage and data resources aligned with the EOSC and FAIR policies and practices.

EOSC-synergy builds on the expertise of leading research organizations, infrastructure providers, NRENs and user communities from Spain, Portugal, Germany, Poland, Czech Republic, Slovakia, Netherlands, United Kingdom and France, all already committed to the EOSC vision and already involved in related activities at national and international level. Furthermore, we will expand EOSC's global reach by integrating infrastructure and data providers beyond Europe, fostering international collaboration and open new resources to European researchers.

The project will push the EOSC state-of-the-art in software and services life-cycle through a quality-driven approach to services integration that will promote the convergence and alignment towards EOSC standards and best practices. This will be complemented by the expansion of the EOSC training and education capabilities through the introduction of an on-line platform aimed at boosting the development of EOSC skills and competences. EOSC-synergy complements on-going activities in EOSC-hub and other related projects liaising national bodies and infrastructures with other upcoming governance, data and national coordination projects.

1. EXCELLENCE

EOSC-synergy envisions EOSC as a coordinated effort to provide an European, open environment for scientific data and related processing that promotes convergence of infrastructures and thematic services provided at national or European level. The consortium members, as leading institutions in the European Scientific and Technological landscape, understand the current challenges as opportunities to remove the obstacles that currently prevent an effective sharing and reuse of scientific data, enabling exploitation of the full data potential and harnessing the capacity of regional investments.

Realizing this support to the European Research Area requires coordinated actions to provide services to access, process and compute on research data in a scalable way at the European level, ensuring research data is broadly exploitable as a public good. For that, a key step is to develop strong pillars at the national level that will foster the adoption of EOSC as an ecosystem to support Excellent Science across Europe. EOSC-synergy is about strengthening the foundations of EOSC, by providing the appropriate integration and harmonization mechanisms at the level of national resources and services.

In this regard integration and harmonization are especially important points that will be addressed by promoting the synergy between European-wide Research Communities and scientific computing infrastructures to bridge the "digital gap" between EC countries with different readiness levels.

EOSC-synergy will expand the capacity and capabilities of EOSC by leveraging the experience, effort and resources of national publicly-funded digital infrastructures in a coherent way, therefore acting also as an incentive for national resource providers.

The synergy between the national infrastructures and the current EOSC ecosystem will contribute to the convergence and alignment towards EOSC common policies and practices and promote an open environment for data sharing and exploitation potentiated by new services, new users and new data. EOSC-synergy will join relevant multidisciplinary and thematic initiatives enriching EOSC and contributing to its uptake.

1.1 Objectives

The EOSC should materialize as a lightweight and flexible federation of distributed interoperable data and computing resources. That is, an innovative data and compute integrated infrastructure that responds to the problem of large scale processing of research data, unleashing the potential of collaborative science across Europe, and maximizing the return of investments in publicly funded e-infrastructures by enabling shared open access.

With this vision in mind, we define the following Objectives supported by Key Performance Indicators (KPIs), defined in Table 3 (page 24):

Objective 1 – Expand EOSC Capacity	
Open access to nationally-funded services and resource capacity, with the participation of key stakeholders from Europe and South America. Address the need of accessing and processing research data close to the scientific computing facilities in a scalable way, by providing standard mechanisms and tools for deployment, integration and operation of regional data and computing infrastructures federated at European level. *Related KPIs (Table 3): 1, 2 and 4*	
Sub-objectives	Rationale
O1.1 Expand EOSC capacity (research data, compute and storage) to support provisioning of thematic services through EOSC to a broader user group.	- Capacity expansion based on the integration of national resources and national services of broad interest.
O1.2 Streamline access to the generic national infrastructures and thematic resources.	- Simplified access to infrastructure resources. - Deployment of services on national resources, and/or multi-national scale up where necessary, as requested by the user.
O1.3 Seamless access procedures, policies and protocols across national resources.	- National alignment towards common policies and practices. - Harmonizing policies and alignment with EU directives in data privacy and cloud security.
O1.4 Support resource provisioning through the EOSC-hub Portal	- Making national services of wide interest easily findable, accessible and usable via the EOSC-hub marketplace.

Objective 2 – Building EOSC Capabilities	
Addressing multidisciplinary research areas, by "connecting" production ready discipline-specific services and data, and scaling up their infrastructure capacity (in connection with Objective 1) sharing the benefits of these services and serving an increasing number of national/European user communities. *Related KPIs (Table 3): 3, 6 and 12*	
Sub-objectives	Rationale
O2.1 Provision of technical and operational integration tools to facilitate the inclusion of new thematic/community services in the EOSC Portal, in collaboration with the EOSC-hub.	- Easier technical integration of community services in EOSC. - Reduced time to market via ready-to-use solutions based on interoperable and reliable building blocks.
O2.2 Streamline the processes to support turning data into value (Data Productivization) in EOSC	- Decrease the time-to-product in the development of innovative data products
O2.3 FAIR integration of National Data Repositories in EOSC	- Sharing of national data and uptake of FAIR data practices.

Objective 3 – Foster EOSC services integration and promote quality	
Ensure that a wide range of present/future services in addition to the current service and resource portfolio for the project can be continually validated and integrated into EOSC by providing generic tools and the human networks necessary to create, integrate and validate new services into EOSC ecosystem. These tools will be applicable to a wide range of services and resources, both generic and specialized ones like data indexes, scientific portals, processing workflows, and integration of virtual research environments. *Related KPIs (Table 3): 4, 5, 6 and 8*	
Sub-objectives	Rationale
O3.1 Define processes, documentation and tools as applicable to facilitate and automatize the onboarding of additional providers, including lightweight certification and service quality auditing procedures.	- Implement a quality-driven integration process for new providers that promotes adherence to EOSC standards. - Enable service quality, conformance and compliance to be assessed. - Facilitate the onboarding of additional providers.
O3.2 Foster the integration of services in EOSC by implementing an automated Software Quality Assurance (SQA) validation mechanism, harmonized with common SQA standards and best practices.	- Provide a complete software management lifecycle for EOSC services. - Promote the uptake of quality standards and best practices applied to services, reducing issues and improving maintainability. - Improve services quality and reward adherence to EOSC standards through an EOSC SQA qualification/stamp.
O3.3 Bridge with the project approved in INFRAEOSC-05 subtopic (c) (FAIR) Implement processes, guidelines and tools for data provider nodes adopting FAIR principles.	- Propagate the policies developed in INFRAEOSC-05 subtopic (c) (FAIR) in the participating countries. - Definition of a FAIR data baseline implementation. - Better understanding of FAIR principles and requirements. - Ease the path towards the adoption of FAIR data practices.

Objective 4 – Promoting EOSC policy harmonisation	
Make possible the sharing of nationally funded services and resources by international communities. This will require the evolution and alignment of existing national policies for funding, provisioning and access to services and resources. In cooperation with the other projects approved in the call, generate the appropriate information exchange mechanism that will foster synergies at the European level, with an impact on the harmonization of country policies in the context of EOSC. *Related KPIs (see definitions in Table 3): 7, 13, 14 and 15*	
Sub-objectives	Rationale

O4.1 Ensuring that EOSC policies and practices are propagated to the national level. Special emphasis will be given to policies and practices that address the needs of different EOSC stakeholders.	- Consider specific requirements of national funding agencies related to national roadmaps, and national organizational and regulatory frameworks. - At the national level, focusing on user communities that are part of national research roadmaps.
O4.2 Provide recommendations aiming at evolving national policies and practices according to the needs of Open Science practices and international research collaborations	- Recommendations need to be identified by comparing the different landscapes at national levels. - Selection of the national best practices and strategies that will be considered more viable and more effectively applicable across the countries in the project scope.
O4.3 Contributing to the national and international organizational coordination and alignment in funding and provisioning of services in the EOSC scope.	- Facilitating, through compatible policies and operational conditions, the access, sharing and exploitation of nationally funded services and resources by international user communities in the context of EOSC.

Objective 5 – Develop the EOSC Human capital	
Establish a skills development framework that promotes EOSC adoption by the research communities (e.g. training and adoption of common interfaces, standards and best practices). Incorporate online training as a new full-fledged component of the EOSC Portal, facilitating access to high quality technical information and tailored training materials. *Related KPIs (see definitions in Table 3): 9, 10, 11 and 12*	
Sub-objectives	Rationale
O5.1 Expand human capacity through a training program aiming at promoting EOSC services. Creating the necessary skills that stakeholders (infrastructure owners, scientific users, software developers and service integrators) need to in order to become active in EOSC. Provide policies & guidelines & certifications to train the trainers.	- Easier access to better training for all EOSC stakeholders. - Reduce the learning curve to acquire the skills to become active in EOSC. - Expand the EOSC training network.
O5.2 EOSC Online training: create a MooC based, central point in the EOSC Portal for training.	- A central point for EOSC training courses based on a sustainable platform for massive online courses.
O5.3 Provisioning of a user self-deployable training infrastructure to foster EOSC adoption.	- Simplify hands-on training by offering a training infrastructure as a service. - Expanding EOSC portfolio with specialized training oriented services

1.2 Relation to the Work Program

EOSC-synergy will help addressing the challenges described in INFRAEOSC-05 subtopic (b) of this call with specific activities designed to facilitate, incentivize and boost the adoption of EOSC by different stakeholders with a special focus on: Infrastructure Managers, Researchers & Research Infrastructures, National Policy makers, Software developers and Service integrators.

EOSC-synergy brings together key research institutions in Europe sharing a strong background and experience in delivering distributed computing and data services to researchers from practically all scientific domains.

The first step towards convergence in federating data and services is interfacing the infrastructure resources themselves. In EOSC those resources will be provided by the large, well-established national e-infrastructures involved in this proposal (see distribution map in Figure 1), several of which are already involved in initiatives that also aim at federation and convergence at national or regional level.

Therefore EOSC-synergy is in a good position to contribute and push forward the interfacing and convergence within the EOSC ecosystem, and at the same time providing a substantial expansion in capacity. Section 4 provides a detailed overview of capacity in terms of resources, national repositories and services that the consortium partners bring in to support the project activities.

Several countries involved in the project (Spain, Portugal, Czech Republic, Poland and Slovakia) are planning in the upcoming 1-2 years substantial investments in infrastructure. The source of funding for these investments is also very diverse: from user community funding, to research performing organizations funding, or national investment, some of them articulated via EU structural funds. Therefore the project is in a good position to leverage those investments with beyond the state of the art IT technology, and analyse them in the light of cost-efficiencies and synergies arising from the possibility of sharing those infrastructures at the EU level.

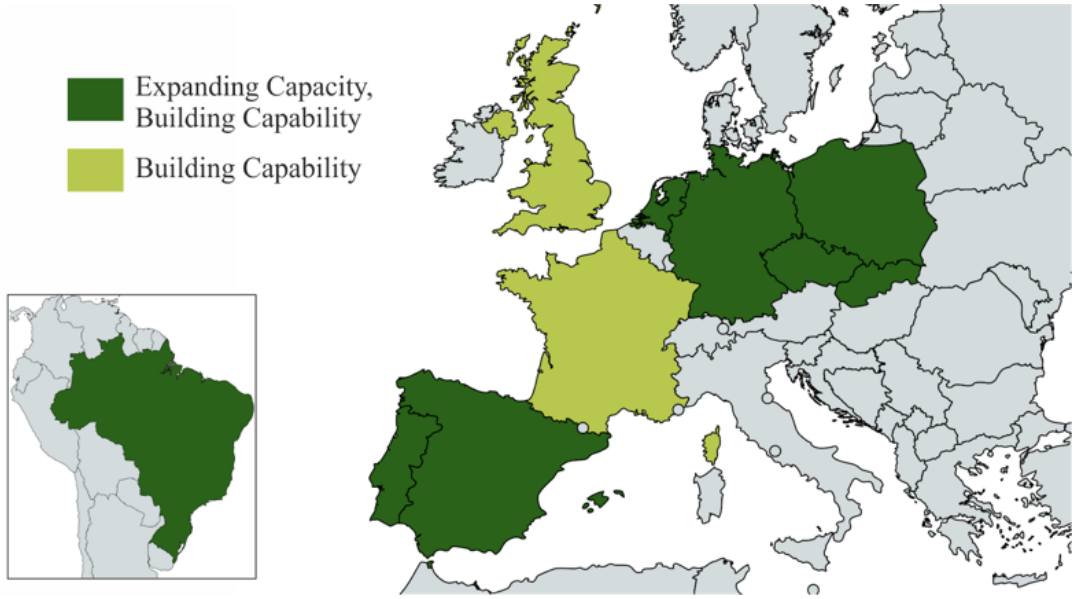


Figure 1. Geographical view of the EOSC-synergy consortium members

1.2.1 Cooperation with EOSC-hub

Regarding delivery of horizontal and core services, we will work in cooperation with the project EOSC-hub (EINFRA-12-2017), closing the existing gaps in the harmonization activities, best practices recommendations, and technology and standards follow-up. The project includes a Letter of Support from the EOSC-hub project, detailing the areas of cooperation between both projects, and which are summarized here

- EOSC-hub will offer a mechanism to promote access and re-use of the key exploitable results to EOSC stakeholders, by building a joint service and resource portfolio that includes EOSC-synergy national services and certified data repositories in the EOSC-hub Marketplace.
- Setting up and implementation of a joint events and communications programme aiming at harmonizing and strengthening the projects' innovation management plans.
- EOSC-hub will promote EOSC-synergy services via the EOSC Digital Innovation Hub, the EOSC stakeholder forum and other relevant engagement channels.
- EOSC-hub is also committed to provide services and resources for training.

1.2.2 Contribution to Service harmonization

Facilitating the integration of horizontal services is a key activity from the point of view of innovation. In scientific discovery the difference between standard and excellence science often relies on the possibility of accessing cutting-edge technological IT services. In order to foster the integration of those services in the EOSC Portal, EOSC-synergy proposes a *quantum leap* with respect to the state-of-the-art in Quality of Service management.

In this context, we will address specific activities in the area of automation of service and software quality assurance to promote a harmonized environment for service integration that covers the full lifecycle of the services for EOSC. In particular, we will address this challenge by proposing an EOSC Service Integration platform. Verification of compliance with standards, harmonization of procedures, promotion of FAIR data uptake, Service Level Agreements, certification and compliance with EU regulations, and interoperability at all levels are part of the work programme described in the relevant work package on Fostering Service Integration in this proposal.

1.2.3 Expanding EOSC Human Capacity

All these solutions will be tested by real user communities that are already heavy consumers of resources at the national level, coming from areas identified as of social importance such as Earth Observation, Genomics or Climate Change. The consortium includes national publicly funded digital infrastructures across Europe that will integrate resources, support the project activities and provide a sustainable added value to the integrated services.

The development of skills to foster EOSC adoption requires a comprehensive approach that needs to cover the technical, the policy aspects, and the interaction with the national educational programs. EOSC-synergy aims at providing the means to support the deployment of a training platform using a suitable and sustainable Open

Source solution. The policy aspect will be covered by the development of guidelines and best practices in the creation of content, with a special focus in training the trainers. Last but not least, we plan activities to interact with Universities via the consortium partners, in order to analyse the possibilities of mutual feedback regarding the creation of content with an appropriate level of quality.

1.2.4 Alignment of national policies

The proposal involves the main e-infrastructure providers and top-level research performing organizations in Spain, Portugal, Germany, Poland, France, UK, Netherlands, Czech Republic and Slovakia. These institutions have in their respective countries the official mandate to participate in the EOSC and represent the country as national initiatives. Therefore joint activities related to harmonization and alignment of practices requiring national engagement will be facilitated through the national contacts.

The alignment of policies would borrow key elements of the **European Interoperability Framework** (EIF) as a foundation for the work, by considering the overlap of national policy frameworks, particularly those applicable to the Legal, Organisational and Semantic interoperability layers.

It would also ensure policy alignment according to the twelve key principles of the EIF: (1) Subsidiarity and proportionality, (2) Reusability Inclusion and accessibility, (3) Administrative simplification, (4) Openness (5) Technological neutrality and data portability, (6) Security and privacy, (7) Preservation of information (8) Transparency (9) User-centricity, (10) Multilingualism, (11) Assessment of Effectiveness and (12) Efficiency

The project will facilitate - through compatible policies and operational conditions - the access, sharing and exploitation of nationally funded services and resources across international user communities in the context of EOSC, addressing Service Level Agreements, seamless access policies, user support, certification and compliance with the EU GDPR and with the EU NIS Directive regarding network Cloud Security.

The integration of national data repositories of broad interest into EOSC is an opportunity to implement and promote the uptake of FAIR data practices across national scientific communities. The proposal addresses this challenge with dedicated activities to integrate selected national repositories of broad interest following FAIR principles, and providing incentives for the uptake of FAIR data practices via the deployment of pilot catch-all repositories at national level. This activity will build on the policy recommendations coming via the project approved in INFRAEOSC-05(c).

1.3 Concept and Methodology

1.3.1 Concept

The two concepts underpinning this proposal are Capacity Expansion and Capability Building. The use of stakeholder knowledge in the conceptual design is graphically shown in Figure 2.

Capacity Expansion concept

Capacity Expansion refers to both enlarging EOSC with a critical mass of computing and storage resources, and fostering service integration in the EOSC, in such a way that those resources and services can be seamlessly exploited.

Service integration in EOSC should be embraced by the infrastructure managers as a strategic commitment and must be based on a clear understanding and evaluation of the benefits and responsibilities of joining EOSC. Incentivizing service integration requires lowering the adoption barriers by simplifying processes and providing comprehensive and yet easy to follow policies and procedures backed by access to adequate technical documentation, training, deployment guidelines and support.

EOSC-synergy builds on, and complements the EOSC-hub project activities by fostering the uptake of the EOSC core functions and horizontal services at national level. In particular it targets the need for a higher grade of automation in service certification and integration. In this sense, it proposes the definition of processes, documentation and tools as applicable to automatize as much as possible the onboarding of additional providers during the lifetime of the project, including lightweight certification and service quality auditing procedures both for Software and for Services.

Since providers are free to choose the components that comprise their present and future infrastructures, the process of integration in EOSC must be lightweight supporting a wide range of well-established interfaces in-line with the infrastructures expectations, thus reducing adoption barriers and simplifying the technical integration.

To provide the necessary level of integration at the infrastructure level we will use state-of-the-art resource management tools, able to interact with the native interfaces of the most used Cloud Middleware Frameworks, and also with batch-oriented infrastructure such as Grid and HPC resources

The heterogeneity of data storage facilities is even bigger. In order to provide an harmonized access, we will exploit tools capable of performing data transfer and synchronization supporting multiple protocols with two main objectives: simplify data access from the services and user perspective, and provide an abstraction layer enabling services to better cope with the infrastructures heterogeneity and technology evolution.

An additional challenge to reach the desired harmonization level is being able to integrate storage facilities with the standard Authorization and Authentication Infrastructure (AAI) as recommended in the AARC Blueprint.

AAI integration is a particularly interesting concept as it allows federating access to centralized storage facilities, which already exists at the national level. As an example, federating access to the Helmholtz Data Federation in Germany, with pools of resources belonging to other research organizations.

We will leverage knowledge on AAI and infrastructures federation to facilitate the desired harmonization.

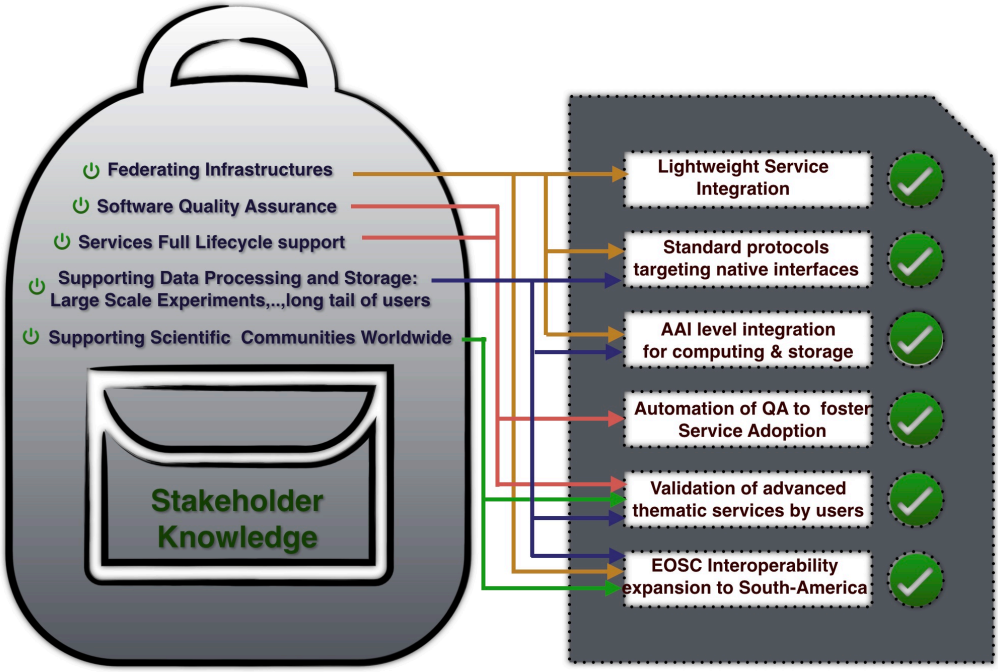


Figure 2. Usage of stakeholder knowledge background in the project conceptual design

Capability Building concept

EOSC Capability means "connecting" production ready discipline-specific services and scaling up their underlying infrastructure capacity, so that an increasing number of national/international user communities can be served. Connecting includes policy, technical and operational integration, so that new thematic services can become part of the EOSC catalogue (in collaboration with EOSC-hub).

EOSC-synergy will introduce new thematic services in several domains expanding the EOSC catalogue and contributing to raise the awareness of EOSC near the related research communities. Furthermore these services will act as validators of the proposed integration approach and as examples for further services to become integrated.

To this end we have selected several thematic services in scientific areas (Earth Observation, Biomedicine, Astrophysics and Climate Change – see Table 11 for complete list and added value for EOSC) providing open research data and services of interest that would benefit from a wider availability. These services have been selected for their potential of attracting further users and interest in EOSC.

During the project lifetime the communities will progress towards the definition of best practices for the adoption of common EOSC guidelines, tools, interfaces and services. This includes the promotion of incentives for the uptake of certification and FAIR data practices in coordination with other activities in the call.

Another important aspect to strengthen the communities is increasing the capacity, performance, reliability and/or functionality of these thematic services through their integration in EOSC. This is especially important to increase the number of users of these thematic services substantially.

In the framework of the capability expansion program, EOSC-synergy proposes activities to promote interoperability with non-European data providers and cloud infrastructures. This activity addresses the recurring problems and criticisms detected across the EU according to which Europe is capable of producing large amounts of data, but often that data is exploited elsewhere. EOSC-synergy will contribute to close this competitiveness gap by starting the expansion of EOSC to include resources beyond Europe.

To help in this endeavour EOSC-synergy will seek the alignment of the policy and technical aspects required to enable European researchers to exploit data that is generated and/or stored in other continents. For this the methodology and work program builds on the background and collaboration history of the Iberian region with South America.

The capacity expansion to South America is based on interoperating EOSC with the Brazilian Research Cloud operated by the Brazilian National Research Network (RNP)[1]. This is an important addition to the EOSC ecosystem in terms of future exploitation. Notably, the BELLA underwater optic fibre cable is expected to be operational in 2020, and will increase the capacity of data transfer between Europe and South America to unprecedented levels.

Our capability expansion program envisions providing access, through EOSC, to services developed by the Universidade Federal de Campina Grande (UFCG) for EMBRAPA (Brazilian Agricultural Research Corporation). This activity implies interfacing with the LANDSAT database, the world's largest collection of space-based moderate-resolution land remote sensing data[2]. In the area of Astrophysics, the project aims at providing access and exploitation services through EOSC to the data stored by the LAGOS experiment[3].

Therefore, having EOSC ready at the level of services and interoperation, to exploit this new data transfer e-infrastructure, is of paramount importance to consolidate the role of EOSC at international level, and to promote it as a hub for data access, exploitation and discovery.

Finally, the project will contribute to the development of the EOSC human capital through the integration in the EOSC Portal of a platform dedicated to Skills Development. In this context we propose a central point for EOSC online training, the adoption of massive online courses technology to reach efficiently the intended audiences, and an expansion of the training network by training the trainers.

1.3.2 Methodology

EOSC-synergy will support Capacity Expansion and Capability Building by interoperating existing national resources, and enabling the coherent integration of added value services, exploiting the federated capacity.

1.3.2.1 Technical integration of national and thematic resources

National infrastructures and thematic resources are tailored to the user communities owning those resources, or the national roadmap strategies when it is about supporting a wide range of users. Most of this capacity has been

[1] Brazilian institutions will participate in the integration of services at consultancy level not as funded partners.
[2] *https://landsat.usgs.gov*
[3] *http://lagoproject.org*

built without the larger context of EOSC in mind. The integration with EOSC requires normalization to the EOSC core services, such as AAI, Information System, Monitoring and Accounting.

Building on the experience and technological advancements driven in past projects and by industrial solutions, we believe that the integration with EOSC needs to be structured along organisational levels, rather than the classic approach via technological categories (e.g. storage, network, compute). The integration of Capacity will be driven on several levels as in shown in Figure 3.

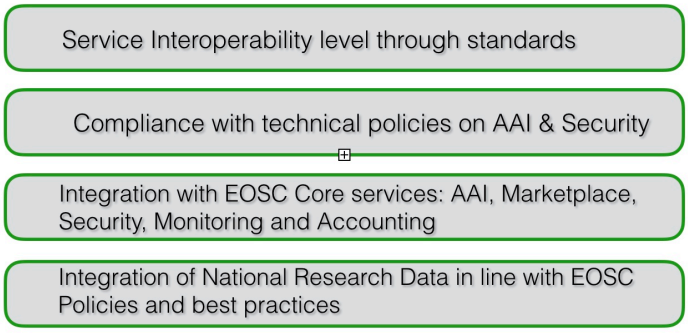


Figure 3. Organizational levels to structure the Capacity integration activities in EOSC-synergy

Leveraging the available capacity requires being able to interoperate between services. While it is easy to mandate one interface for one service class, this is not feasible, because existing users already rely on their interfaces in production. Therefore our methodology to leverage capacity will focus on integrating and promoting Open Source solutions that support standard protocols. This will lead to an overall enhancement of the EOSC services sustainability and usability.

1.3.2.2 Addressing Software & Service Quality Harmonization

Following the recommendations of the *EOSC 2nd High Level Expert Group*[4], software sustainability and service quality certification are of paramount importance to build a trustable ecosystem between service providers and infrastructure managers. Recognition to software developers and service integrators is also fundamental to guarantee that the academic and industrial European software development communities get engaged with EOSC. Our methodology will also work towards the implementation of an *EOSC-ready* label for services.

Service building blocks need to be mature software components, which means that from a quality standpoint EOSC needs to incorporate good practices to guarantee such maturity. Initiatives such as the Software Sustainability Institute[5] define sustainability in this context as software, which is easy to evolve and maintain, fulfils its intent over time, survives uncertainty, supports relevant concerns (Political, Economic, Social, Technical, Legal, Environmental). A baseline for software maturity, that helps promoting sustainability in EOSC is therefore of paramount importance.

[4] See EOSC HLEG documents in https://ec.europa.eu/info/events/2nd-eosc-summit-2018-jun-11_en
[5] https://danielskatzblog.wordpress.com/2016/09/13/defining-software-sustainability

Figure 4 describes the basic workflow of our approach. Fostering the integration of services in EOSC relies on a state of the art automation mechanism to speed the validation and certification of those services. EOSC-synergy aims to introduce an automated software quality assurance as a service (SQA-as-a-service) process that will leverage developments on continuous integration and continuous delivery from the INDIGO-DataCloud consortium. The SQA-as-a-service will assist in the steps necessary for service development, delivery and integration while facilitating the services maintenance and enabling conformance validation.

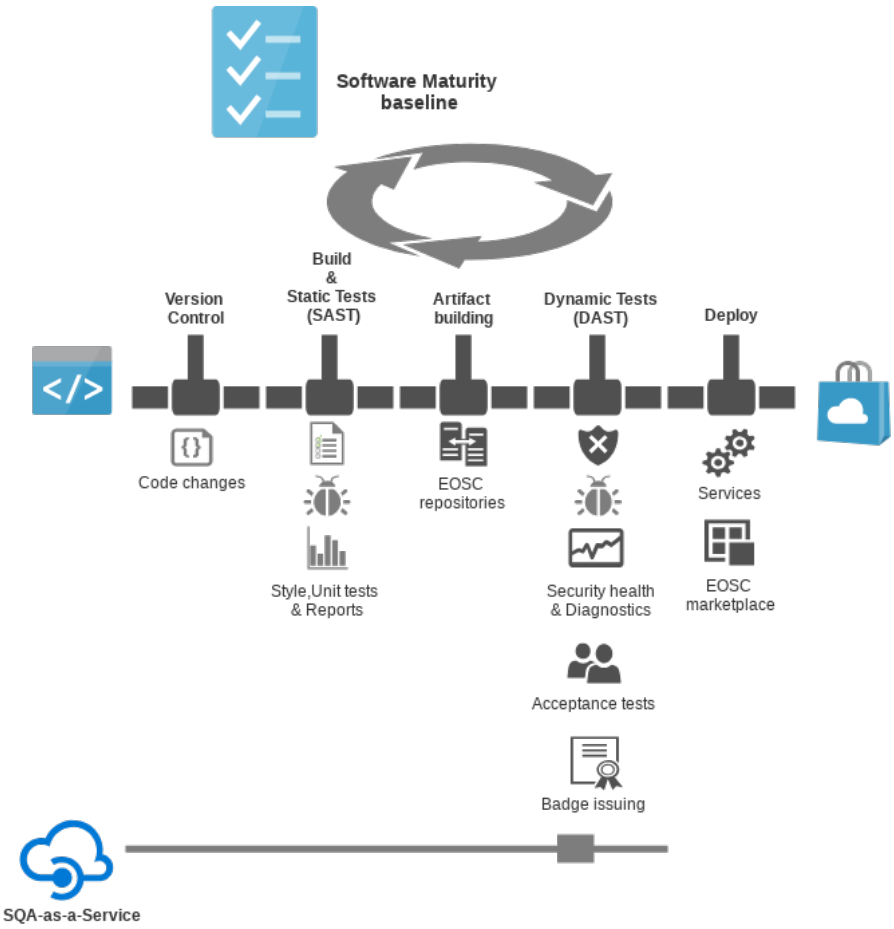


Figure 4. Fostering integration of software sustainable quality services into EOSC

1.3.2.3 Expanding EOSC Research Potential and Capabilities

Thematic services are frequently adapted to very specific needs of the scientific communities that built them. The targeted global user communities are supported by dedicated capacity (computing, network, and storage) provided by national or regional funding. This created sets of disconnected islands at the European level and results in a lack of harmonization that limits exploitation opportunities and collaboration potential.

In order to overcome the current obstacles, the project devises a set of technical and operational activities to simplify the coherent integration of national services in EOSC, complementing the activity of EOSC-hub at

national level. The project will identify mature "EOSC-relevant" national services, and provide the resources for integration, validation and scaling up making them available at European level.

In Figure 5 we show the methodology workflow to foster integration of new thematic services:

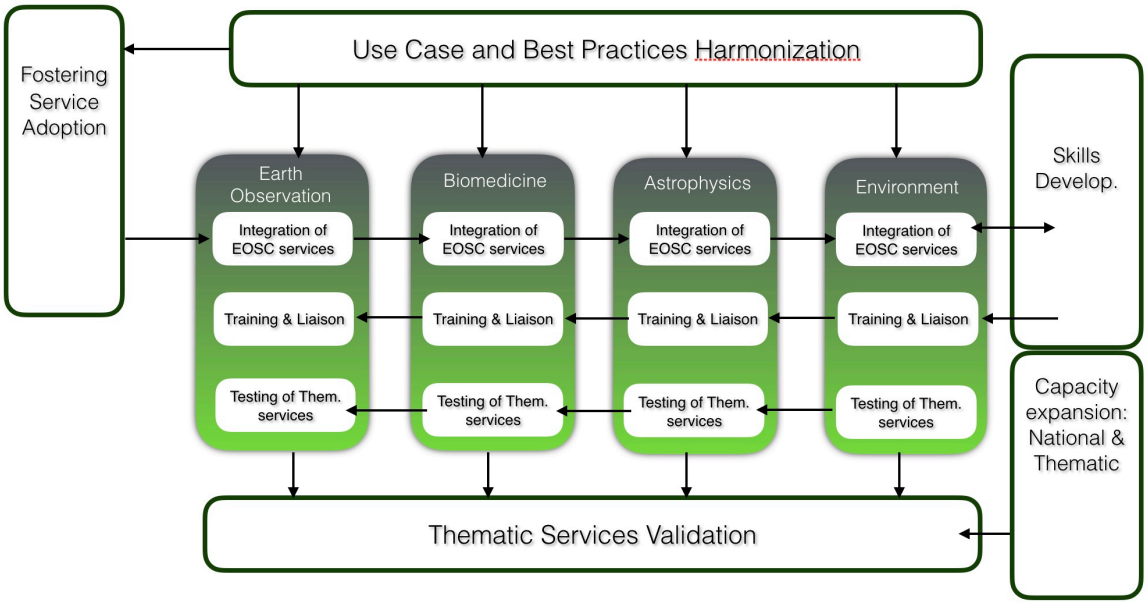


Figure 5. Methodological approach to Capability Expansion in the global framework of the project

1.3.2.4 Building Human Capacity

At the training level, EOSC-synergy aims to smooth the learning curve and reduce learning time to become skilled and active in EOSC. Our approach incorporates online training as a new fully-fledged component of the EOSC Marketplace, facilitating access to high quality technical information and tailored training materials. This approach is complemented by a self-deployable training infrastructure that will provide support for massive open online courses and hands-on training capabilities. The final objective is to improve training practices, enhance content quality and make EOSC training easily available to a much wider audience.

The improvement in quality will be achieved by training the trainers as content developers, as well as by cooperating with universities to share training materials of interest. To this end, parallel national activities will establish links with Universities and other educational institutions that are focusing on cloud computing and data science on their programs. This bi-directional collaboration will enrich the University programs with Open Science baseline knowledge, and will be an important channel to receive feedback on the tutorials methodology. This will allow developing EOSC-related skills, while increasing awareness among students.

1.3.2.5 Raising EOSC awareness at the level of national policies

At the policy level EOSC-synergy proposes an integrated approach liaising with national authorities and relevant initiatives to push the uptake and harmonization of EOSC at national level across the participating countries.

National infrastructure roadmaps are the usual way for countries to define and design their scientific strategy. Propagating EOSC at the national level also implies working in the context of those national roadmaps. In this respect our approach is based on the implementation of surveys at the national roadmap level to first understand the situation of those infrastructures with respect to EOSC integration. This includes a prospective of the compliance towards EOSC and FAIR practices at national level.

For this last point we will bridge with the project approved in subtopic (c), in particular by proposing activities to practically implement project recommendations (see Figure 6).

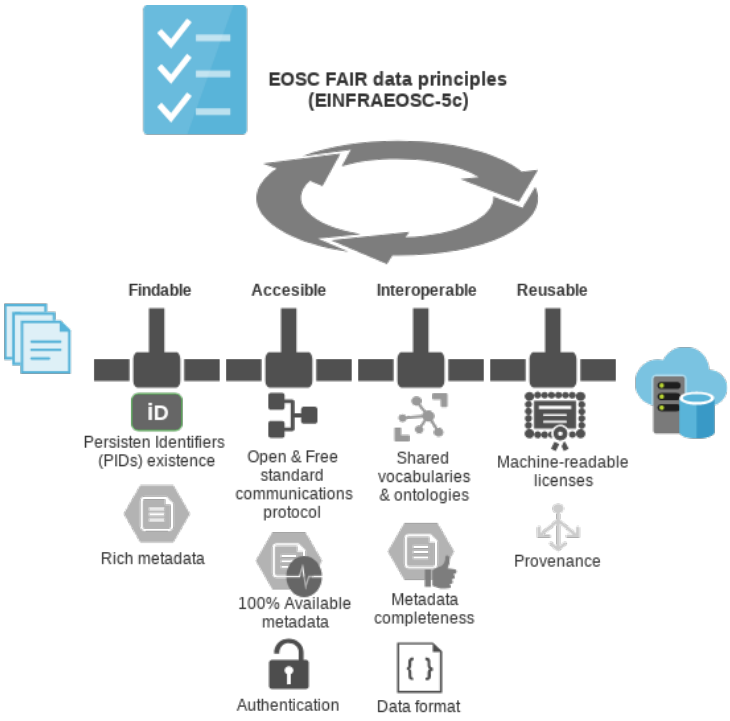


Figure 6. Supporting and fostering data FAIR adoption in cooperation with EINFRAEOSC-5c

These analysis needs to be followed by an identification of gaps in the national policies, and also by the identification of successful best practices that some countries might have implemented already, and that may serve as a baseline for others to build upon.

In this context, we will bridge with the governance project approved in INFRAEOSC-05(a), and other projects in the EOSC ecosystem, to align and expand policy recommendations at the country level. Our policy activity will provide the means to identify potential issues in the harmonization in dialog with national infrastructures and communities, and push for the alignment of national policies and practices including access to services, data protection and privacy, infrastructure security, leveraging public investments, alignment of practices and service management.

1.4 Ambition

EOSC-synergy will ensure the coordination of EOSC relevant national and thematic initiatives through a process that will promote their convergence through innovation contributing to expand the EOSC state-of-the-art. Table 1 shows how the project ambition goes beyond the state of the art.

EOSC-synergy joins organizations that deliver advanced scientific computing, data and network services at national level and have long-standing experience in developing, integrating and supporting services for massive data processing.

In many cases, these organizations have mandates from their governments to participate in research infrastructures at national and international level. These organizations have the competences to expand the state-of-the-art and are highly motivated to contribute in establishing EOSC as a ground-breaking initiative. They share the vision of federating data and compute infrastructures in Europe and elsewhere to support innovative added value services that enable the exploitation of scientific data.

EOSC-synergy will address the convergence at several levels in close interaction with other relevant projects and initiatives. Through the connections of our partners with projects, infrastructures and initiatives such as EOSC-hub, EGI, EUDAT, INDIGO-DataCloud, GÉANT, AARC, and OpenAIRE, we are in a privileged position to become a focal pointing facilitating the coordination and alignment towards EOSC. Additionally, a strong participation of NRENs will complement the infrastructure providers and promote convergence in areas such as AAI, security and data access.

Table 1. EOSC-synergy ambition beyond the state of the art

Area	Current Status	EOSC-synergy foreseen development
EOSC policy and coordination	Performed via the participation of individual organizations in EOSC projects and related European infrastructures. Currently fragmented and missing liaison between EOSC and the member states.	- Extend EOSC coordination to the national level within the participating countries. - Promote the alignment of policies and practices at the national level. - Join NRENs, resource providers and user communities. - Liaison with national authorities. - Bridging to national users.
Cloud IaaS	EOSC capacity is limited by the resources committed in current EOSC-related projects.	- Streamlining of policies and procedures as a means to simplify and promote integration. - Enlarge capacity with new resources and new providers. - Exploiting new public investments using European funds. - Extend providers to South America.
Cloud and data interfaces	Proliferation of specifications and interfaces many of which not fully interoperable. Federation and combined usage of	- Converge on a subset of well-proven and reliable interfaces. - Improve federated access.

	resources requires effort. Users and integrators do not know which services and interfaces work together.	- Identify interfaces and service combinations that work together.
Services quality	Current focus mainly on services integration and operation. Quality is mainly delegated to the developers and providers.	- Establish an SQA baseline for EOSC services. - Provide SQA for services enabling simplified software - maintenance and conformance validation. - Definition of a baseline towards an *EOSC-ready* seal for services.
Thematic services	Limited set of services provided within ongoing projects. Considerable effort and expertise required for service integration.	- Provisioning of well-proven deployment scenarios and recipes to drive service integration. - Tools to facilitate usage and service integration. - New thematic national services enlarging the EOSC catalogue.
Data repositories	Few data repositories in EOSC. FAIR principles still poorly understood.	- Enable conformance verification - Liaison with other data initiatives. - Promote FAIR data practices at national level. - New data providers across Europe, and expansion to global data providers.
Training	Lack of a global EOSC wide coherent approach to training. Training materials scattered. Lack of funding for training in situ.	- Add training as fully-fledged EOSC Portal component. - Provide guidelines and recommendations for high-quality training materials, oriented to training the trainers. - Provide infrastructure to support self-learning. - Provide infrastructure for massive open online courses (MOOCs).

2. IMPACT

The strongest impact of the EOSC-synergy project will be a measurable increase of the number of resources, services and data repositories offered to researchers through EOSC, paired with concrete, contributions to the EOSC governance structure from national perspectives and a new channel to support the build-up of EOSC human capabilities.

EOSC-synergy's impact will be achieved by the implementation of the project's mission to:

- **Expand capacities** by adding compute, storage and data resources to the EOSC ecosystem, and
- **Build capabilities** by augmenting the EOSC offer with new Thematic Services targeted at relevant emerging communities.

This will be complemented with actions directed at: harmonizing national and European policies, streamlining the integration of new resources and services, promoting the uptake of FAIR practices, and strengthen the human capabilities within the EOSC landscape.

The following sections will explore the details behind this ambition.

2.1 Expected impacts

The project proposed objectives, when implemented, will generate impact at the following levels, whose corresponding KPIs can be found in Table 3 (section 2.1.1):

At the level of capacity

EOSC-synergy will make it simpler for new infrastructure providers to federate resources in EOSC. This will be done by streamlining processes and procedures, establishing guidelines on the services to be deployed, and providing close support to national providers facilitating integration in EOSC.

EOSC-synergy will have an impact on EOSC capacity through the integration of new infrastructure resources. The project will open the path to the expansion of EOSC beyond Europe through the integration of research and academic cloud providers from Brazil, building on the value of EU-Brazil projects.

Related KPIs: 1, 2, 3 and 4

At the level of capabilities

EOSC-synergy will facilitate thematic service integration by providing a set of well-proven deployment scenarios that address common thematic service requirements. These scenarios will be made available as recipes to drive the integration of new services.

EOSC-synergy will augment the service portfolio of the EOSC Portal, by adding new thematic services aimed at diverse user communities.

Related KPIs: 3, 4, 5 and 6

At the level of data repositories

EOSC-synergy will increase the adoption of FAIR data management principles by providing. This will be done by bringing data repositories currently operating in closed silos into the EOSC ecosystem and, in the process, making them findable, accessible, interoperable and reusable.

EOSC-synergy will expand the EOSC open data portfolio with databases from the Brazilian Agricultural Research Corporation (EMBRAPA) and data from the LAGOS astroparticle physics experiment. Adding South American data providers to EOSC is an important step towards the integration of non-European data in EOSC.

Related KPIs: 2, 4 and 8

At the level of services quality

EOSC-synergy will increase software quality in the EOSC ecosystem through the introduction of an automated Software Quality Assurance as a service (SQA-as-a-service) process. This impact leverages the developments on continuous integration and continuous delivery from the INDIGO-DataCloud project.

EOSC-synergy will deliver transparency by defining a baseline towards an *EOSC-ready* seal for services as recommended by the EOSC high-level expert group. With the seal, users and infrastructure managers will be able to recognize the services that fulfil the SQA-as-a-service as a proof of conformance, while service providers will be rewarded and further motivated for the excellence of their services.

Related KPIs: 5 and 6

At the policy level

EOSC-synergy will increase the uptake of the EOSC vision at national level filling a major gap in the current EOSC setup by liaising with national policy makers, digital infrastructures and user communities.

EOSC-synergy will maximise the impacts of the projects approved in all INFRAEOSC-05 sub-topics by providing coordination with the countries in the consortium and open channels for dialogue with national infrastructures and communities, and by pushing for the alignment of national policies and practices including access to services, data protection and privacy, infrastructure security, leveraging public investments, alignment of practices and service management.

Related KPIs: 8, 10 and 13

At the training level

EOSC-synergy will significantly improve the EOSC training capabilities by providing an online training platform as a fully-fledged component of the EOSC Portal to facilitate access to high quality technical information and training materials. This will be complemented by a self-deployable training infrastructure, support for massive open online courses and self-learning capabilities.

EOSC-synergy will improve the sustainability and scalability of training for EOSC stakeholders by creating a training programme, support materials and quality assurance to assist trainers and educators in developing their own high quality online resources.

Related KPIs: 9, 10, 11 and 12

2.1.1 Contribution to expected impacts in the work programme

This section focuses on the contribution that EOSC-synergy brings to the impacts expected by the work program (Table 2). Table 3 lists the Key Performance Indicators (KPIs) used to measure the progress towards achieving those impacts.

Table 2 - Expected impact extracted from the call text (in italic) and contribution from the project

"Streamlining advice to policy making and funding authorities concerning the EOSC (...) "	
Contribution to impact	
EOSC-synergy will provide detailed insights about national policies, practices, roadmaps and strategies around funding, procuring, providing, accessing and sharing of services and resources in the countries covered by the project. This will help the policy makers and funding authorities to have a clear view on the state of the art of national infrastructures in the context of EOSC. EOSC-synergy will increase national and international coordination in funding and provisioning of services in the EOSC scope. EOSC-synergy make a contribution to the EOSC governance, in the shape of advice on Rules of Participation.	KPI 10 KPI 13 KPI 14 KPI 15
"Achievement of a sustainable support to the EOSC governance via the (...) federation of EOSC-related initiatives with capacity of delivering the EOSC core functions to a wide range of users across Europe."	
Contribution to impact	
EOSC-synergy will make a measurable contribution towards the delivery of EOSC's core functions by adding compute and storage capacity to the EOSC ecosystem, in parallel with augmenting the EOSC Portal service offer with mature thematic services for researchers. The scientific areas that will benefit the most from this expected impact are: Earth Observation, Biomedicine, Astrophysics and Climate Change. EOSC-synergy will federate additional data repositories with the EOSC Portal. EOSC-synergy will federate extra capacity and capabilities into the EOSC ecosystem through technical integration focused on the use of open standards, guaranteeing, where relevant, the adherence to the FAIR principles of data management. EOSC-synergy will improve the performance, reliability and functionalities of the newly federated thematic services by establishing quality baselines and a user-facing programme of service validation.	KPI 1 KPI 2 KPI 3 KPI 4 KPI 5 KPI 6 KPI 9 KPI 12
Piloting and establishing future (co)funding strategies and business/usage models to ensure long-term sustainability.	
Contribution to impact	
EOSC-synergy will identify gaps and successful best practices to ensure that recommendations are in line with the needs of open science practices and international research collaborations.	KPI 13 KPI 14 KPI 15
Wide adoption of processes and practices for optimal use of the resources ensuring open, fair and unbiased access of services through the EOSC hub.	

Contribution to impact	
EOSC-synergy will bridge with other relevant projects and activities to promote the harmonization and uptake of EOSC wide policies and best practices. EOSC-synergy will implement policies on operational security, incident response, acceptable authentication assurance and acceptable use policies. EOSC-synergy will facilitate the access, sharing and exploitation of nationally funded services and resources across international user communities in the context of EOSC. EOSC-synergy will define software quality assurance baselines for the thematic services federated in the EOSC hub. This will improve: i) the quality, reliability and maturity of software; ii) the reusability, findability and transparency of software; iii) the delivery and integration of services built from these software products	KPI 5 KPI 6 KPI 9 KPI 10 KPI 11 KPI 12 KPI 13 KPI 14 KPI 15
Progressive removal of technical and organisational barriers to ensure findability, accessibility, interoperability and re-use of research data.	
Contribution to impact	
EOSC-synergy will expand the EOSC capacity by adding FAIR-compliant data repositories to the EOSC Portal. EOSC-synergy will remove organizational barriers by disseminating expertise of federated infrastructure management and operation to national and thematic infrastructures and resources. EOSC-synergy will reduce accessibility barriers by implementing a common AAI solution to the newly federated thematic services. EOSC-synergy will reduce technical barriers by harmonizing the use of standards and enhancing interoperability. EOSC-synergy will integrate data providers in non-European countries widening the scope of research data available through EOSC.	KPI 1 KPI 2 KPI 4 KPI 7 KPI 8 KPI 9 KPI 10 KPI 11 KPI 12
Enhancement of FAIR data uptake and contribute to harmonisation of related policies in Europe and facilitate alignment with international initiatives on research data sharing.	
Contribution to impact	
EOSC-synergy will promote FAIR best practices on data management at national level through the policy harmonization work of WP5. EOSC-synergy will provide tools to assess the level of FAIRness of EOSC data repositories. EOSC-synergy will take data repositories currently isolated and locked in national infrastructures with no interoperability, and bring them to the EOSC ecosystem, making them Findable, Accessible, Interoperable and Reusable. EOSC-synergy will ensure that EOSC policies and practices are propagated to the national level, taking into account the requirements of national funding agencies, the needs of the national roadmaps and the constraints of national organizational and regulatory frameworks.	KPI 7 KPI 8 KPI 13 KPI 14 KPI 15

Table 3 - Key Performance Indicators. Base Target=BT, Stretch Target=ST

KPI	Target (BT/ST)
KPI 1 – Number of new data centres federated into EOSC	7/11
KPI 2 – Number of non-European resources added to EOSC	3/5
KPI 3 – Number of services integrated with the EOSC Portal	5/8
KPI 4 – Number of FAIR-compliant data collections	8/12
KPI 5 – Software Quality Assurance (SQA) mechanism in place	TRUE / na
KPI 6 – Number of services that went through SQA validation	10/20
KPI 7 – Collaboration in place with EOSC-related projects	4/8
KPI 8 – FAIR recommendations implemented	TRUE / na
KPI 9 – MOOC platform in production	TRUE / na
KPI 10 – Number of universities collaborating in skills development	6/12
KPI 11 – Number of tutorials uploaded to the platform	7/15
KPI 12 – Number of requests for online self-training	50/100
KPI 13 – Number of contributions coming from national roadmaps analysis	8/10
KPI 14 – Organization of meetings/workshops with national stakeholders	5/8
KPI 15 – Consolidated input documents to the EOSC Coordination structure	5/8

2.1.2 Enhancing innovation capacity

EOSC-synergy will make significant contributions to the EOSC service offer by adding services and data repositories targeted at a diverse number of domains, including environmental science, social science, medicine, astrophysics.

The impact of capability building will be bigger than just an augmented service offer because these resources, previously isolated in the niche of the communities that developed and managed them, will become available to a wider community of researchers, opening new possibilities for innovation. For example:

- The standardization of the services will enable creating processing pipelines combining data from different services, creating added-value products such as the integration of the water flooding (WorSICA) with toxicity analysis (MSWSS) and evapotranspiration (SAPS). With this, researchers can, for example, derive the collateral effects of the flooding in carrying out toxic products into water distribution or the short-time effects of water abundance.
- The benchmarking system for bioinformatic products in the genomic thematic service (OpenEBench) can also be used by disciplines, for example, to characterize their workloads.
- The g-Core service from INDRA is suitable for managing all types of data with properties similar to earth observation data. This includes, for example, climate data or astrophysics (LAGO).

Additionally, we can envision the following enhancements to the EOSC innovation capacity:

- The efforts aimed at national coordination will generate a momentum in the countries involved in the project that will lead to a greater harmonisation of policies and understanding of the benefits of EOSC. In the long-term, this will lead to more actions and more value generated locally from the EOSC vision.
- The project will expand the coverage and the capacity of EOSC. This will create future opportunities for all domains of science, nationally and at European scale.
- The introduction of an innovative quality driven process applied to the services integration will streamline the addition of new services and has the potential of changing the way scientific services are built and maintained by the communities.
- The creation of an on-line training platform for EOSC will open opportunities for diversifying and improving the training channels and materials including the new collaborations with universities and other entities.
- The project will create an environment that stimulates innovation by reducing barriers to adoption by users. This, in addition to an enriched EOSC catalogue will offer more opportunities for serendipitous discoveries.
- The integration of infrastructures and data providers in non-European countries will pave the way to the international expansion of EOSC opening access to data and services at the global level.
- EOSC-synergy will be instrumental in establishing the EOSC as a reference for data exploitation and analysis. In the long term, a better offer will enhance the EOSC status as a global hub for research, thereby generating more interest from both users and providers.

2.1.3 Strengthen the competitiveness and growth

EOSC-synergy will support Open Innovation in accordance with the broader European vision. Sources of innovation from the project will come from the areas of providing a high level of automation in quality of service, and service deployment procedures. Increasing the readiness and maturity level of marketplace services will have an immediate impact in terms of fostering the adoption of EOSC.

The creation of a full-fledged marketplace component dedicated to training has the potentially of becoming a key element to foster innovation, as it can be used and picked up by external stakeholders to provide their own material to develop human skills.

The project will also promote interoperability with non-European data providers and cloud infrastructures, and thus will further develop the European research innovation capacity, in terms of global data exploitation.

2.1.4 Contribution to environmental and societal impacts

The project addresses environmental and societal challenges in the areas of earth observation with applications to climate change and environmental policies. In particular, the project will integrate thematic services such as:

- Addressing the detection of the coastlines, coastal and inland inundation areas: WorSiCa service.

- The *Sand and Dust Storm Warning Advisory and Assessment System* of the WMO (World Meteorological Organization): SDS-WAS service.
- Mass-spectrometry analysis from the Research Centre for Toxic Compounds in the Environment at Masaryk University) in the Czech Republic: RECETOX service.
- Analysis of water network distribution with regards to the mitigation of hazardous events, by the integration of existing on-line analysis of toxics in drinking water supply networks with distribution network simulation of EPANET: MSWSS service.
- Analysis of ozone projections through the O3AS system to forecast the evolution on a long-term basis.
- Estimation of the Surface Energy Balance Algorithm for Land (SEBAL) and similar information for predicting the evolution of forest masses and crops, in collaboration with Brazil: SAPS service.

2.1.5 Barriers, obstacles and other framework conditions

The barriers that may determine whether and to what extent the expected impacts will be achieved are:

- **Political commitment**: At the moment, the EC is strongly supporting the EOSC vision and Member States are being engaged to collaborate and join forces. But the 30 months proposed for this project are a long time in politics. The consortium will work on influencing the public authorities and funding agencies to communicate the importance of EOSC and to collaborate in keeping the agenda high in the priority list.
- **National versus Transnational access policies**: Sustaining access provision of national services in EOSC, while national policies restrain access to international users: lack of a practical business model to support transnational access.

2.2 Measures to maximize impact

In this section, we identify an initial list of the expected results from the project and the key target groups to which these results may apply. We also describe a preliminary dissemination strategy aimed at making sure that these target groups become aware of the results and adopt them for maximum impact. Finally, we describe the communication activities required to disseminate these results to their target groups and to make the project known to the EOSC ecosystem. The main categories of project results are listed at the beginning of section 2.1. The groups to be targeted are:

- **Researchers (sensu lato)**: this includes individual researchers, as well as organized Research Communities and Research Infrastructures.
- **National e-Infrastructures**: the national providers of computing resources to research.
- **Policy makers**: Governmental institutions at local, regional, national levels, as well as the European Commission, European Parliament, Committee of Regions.
- **Service providers**: the maintainers of the services that the project proposes to federate in EOSC.
- **EOSC ecosystem**: projects that are collectively working towards building the EOSC. This includes: the other projects selected from the INFRAEOSC-05-b call, as well as the projects selected from the INFRAEOSC-05-a (EOSC governance) and 05-c (FAIR implementation) calls, and the related projects selected from the upcoming 2019 calls.

2.2.1 Dissemination and exploitation of project results

2.2.1.1 Exploitation of project results

The table that follows describes, for each expected project result, their corresponding IP owners, any foreseen or expected protection and licensing options. It also includes the targeted stakeholder groups that can be users/customers, the exploitation paths that we see as most prominent at this stage. The dissemination strategy for category of results is outlined in Table 5

Table 4. List of exploitable results

Result	Owners	IPR	Target stakeholders	Benefits	Possible exploitation path
Category: Research clouds federated in EOSC					
Roadmap for integration of national capacities into the EOSC	Participants of WP2	Public domain	National Infrastructures, Thematic Services	Capacity planning and prioritization of services integration.	Definition of national roadmaps and EC investments
Handbook for service integration	Participants of WP2	Public domain	National Infrastructures, Thematic Services	Streamlining of capacity integration.	National Clouds available through the EOSC Portal.
National clouds integrated in EOSC	Infrastructure resource providers	Public domain	National Infrastructures, Thematic Services	Optimization of National and EC investments. Open Access to resources.	
Category: Data, Thematic and Software services					
Thematic Services (Table 11 section 3.4)	Their respective developers	Open Access	Research communities and EOSC ecosystem	More analytic tools for the research community; more value add for EOSC	The services and repositories will be available through the EOSC Portal
Data repositories (Tables 9 and 10 section 3.4)	Their respective owners	Open Access	Research communities and EOSC ecosystem	More FAIR-compliant data available for research	
Category: Training platforms and training materials					
EOSC Training Platform	PSNC	Open source	Research communities, EOSC ecosystem, service providers	Enables online events and self-study learning in EOSC	The platforms will be integrated in the EOSC Portal

Hackathon as a Service platform	CSIC, BIFI	Open source	Research communities, EOSC ecosystem, service providers	Facilitates organizing and conducting the hackathons based on EOSC infrastructure	
How to guides for trainers	Participants of tasks 6.1 and 6.2	Open source	Research communities, service providers	Facilitates the creation of high-quality materials	Can be widely adopted in and outside the EOSC ecosystem
Accreditation scheme for training materials and trainers	Participants of task 6.2	Open source	EOSC ecosystem, Universities	Creates a check list for quality check of the training materials	
Category: Policy recommendations to EOSC					
Landscape analysis of EOSC-related national policies and practices	Participants of WP5	Public domain	EOSC ecosystem, policy makers	Understanding differences between national policies; the first step towards their harmonization	To be shared with the Infra-EOSC5a project as input for their work. Also of interest to countries not in the consortium and bodies such as e-IRG
Recommendations for the national alignment	Participants of WP5	Public domain	EOSC ecosystem, policy makers	Foster harmonization of national policies on resource provisioning, access and allocation	
Category: Certification and quality auditing schemes					
A common baseline for Software Sustainability	Participants of WP3	Public domain	EOSC ecosystem, service providers	Quality-assured and secured services are delivered through the EOSC Portal	Base for an EOSC wide quality driven process for services integration and maintenance.
Quality badge scheme for software components	Participants of WP3	Open source	Research communities, EOSC ecosystem, service providers	Motivates service providers to develop reliable software Gives feedback to users on the level of quality expected	Foundations for the establishment of an EOSC-ready stamp
Software Quality Assessment (SQA) for on-demand validation	Participants of WP3	Open source	EOSC ecosystem	Facilitates service integration in EOSC portal by automating validation procedures	Offered “as a service” through the EOSC portal

Framework for validating EOSC FAIR data requirements.	Participants of WP3	Open source	EOSC ecosystem, service providers	Facilitates data integration in repositories according to FAIR	Can be used as mandatory requirement for EOSC FAIR data validation

2.2.1.2 Dissemination of project results

Successful delivery of project results does not guarantee impact. To maximise the project impact on Europe's research landscape, the results need to be disseminated to their respective target audiences. Table 5 outlines the basic components of EOSC-synergy's dissemination strategy.

Table 5 - Approach to disseminate the project results

Category: Research clouds federated in EOSC		
Target	Dissemination objective	Dissemination strategy
Researchers (sensu lato)	Motivate the research community to use the clouds federated in EOSC	> Use infographics to show the scale of available resources > Write case studies describing successful use of these resources > Leverage EGI Foundation contacts in the research and service provider communities
Service providers	Motivate service providers to join the EOSC ecosystem using the added capacity	
EOSC ecosystem	Set the example of how to integrate clouds in EOSC for further capacity expansion efforts	> Promote the Handbook for integration > Share the Roadmap for integration of national capacities with the EOSC governance
Category: Data, Thematic and Software services		
Target	Dissemination objective	Dissemination strategy
Researchers (sensu lato)	Motivate the research community to use the data and services added to the EOSC	> Write compelling service descriptions for publication in the EOSC Portal and case studies describing successful use > Leverage EGI Foundation contacts in the research communities > Provide 'how to guides for each service > Promote the data repositories in the EOSC Portal with overviews of their content
Service providers	Motivate new service providers to add their products to EOSC	> Write case studies of integration, detailing the steps required, benefits for integration and examples of usage
EOSC ecosystem	Motivate the EOSC bodies to use these results as examples for further capability building exercises	

Category: Training platforms and training materials		
Target	Dissemination objective	Dissemination strategy
Researchers (sensu lato)	Encourage researchers to learn about services through the platforms	> Promote the creation of high-quality appealing contents through training. > Write case studies of successful usage (e.g. user journeys) > Provide an overview of the resources available (in the EOSC Portal)
Service providers	Motivate providers to use the platform to train users of their services	> Promote the 'how to guide' for trainers > Deliver a report with an overview of usage during the project span and key benefits of the platform
EOSC ecosystem	Demonstrate the benefits of the platforms to motivate the EOSC boards to promote them for further use	
Category: Policy recommendations to EOSC		
Target	Dissemination objective	Dissemination strategy
EOSC ecosystem	Motivate the EOSC governing bodies to consider the recommendations of the project	> Provide clear write-ups of the policy work, with schematic overviews > Establish a working collaboration with the INFRAEOSC-5a project
Policy makers	Contribute to the global discussion on service integration at the international level	> Promote the project policy results as white papers to be shared with national authorities and international forums, for example e-IRG
Category: Certification and quality auditing schemes		
Target	Dissemination objective	Dissemination strategy
Researchers (sensu lato)	Encourage users to assess the maturity of services through the badge scheme	> Liaise with the EOSC Portal Editorial Board to publish an overview of the different types of badges and what they mean
Service providers	Motivate providers to implement FAIR principles and the recommendations of the Software Quality Baseline from day one	> Provide a digested overview of the recommendations > Inform developers about the badge scheme
EOSC ecosystem	Promote the EOSC to use the Software Quality Baseline as a standard for validation	> Communicate the advantages of the baseline and the SQA-as-a-service to EOSC governing bodies

2.2.2 Communications activities

EOSC-synergy will define two types of communication activities:

- **Project-related communications:** These activities will focus on establishing the public image of the EOSC-synergy project, its identity and project website, as well as the appropriate communication channels. This activity also aims to provide communications support to the consortium to ensure a strong sense of community in the team, facilitate intra-project work and ensure that everyone in the consortium is aware of developments and speaks with the same voice. Project-related communication activities are described in Table 6.
- **Communication activities to support the dissemination & exploitation of project results**: Building on project-related communications, the communications team main focus will be to support the objectives defined in the dissemination plan (drafted in section 2.2.1.2) with specific actions. Table 7 provides an overview of the measures we can already envision (to be updated when the Dissemination Strategy is finalised.)

Table 6 - Measures to communicate project activities

ACTIVITY	Outputs
Objective: Define the image of the project	
Develop the project corporate identity	> Project logo, presentation and document templates to support consortium communications
Establish the key messages about the project and ensure consistency	> Project one-liner and pitch to each target audience > Generic project poster & flyer
Summarise the project in a short video	> 1 short (1 minute) video about the project and its goals > To be posted on the social media feeds of 10 partners
Objective: Establish and maintain communication channels	
Develop the project website	> Web platform with up-to-date information on the project; to be repurposed after the project's end as a repository of information
Develop contacts to be deployed in dissemination and exploitation	> A database of project contacts for announcements and news > A database of news feeds, newsletters and other channels where the project can promote its results (e.g. EGI newsletter)
Establish communication channels and keep them up to date	> Blog to publish articles about developments, announcements, news > 24 contributions per year to the blog
Objective: Facilitate the project work and inspire a sense of community	
Organise the project meetings	> 6 project meetings: 1 kick-off meeting and 1 all-hands meeting every six months
Keep the consortium informed and engaged	> 15 'Coordinator Briefs' detailing the latest developments and actions > 1 area on the website with generic information for the consortium (e.g.: quality plan, status of deliverables, etc) > 1 communications package of the consortium with project messages, description, slides, templates

Table 7 - Measures to communicate project results

ACTIVITY	Outputs
Demonstrate scale of resources added	> 1 infographic to show the degree of capacity expansion > 1 infographic to show the diversity of capabilities built onto EOSC > 1 infographic to promote the span and scale of EOSC Portal resources
Write case studies	> 2 case studies focused on successful use of cloud resources > 4 case studies focused on the integration work > 2 case studies focussed on the usage of the training platforms (user journeys)
Leverage EGI contacts	> 4 articles about project results published in the EGI Newsletter (readership: ~2000 per issue) > 6-10 emails to specific contacts requesting redistribution in their communities (e.g. the SAPS service is of interest to the ICOS infrastructure)
Promote the project's "How to" guides	> 1 edited and published version of the 'Handbook for integration' > 1 edited and published version of the 'How to Guide for trainers' > For each: 1 version of the publication in slide format
Promote the thematic services	> 1 service description optimised for the EOSC Portal (per service integrated) > 1 'how to access' guide for each service
Promote the data repositories	> 1 overview of the data contained in the repository (per repository added) to be published in the EOSC Portal
Promote training platforms	> 1 overview of usage during the project span and key benefits of the platform
Promote the project's policy work	> 6 presentations of the project's results in EOSC-related conferences (e.g. DI4R, EOSC Stakeholder forums, e-IRG meetings) > 1 mail distribution to all national infrastructures (leveraging consortium contacts) per policy work
Promote the quality badge scheme	> 1 overview of how the scheme works. > 1 publicly available table of badges awarded.
Promote the Software Quality baseline	> 1 digested overview of what makes a software component compliant
Promote the SQA-as-a-service	> 1 article on community newsletters about the service (e.g. EGI newsletter) > 1 page on the EOSC Portal describing the service with messages for users and for developers

3. IMPLEMENTATION

3.1 Work plan: Work packages and deliverables

The project is structured around six work packages addressing management, service, networking and joint research activities. Figure 7 shows the interrelations among the work packages and respective tasks.

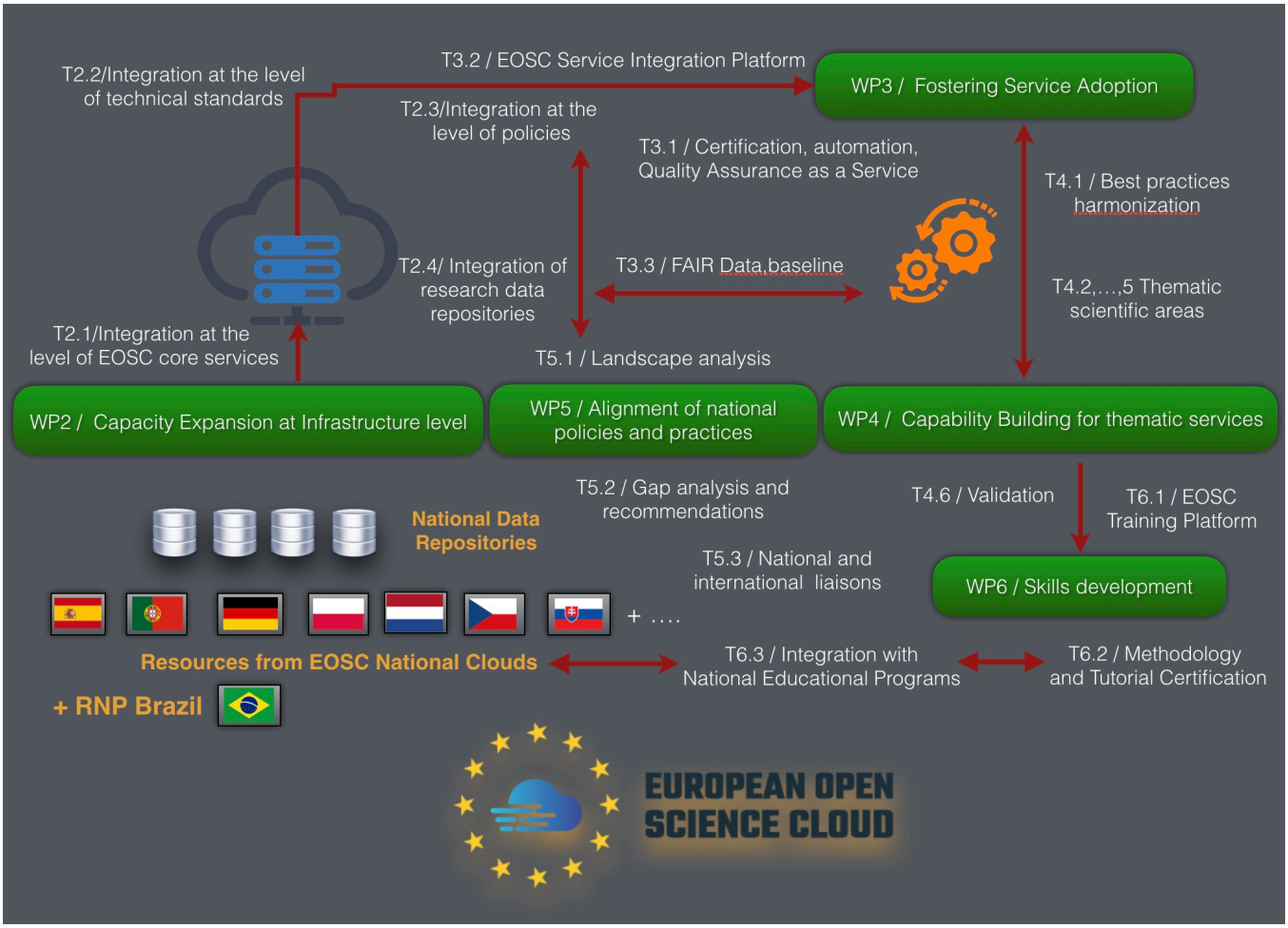


Figure 7. PERT chart: Project Implementation in terms of Work Packages and Tasks.

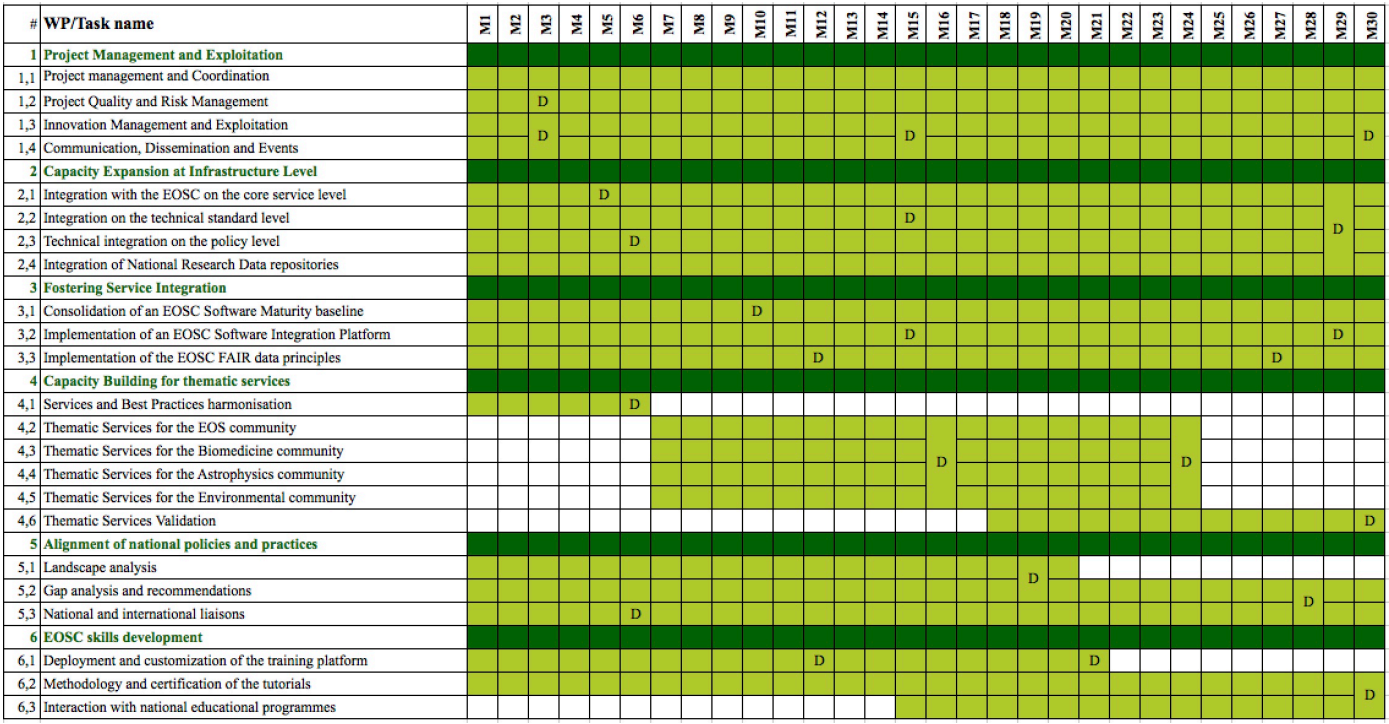


Figure 8. GANTT Chart

Table 3.1a: List of work packages

WP	WP Title / Type	Lead Part. #	Lead Part. Shortname	Person-Months	Start Month	End Month
1	Project Management and Exploitation /MGMT	1	CSIC	**53**	1	30
2	Capacity expansion at Infrastructure Level / SA	3	KIT	**216**	1	30
3	Fostering Service Integration in EOSC / JRA	2	LIP	**138**	1	30
4	Capacity Building for Thematic Services / SA	10	UPV	**239**	1	30
5	Alignment of National policies and practices / NA	5	CESNET	**86**	1	30
6	EOSC Skills development / NA	4	PSNC	**94**	1	30
TOTAL Person-Months				**826**		

3.1.1. Detailed description of work packages

WP1 - Project Management and Exploitation [M1 – M30]				
Lead beneficiary: CSIC				
Participant number	**1**	2	14	
Short name	**CSIC**	LIP	EGI.eu	
PMs per participant	**25**	8	20	

Short description

Work Package 1 is dedicated to the overall project administration including definition of the project Quality Assurance plan, Communication and Exploitation strategy. The objectives include the overall efficient operation of the consortium, careful monitoring of resource and financial expenditures, fulfilment of contractual obligations, periodic reporting and relationship with the European Commission.

CSIC is the Consortium Coordinator and Project Manager. CSIC will setup the Project Office that and will carry out all the financial and administrative activities; the Consortium coordination will be achieved through the Collaboration Board where each partner is represented; the Quality Assurance will be carried out also through the Project Management Board.

LIP as deputy coordinator will take care of the development the Quality Assurance plan, metrics and KPIs follow-up, to ensure that the integrated services are fulfilling the requirements of generic user communities.

The EGI Foundation will coordinate the dissemination and communications activities and will also contribute to the tasks related to exploitation of project results.

Objectives

- Provide an effective consortium coordination and project management support.
- Provide a comprehensive communication and dissemination strategy, covering the relevant segments of EOSC stakeholders.
- Define and implement an Innovation Management strategy geared towards effective exploitation of the project results.

Description of work. All tasks M1-M30

Task T1.1: Project management and coordination (Lead partner: CSIC [10 PMs]; Contributor: LIP [2 PMs])

The main activities will be:

- Delivery of the scientific and technical objectives of the project in a timely and effective manner. This includes the preparation of periodic activity reports and supporting the preparation of EC reviews.
- Financial Management of the project: monitoring of resources, preparation of periodic consolidated financial reports to be delivered to the EC.
- Provide the framework for a clear and effective communication between partners;
- Establishment of intra-project communication and information networks.
- This includes the provision of mailing list, tools to track the project evolution, support to project

meetings and follow-up.
- Manage relations with other EOSC relevant initiatives and projects at national and international level.

Task T1.2: Project Quality and Risk Management (Lead partner: LIP [6PMs]; Contributors: CSIC [5PM]; EGI.eu [3PMs])

The objective of this task is to establish and enforce effective management and quality procedures that will result in high quality project deliverables. The main activities will be:

- Definition of a Quality Plan (tools and metrics), including specific measures to follow up the fulfilment of KPIs.
- Encouraging and verifying that standards, procedures and metrics are defined, applied and evaluated.
- Assess the work and achievements of the different WPs
- Review deliverables and reports and verify fulfilment of Milestones
- Monitoring the risk and develop contingency plans, as well as ethics and gender equality

Task T1.3: Innovation Management and Exploitation (Lead: CSIC [10PMs]; Contributor: EGI.eu [5PMs])

The main activities will be:

- Capture project results, support the definition of the best IP protection approach, and define and maintain the exploitation plan.
- Definition an innovation management plan with guidelines on how to collect information and defined approaches to.
- Maintain a catalogue of project results to support the definition of the most appropriate exploitation routes. This will feed into the development of a dissemination and exploitation plan.

The task will rely on T1.4 for communications-related activities.

Task T1.4: Communication, Dissemination and Events (Lead partner: EGI.eu [12PM])

This task will define and implement the project's communications and dissemination activities, focusing on two areas:

Project-related communications. This will include the definition of the project's identity and key messages, and the development & operation of the project's website and communication channels.

Communication support for the dissemination & exploitation of results. The main activities of this area will be:

- Writing the content for the project's website, articles, publications and blogs.
- Develop the concept and content of the project's publications (e.g. booklets, leaflets, posters) as appropriate to the dissemination of results.
- Leveraging the EGI Foundation communications network pitching articles about the project's results to the communication channels of EOSC stakeholders.
- Supporting the project coordinator and the technical teams in their dissemination work, for example support for presentations, editing of papers or abstracts, managing submissions to conferences.

This task will also organize the project's all hands meetings (one kick-off, and one meeting every six months) and manage the contribution of the project to external conferences (e.g. DI4R, community-specific conferences). The team will also support the WPs in organising, when required, focused workshops.

Deliverables

D1.1 – Exploitation and dissemination plan (R, CO) – M3 – Lead: CSIC

D1.2 – Exploitation and dissemination report (R, CO) – M15 – Lead: EGI.eu

D1.3 – Final exploitation and dissemination report (R, CO) – M30 – Lead: CSIC

These reports will focus on the project exploitation strategies and associated communication activities

– Quality Plan (R, PU) – M3 – Lead: LIP *Includes innovation management and risk management*
Milestones (mov = means of verification)
M1.1 – Project website implemented – M3; mov: website online
M1.2 – Communication package available to the consortium – M3; mov: added to the website

WP2 - Capacity Expansion at Infrastructure Level [M1 – M30]								
Lead beneficiary: KIT								
Participant number	1	2	**3**	4	5	6	12	15
Short name	CSIC (ES-JRU)	LIP	**KIT**	PSNC	CESNET	IISAS	CYFRONET	INCD
PMs per participant	54	13	**37**	25	10	20	30	27

Short description

National infrastructures and thematic resources are often built for a specific community, focused on specific use cases. This is why they work well for the communities that built them, but not within the larger context of EOSC. The integration with EOSC requires adaptations to the basic EOSC services, such as AAI, Information System, Monitoring and Accounting.

This work package will address this issue focussing on:

- Integration with the EOSC on the core service level: AAI, EOSC Portal, security, monitoring and accounting.
- Integration on the technical standard level: for example, APIs, tools and software of the repository, versioning of the datasets.
- Technical support for implementing policies from a higher level.
- Integration of research data: accessibility of research data in national repositories, prioritising measures to comply with the FAIR principles.

Objectives

- Integrate national cloud infrastructures, resources and national data repositories within the EOSC platform, so they can be accessed in a streamlined, generic way.
- Expand EOSC capacity to support provisioning of thematic services through EOSC.
- Provide a handbook to disseminate expertise of federated infrastructure management and operation to national and thematic infrastructures and resources.
- Involve the NRENs to support adherence with policies and practices relevant in the EOSC context of international collaboration. This will especially focus on policies, such as Operational Security, Incident Response, Acceptable Authentication Assurance and Acceptable Use Policies.
- Support open source protocols and envisage long term availability of the integration effort.
- Support Thematic services and research data repositories increase their adherence to the FAIR

principles. Priority will be given to those communities committed to making their data "FAIR".

Description of work. All tasks M1-M30

Task T2.1: Integration with the EOSC on the core service level (Lead partner: CYFRONET [20PM]; Contributors: CSIC[18PM];KIT [15PM]; PSNC [15PM]; LIP [5PM])

This task will analyse the status of the national and thematic capacities to be integrated to identify gaps that need to be overcome for a full interoperation with existing services in the EOSC context. This will result in an individual integration plan per national service to be captured in D2.1. We expect that work for one capacity will be beneficial for others. Therefore, it is expected that integration of most services will require cooperation and collaboration with their operators.

Following the integration plan, this task will conduct targeted developments for tools to foster the integration of services with EOSC, using different approaches for each class of service. Based on our experience with building up federated services in international research projects, we envisage the following approaches for the different classes of services:

- Compute at the IaaS level: the majority of existing capacity is based on two major products. To fill the gap to EOSC, they require integration with core EOSC services: AAI, Information System, Accounting and Monitoring.
- Openstack, Kubernetes, Mesos, are all mature components for providing cloud resources. Integration with EOSC requires integration with the information system. These services require additional integration effort, because they are not multi-tenant capable. Proper support to match EOSC requires integration at conceptional levels.
- Storage is provided using many different protocols on the frontend. Also storage sites are very complex (SSD, disk, tape) on the backend. Integration of existing installations requires:
    - Identification of the most relevant protocols (e.g. WebDAV)
    - Integration or extension of one selected protocol to support EOSC AAI
    - Integration of existing national or thematic storage services with the Accounting and Monitoring of EOSC.
- Advanced capabilities such as resource orchestration, provided for example via the INDIGO DataCloud service Infrastructure Manager (IM, now part of the EOSC-hub portfolio), require integration with the information system so it can schedule resources dynamically. In addition scenarios of IM-on-Demand and IM-as-a-Service can be enabled by EOSC integration.

Task T2.2: Integration on the technical standard level (Lead partner: CSIC [16PM]; Contributors: LIP [5PM]; CYFRONET [5PM]; INCD [10PM])

This task focuses on homogeneous access to services via tools on the client side. This drives the integrated accessibility of the integrated services. In other words, even with an integrated AAI, there are still many different protocols that hinder, for example, transfer of data from one storage to another. We will take different approaches according to service type. In many cases it will not be possible to change the interfaces offered at the services, therefore client-side approaches will be followed. For storage, we will pursue the so called "rclone approach". Rclone is a client tool that supports a large number of transfer protocols. This tool will be integrated, extended and configured to work with all of the supported services.

For computing there are a lower number of services to be supported, hence we mandate to support a common set of APIs, use compatible tools to provide common basic services (e.g. OpenStack, Kubernetes, or similar.

On top of these basic services, more complex services can be integrated. This task will ensure that they are provided and accessible in an integrated way. This may include simpler services, such as SQL (e.g. MySQL, Postgres) but also more complex services such as state-of-the-art repository services (e.g. linked data, scalability,

versioning of the datasets). The focus for all of them is the integrated approach, making use of (or extending) EOSC central services, such as AAI, Monitoring and Service Catalogues.

This will lead to interoperability between services, so that European and Global users can leverage the capacity that is available.

For detailed planning (for storage, computing and containers) we will identify the minimum subset of well-proven native interfaces that are required to enable infrastructure access and thematic service integration in EOSC in D2.1, and promote their adoption. Once the basic integration is done, we will integrate the more complex services, likely in a *aaS fashion.

Task T2.3: Technical integration on the policy level (Lead partner: KIT [15PM]; Contributors: CSIC [12PM]; PSNC [5PM]; IISAS [5PM]; INCD [10PM]; CESNET [5PM])

This task addresses the technical measures that are required in the EOSC context to fulfil the policies governing collaboration across infrastructures, such as policies for data access, policies for resource allocation but also those for enacting policies for authentication, authorisation, and security incident response.

Work here is clearly separated from WP5. While WP5 deals with the harmonisation of the policies themselves, this task focuses on the implementation of technical measures to fulfil the policies.

Technical integration work is required to make sure that the relevant policies can be implemented in similar ways at the different participating centres. The national research network providers (NRENs) will be involved, to exploit their expertise in technical implementation of policies in international and inter-organisational environments.

Work in this task will therefore start with the gap analysis and document the results in D2.2. This will include a comparison between existing policies and recommendations from EOSC and specifically the AARC project. Based on that, recommendations will be made to each service that contains recommendations for harmonising policies in the federated EOSC context. In particular:

- How to harmonize resource allocation in the EOSC context
- How to organize data access scalable in a federated environment
- Which measures to take to enable security among collaborating infrastructures
- How to authorize and identify users in a consistent and scalable way

These results will be presented in D2.3.

We will prioritise implementation of the recommendations by the number of services that benefit from the implementation. The report about the integration with the services will be given in D2.4.

Task T2.4: Integration of National Research Data repositories (Lead partner: IISAS [15PM]; Contributors: KIT [7PM]; CSIC [8PM]; PSNC [5PM]; LIP [3PM]; INCD [7PM]; CYFRONET [5PM]; CESNET [5PM])

This task will perform the technical integration of national research data repositories of wider interest into the EOSC ecosystem, with the goal of making the data available to the European researchers and exploitable by the EOSC services. This work will leverage established and emerging data repositories that will be identified in close collaboration with the project funded under INFRAEOSC-5 (c).

These are community specific repositories or national repositories that aspire to comply with FAIR data practices and become openly available. The activities include:

- Work with WP5 to identify candidates and prioritise the integration;
- Push the alignment of those candidates towards EOSC FAIR data principles, relying on the interaction of WP5 with the INFRAEOSC-5 (c), so that they achieve the necessary level of maturity for integration in

EOSC, addressing and verifying practical aspects of data representation, metadata, data provenance, licenses, access policies and legal.
- Work with WP3 to ensure the required quality baselines.
- Support the technical integration of those repositories in EOSC, meaning that data will need to be findable in a standard way by the users and other services, accessible via standard interfaces supported in EOSC, combinable with other data in EOSC repositories, and exploitable by the EOSC services enabling data analysis to be performed using resources and services from the EOSC infrastructure providers.
- Work with T2.1 and T2.2 to ensure the necessary adaptations in terms of AAI, APIs, communication protocols, data identifiers and interoperability with EOSC services and may include data hosting aspects when needed.

This task will apply the EOSC FAIR data principles in collaboration with WP3 that will provide the technical framework to support the best practices at the level of implementation, validation and monitoring. This will ensure that the data repositories will meet the desired quality baseline and remain compliant.

The framework of principles, tools and services will be put into practice to demonstrate how they can facilitate the process of adding new data repositories to EOSC while simultaneously helping repositories to become "FAIR" and enlarge the EOSC capacity with new data.

Deliverables

D2.1 – Roadmap for integration of national capacities into the EOSC (R, PU) – M5 – Lead: CYFRONET

This report identifies individual integration plans per national service. It will provide a roadmap on which steps to take for the integration of all targeted capacities within EOSC.

D2.2 – Policy Gap Analysis (R, PU) – M6 – Lead: KIT

This report identifies the existing gaps between policies of national infrastructures and those recommended by AARC for the EOSC context.

D2.3 – Intermediate report on integration efforts (R+DEM, PU) – M15 – Lead: CSIC

This report will provide information on the implementation status of Integration Plan. The Integration Plan will be updated where appropriate.

D2.4 - Final report on EOSC integration (R+DEM+DEC, PU) – M29 – Lead: IISAS

This report will be a valuable "Handbook" on how to integrate national clouds, thematic resources, and data repositories conformant to common quality standards, and harmonized in terms technological, policy, and legal aspects. (R+DEM+DEC, PU)

Milestones (mov = means of verification)

M2.1 – Outlined Roadmap for Integration available for discussion – M3; mov: draft plan for roadmap

M2.2 – Joint team with WP5 formed to address the Policy Gap Analysis – M4; mov: team composition

M2.3 – First service analysed – M6; mov: service analysis report

M2.4 – First service integrated into EOSC, two more analysed – M12; mov: first service available in EOSC

M2.5 – Internal status update on progress of Integration Plan, Service Integration, Policies and Data integration – M12; mov: internal report

M2.6 – Additional services analysed – M18; mov: services analysis report

M2.7 – Internal status update on progress of Integration Plan, Service Integration, Policies and Data integration – M24; mov: internal report

WP3 - Fostering Service Integration [M1 – M30]

Lead beneficiary: LIP

Participant number	1	**2**	3	7	10	11	12
Short name	CSIC	**LIP**	KIT	DANS	UPV	LNEC	CYFRONET
PMs per participant	34	**34**	5	25	10	15	15

Short description

The reliability of EOSC services and data is a key aspect in motivating adoption from European research communities. This work package puts in place the common processes and tools to define and automatically validate the quality and maturity requirements of the software that composes services, both thematic (i.e.: exposed to researchers) and core (i.e.: needed for advanced capacity integration).

Compliance with FAIR data principles will be tackled automatically, to the extent possible, through metadata analysis and leveraging machine-actionable features on data repositories. The ultimate goal would be to add incentives for EOSC service adoption through the seamless disposal of mature and validated services through the marketplace, setting the foundations for the establishment of an EOSC-ready stamp both for data and software.

Objectives

- Manage the integration process of thematic (WP4) and infrastructure (WP2) services in EOSC, focusing on validation, delivery and deployment of software components.
- Consolidate a Software Sustainability baseline for service validation, harmonized with de-facto standards and specifications of software quality and security assessment.
- Rely on automation to speed up the validation and certification of the Software Sustainability baseline, through the execution of continuous integration (CI) and continuous delivery (CD) pipelines.
- Promote incentives for the adoption of quality practices by the definition and implementation of a badge issuing process to recognize the achievements in relation to the baseline recommendations, towards the establishment of an EOSC-ready stamp.
- Develop a Software Quality Assurance (SQA) as-a-service to facilitate the development, delivery and integration of existing and prospective EOSC services.
- Foster the adoption and compliance of FAIR data practices by the scientific communities through the implementation of a mechanism to support the recommendations of the FAIR Coordinating and Support Action approved under the subtopic (c) of this call.

Description of work. All tasks M1-M30

Task T3.1: Consolidation of an EOSC Software Maturity baseline (Lead partner: LIP [10PM], Contributors: CSIC [6PM], UPV [5PM], DANS [10PM]

This task will tackle the definition, alignment, maintenance and evolution of a common baseline for the implementation of Software Maturity in the EOSC ecosystem. The outcome will be a well-defined process towards labelling software and the corresponding derived services as “EOSC-ready”, meaning that they are suitable to be delivered under EOSC. The baseline will exploit current know-how (described below) and will provide clear guidelines, targeted to service providers, to facilitate such delivery.

The Software Maturity baseline shall contribute to:

- *Quality, reliability and maturity of scientific software*, by the consolidation of best practices that seek software quality. Scientists are not just users of software. They are prime developers as well and consequently there is the need to promote quality from the early stages of the software life cycle. This

is done through the adoption of Software Quality Assurance (SQA) practices. Following this concept, we will ensure that service software components are open and secure (with static and dynamic security testing), with readable source code (with style standards) and compliant with the expected functionalities, both in terms of operational (with test-driven development) and behavioural requirements (with behaviour-driven testing). Therefore, new versions of services could be continuously delivered and integrated in EOSC.

- *Reusability, discoverability and transparency of software*, allowing easy access to specific versions of its source code, thus improving the reproducibility of the experiments by relying on a common metadata schema.

The Software Maturity baseline will:

- Follow the guidelines of the *"A set of Common Software Quality Assurance Baseline Criteria for Research Projects"* document. This task will contribute to the maintenance and continuous enhancement of these guidelines, such as the adoption of FAIR principles applied to software or the reinforcement of quality requirements and security practices.
- Be harmonized and aligned with: EOSC software recommendations, software Technology Readiness Levels (TRLs), know-how gathered in EOSC projects (e.g. EOSC-hub), to ease the integration of services in the EOSC ecosystem.

The application of the Software Maturity baseline for the validation of a given service shall outline its level of conformity with respect to the defined requirements, recommendations and best practices. The software components processed through the SQA pipeline will be awarded specific quality badges to recognise their achievements. The quality badges will be embedded into the service component metadata description.

The evaluation of software will distinguish between *functional maturity* and *quality* of software. This distinction will be useful, for example, in the (expected) cases where the service is solid and works as required for a considerable amount of time (i.e. is mature), but it is not following best practices during its life cycle (i.e. does not meet quality standards).

The purpose of the quality badges is to promote transparency and traceability, as they will allow tracking down the quality achievements for future access and review. This assessment is also an incentive to the adoption of EOSC and FAIR practices as it appears as a symbol of trust and quality of open software.

By demanding minimum requirements and recognizing achievements on SQA and maturity of services, the Software Maturity baseline will contribute to the definition of a software certification process that is capable of issuing an "EOSC-ready stamp" for the integration of prospective services in EOSC.

Task 3.2: Implementation of an EOSC Service Integration platform (Lead partners: CSIC [20PM]; Contributors: LIP [17PM], UPV [5PM], LNEC [15PM], CYFRONET [10PM]

The EOSC Service Integration platform will put into practice the outcomes from Task 3.1 in order to validate the quality and maturity of the services following the EOSC Software Maturity baseline.

The following considerations will be accounted in the implementation of the platform:

- The platform will rely, to the extent possible, on automation in order to be cost and time-efficient. Leveraging CI services, i.e. Jenkins, the validation of the minimum requirements and best practices identified in Task 1 will be executed through CI/CD pipelines as shown graphically in the Methodology section (Figure 4). These pipelines describe, using Apache's Groovy programming language syntax (Pipeline as Code), the sequential and parallel steps needed to perform these validations (e.g. ensuring the openness of the source code by checking its adherence to any of the licenses promoted by the Open Source Initiative).

- Service providers will be able to extend the steps/features of the aforementioned pipelines. These pipelines are meant to be part of the source code's repositories of the services, thus they can be versioned, profiting from all the features of a version control system.
- The platform will provide a shared library for easing the definition of the CI/CD pipelines. The library will be built on an existing implementation, currently used and developed within EINFRA-21 initiatives (e.g. DEEP-Hybrid-DataCloud, Extreme-DataCloud) to test the requirements defined in the SQA baseline delivered by Task 3.1. Extensions for this library will be gradually available, thus fostering reusability and keeping pipeline code DRY.
- The platform will interact, through open APIs, with common, de-facto online software repositories. It is expected to react to events coming from those repositories so that the validation process is started right after a given change in done to source codes.
- The automatic deployment and on-demand provisioning of services will be evaluated, aiming to be prototyped and eventually integrated with the EOSC marketplace.

Towards the definition of an EOSC-ready seal:

- The result of the validation, through the execution of the formerly described pipelines, will report a grade of conformity according to the minimum requirements and best practices described in the Software Maturity baseline from Task 3.1.
- In case of successful validation, and according to the results, online badges will be issued so that new software versions get EOSC-acknowledged and publicized in online repositories, embedded as metadata.
- The implementation will be based on the Open Badges standard.

SQA as a Service:

- The platform will be openly available as-a-service so that the validation of Software Maturity baseline can be triggered on demand for any given software release.
- This service will allow e.g. external/onboarding service providers to check the EOSC conformance at any stage while integrating their services in the EOSC Portal.

Task 3.3: Implementation of the EOSC FAIR data principles (Lead partner: DANS [15PMs], Contributors: CSIC [8PM], LIP [7PM], KIT [5PM], CYFRONET [5PM]

Scientific communities agree that research data needs to follow FAIR principles. To go from theory to practice we need to provide guidance to make data comply with FAIR. This task will build upon the EOSC recommendations on FAIR data practices delivered by the forthcoming CSA project funded under subtopic (c) in this call.

The implementation of the EOSC FAIR data principles implies the timely analysis of the outcomes on FAIR practices coming from the EOSC-relevant context in the CSA project, and the design of a technical framework to support those best practices operationally at the level of: implementation, validation and monitorization. The framework will leverage from the architecture and technical solutions from Task 2, as shown graphically in the Methodology section (Figure 6).

- *Implementation*: The repositories will be require to provide a set of machine-actionable features to automatize the validation process using standard protocols. These functions will facilitate different tools to access both data and metadata to determine the FAIRness level. The data FAIRness will be supported by three basic elements: the system where the data will be stored (repositories), the metadata describing the data, and the datasets or digital objects.
- *Validation*: Once a new dataset or digital object is published in a repository, the data and metadata will be analysed to extract information to determine FAIRness level. The validation process will take into account the four data principles separately, and as an example, it could check the following issues:

- Findable: Check Persistence Identifier existence, resolvable, metadata contained.
- Accessible: Check repository protocol to access data/metadata.
- Interoperable: Check minimum metadata (provided by call subtopic c) existence and correctness, at a different level. Check data format.
- Reusable: Check data provenance and type of license.

- *Monitorization*: Whenever a dataset or digital object is modified, it will be re-checked in order to certify its current FAIRness level. Other monitoring issues like data reusability indexes will be taken into account.

Deliverables

D3.1 – Software Maturity baseline (R, PU) – M10 – Lead: LIP

Describes quality requirements and best practices to be considered when validating software from an EOSC service; describes the badge issuing process.

D3.2 – First prototype of Service Integration platform (R, PU) – M15 – Lead: CSIC

Architecture, first achievements and implementation status of the platform for software validation of EOSC services, with the first CI/CD pipeline definition for the WorSiCa Thematic Service from WP4.

D3.3 – Intermediate report on technical framework for EOSC FAIR data principles implementation (R, PU) – M12 – Lead: DANS

Describes the evaluation of the recommendations for assessing data FAIRness and data repository features coming from INFRAEOSC-5c and provides details about architecture, requirements and a roadmap for implementation.

D3.4 – Final release of Service Integration platform (R, PU) – M29 – Lead: CSIC

Describes the CI/CD pipelines for service validation, the badge issuing process for achievement recognition and the SQA-as-a-service offering.

D3.5 – Final report on technical framework for EOSC FAIR data principles implementation (R, PU) – M27 – Lead: DANS

Implementation details for a technical framework to validate and monitor data FAIRness. Any change or addition to the information gathered in D3.3 will be reported.

Milestones (mov = means of verification)

M3.1 – All infrastructure services controlled by the project (national initiatives) have CI/CD pipelines – M6; mov: Jenkins service

M3.2 – Quality and maturity software requirements & best practices are defined – M8; mov: new release of "A set of Common Software Quality Assurance Baseline Criteria for Research Projects" document.

M3.3 – Working CI/CD pipeline for WorSiCa thematic service (WP4) – M12; mov: Jenkins service

M3.4 – Badge issuing implemented – M20; mov: Jenkins service, EOSC Portal

M3.5 – All thematic services (WP4) have a working CI/CD pipeline – M22; mov: Jenkins service

M3.6 – SQA-as-a-Service allows to plug-in generic software for CI/CD. Integrated with badge issuing – M25; mov: SQA-as-a-Service API

WP4 – Capacity building for thematic services [M1 – M30]

Lead beneficiary: UPV

Participant number	1	2	3	5	6	9	**10**	11	16
Short name	CSIC (ES-JRU)	LIP	KIT	CESNET	IISAS	IRD	**UPV**	LNEC	INDRA
PMs per participant	68	10	15	15	10	20	**30**	47	24

Short description

WP4 leads the technical and operational integration of scientific thematic services so that new services can become part of the EOSC Catalogue and/or Portal.

By thematic services we mean production-level mature services at the national level (TRL7+), relevant for being exploited European-wide, that would benefit from the possibility of scaling up resources and/or sharing data with their international peers in the EOSC ecosystem.

WP4 will also apply service harmonisation and standardisation through common best practices and common interfaces to ease the federation in a wider context, both at the transnational level and across scientific communities.

Objectives

- To harmonize EOSC-relevant thematic initiatives and use cases in relevant scientific areas (Earth Observation, Biomedicine, Astrophysics and Environment) providing open research data and services, to expand the use of the mature national services in an international scope.
- To define best practices for the adoption of common EOSC core tools and services identified in WP3, considering community-wide standards, to increase the capacity of dealing with more complex user demands.
- To increase the capacity, performance, reliability and/or functionality of mature thematic services currently not integrated into the EOSC portal or which have limited integration by means of EOSC core tools and services.
- To increase the number of users of these thematic services and users' satisfaction in an international scale through its integration in the EOSC framework and targeted training providing content for the WP6 skills development activities.
- To increase service quality by supporting the uptake of certification and FAIR data practices, defining and presenting incentives to do so in coordination with other activities in the project.

Description of work.

The structure of the work package follows a combination of horizontal and vertical tasks, so they provide the right degree of coordination. On the one hand, the activities of best practices harmonisation and validation of the thematic services are defined at a global level, fostering the interaction across disciplines and providing a coherent evaluation schema. On the other hand, we define one vertical activity per domain, so integration activities can take place within the domain, fostering the integration at the level of the thematic service.

This approach guarantees a cross-domain fertilization while preserving a reasonable degree of freedom within the communities, so the thematic services can benefit from the commonalities and the existing organization at community level. For the sake of brevity, we include a detailed description of the thematic services and the expected impact in the appropriate sections of the proposal, as well as a clear differentiation between the activities being carried in concurrent projects and the aim in EOSC-synergy.

Task 4.1. Services and Best Practices harmonisation (Lead Partner: UPV (10 PMs); Contributors (PMs): CSIC-IFCA (2PM), CSIC-CNB (2PM), CSIC-BSC (2PM), CSIC-CIEMAT (2PM), LIP (1PM), KIT (2PM), CESNET (2PM), IISAS (1PM), IRD (2PM), LNEC (4PM), INDRA (2PM) – **Task duration: M1-M6**

It targets mainly the first two objectives, by eliciting current use cases, limitations, evaluation metrics and baseline for the validation of the services. As a global task 4.1 will produce consolidated input to WP3 and will gather the necessary information to implement Task4.2-5 activities. The task will identify technical and scientific representatives from each service, obtain user profiles, bottlenecks, service needs and will select a set of representative end-users to obtain the user satisfaction baseline and use them as prescripts.

In the following Tasks (4.2-4.5) we provide a coherent approach at task level for each community scoped in the project. Based on the use cases agreed and the outcome of the Service Adoption Work Package, the activity in each Thematic Service task will go through four sub-tasks:

- Integration of EOSC Services,
- Testing of the services by selected groups,
- Development of community-specific training material and
- Liaison with the user community.

All thematic tasks address objectives that are verified and quantified in Task 4.6.

Task 4.2. Thematic Services for the EOS community. (Lead partner: LNEC (37 PMs), Contributors (PMs): LIP (7PM), IRD (16PM), UPV (5PM), INDRA (18PM) **– Task duration: M7-M24**

In the Earth Observation community (EOS), the project will deal with three thematic services:

- **WorSiCa**, for the detection of the coastlines changes, coastal inundation areas and inland waterbodies water detection, led by LNEC.
- **G-Core**, a production-ready technology used as a service at ESA's and national programs led by INDRA for managing EOSC data in terms of the acquisition, storage, cataloguing and processing data of several levels from different missions.
- **SAPS**, a service to compute the Surface Energy Balance Algorithm for Land (SEBAL) and similar information for estimating the evolution of forest masses and crops. This last use case is based on an international collaboration with Brazil. It requires integrating access to the LANDSAT data repository (Table 10).

In the frame of this task we have two aims:

- To scale from national up to European level the WorSiCA and SAPS services (integration of WorSiCA as an EOSC service and deployment of a federated site of SAPS in EOSC to facilitate European scientists to exploit the evapotranspiration estimation services from remote sensing imagery).
- To explore the sustainability of the EOS services exposed through the creation of added-value products through the integration of G-Core as a data manager.

Task 4.3. Thematic Services for the Biomedicine community. Lead Partner: CSIC-CNB (12PMs); Contributors (PMs): CSIC-BSC (10PM), CESNET (5PM), UPV (3PM) **– Task duration: M7-M24**

This task will deal with two main thematic services focusing on bioinformatics benchmarking and integration of metadata of biomedical information, and the processing of Cryo-Electron Microscopy:

- **OpenEBench** is a platform for supporting scientific communities-led benchmarking efforts across

different domains in Life Sciences by exposing a virtual research environment where communities can deploy workflows for computing scientific metrics, and by facilitating an unbiased technical and scientific assessment of participants by executing participants workflows under the same environment. This case aligns with the role of the BSC as lead entity to facilitate the access to Core Data Resources including EGA (European Genome-Phenome Archive).

- The **Cryo-Electron Microscopy** service serves structural biology community with Scipion (a workflow-oriented image processing software integrator developed by the CNB in Spain) running on resources provided by the Masaryk University in the Czech Republic.

In the frame of this task, we want to integrate those services in the EOSC Portal to expand their capacity. In particular:

- The OpenEBench service for genomics will expose and deploy the existing genomic processing pipelines on top of heterogeneous resources available, to provide access and processing to both data and metadata across different distributed resources to validate, at data and metadata level. This case will also provide some best practices for the harmonisation of data access to clinical trial data of EGA and clinical data warehouses already connected through Informatics for Integrating Biology and the Bedside (i2b2) - https://github.com/i2b2.
- Scipion is able to provide cloud deployments on demand which could deal with the scientist's own data with an automatic setup of the tools and environments. However, currently this has not been widely used due to lack of cloud resources. By integrating it in the EOSC Portal users will be able to access a service that automatically produces a cloud instance of Scipion with all the user data loaded and with all the required software.

Task 4.4. Thematic Services for the Astrophysics community. Lead Partner: CSIC-CIEMAT (18 PMs); Contributor: UPV (3 PM) **– Task duration: M7-M24**

This task will deal with the integration of the Latin American Giant Observatory (LAGO) data within EOSC. LAGO will be represented through CIEMAT, which have intensively worked with this initiative in the frame of previous projects such as CHAIN-REDS.

LAGO is an extended cosmic ray observatory composed of a network of water-Cherenkov detectors (WCD) spanning over different sites located at significantly different altitudes (from sea level up to more than 5,000m) and latitudes across Latin America. The LAGO network is simple and robust, and incorporates several integrated devices to allow time synchronization, autonomous operation, on board data analysis, as well as remote control and automated data transfer.

The LAGO experiment perfectly matches the requirements that need to be exploited on an infrastructure as EOSC as it complies with the EOSC data management standards and simulations can be seamlessly run by mature and well-tested applications on distributed virtualized platforms (e.g. CORSIKA; GEANT4). In this way, the codes used to perform the simulations will be integrated into the EOSC portal and the LAGO measured data can be harvested from the EOSC services.

Task 4.5. Thematic Services for the Environmental community. Lead partner: CSIC-BSC (8PMs); Contributors (PMs): KIT (10PM), CESNET (6PM), IISAS (7PM) **– Task duration: M7-M24**

This task will deal with the integration of four thematic services related to the protection of the environment:

- **SDS-WAS** is a World Meteorological Organisation service for sand and dust storm warning advice and assessment.
- **UMSA** is an untargeted mass-spectrometry analysis service from RECETOX (Research Centre for Toxic Compounds in the Environment at Masaryk University) in the Czech Republic.
- **MSWSS** is a service for analysis of water network distribution with regards to the mitigation of

hazardous events, by the integration of existing on-line analysis of toxics in drinking water supply networks with distribution network simulation of EPANET.

- **O3AS** is a service set up to analyse ozone projections and forecast their evolution on a long-term basis.

In the frame of this task, the previous services will be made available through the EOSC Portal, including added-value processing services. In particular:

- The SDS-WAS applies and offers several modelling techniques for dust forecast through national services (e.g. https://sds-was.aemet.es/forecast-products/dust-forecasts). By the integration of such services in the EOSC, a more complete set of derived services can be built and offered to a wider group of users. The geographical area of interest of this service is reaching less favoured countries and having the potential to increase the life quality.
- MSWSS services require a more powerful and efficient infrastructure for modelling complex drinking water supply system for being offered in a global extend. By the integration of MSWSS in EOSC it will be possible to offer the service to a wider community.
- UMSA will be integrated in EOSC thanks to use the EOSC Federated Cloud Backend, targeting a European-wide scientific community. The drinking water toxicity is currently offered at a reduced scale and not available online. By integrating it in EOSC following the INSPIRE directive we open up a wide range of new services.
- Finally, the integration of O3AS, a complex workflow comprising novel analysis tools that access Climate Data Archives using EOSC methods. For the Scientific Assessment of Ozone Depletion large data volumes have to be analysed to generate key metrics for policy makers (e.g. ozone return dates). This analysis is always time critical and speeding it up will ease the assessment process for future cycles (e.g. 2022, the 2018 Assessment will be published in December).

Task 4.6. Thematic Services Validation. Lead partner: UPV (9PMs); Contributors (PMs): CSIC-CNB (4), CSIC-BSC (4), CSIC-CIEMAT (4), LIP (2), KIT (3), CESNET (2), IISAS (2), IRD (2), LNEC (6), INDRA (4) – **Task duration: M18-M30**

The validation of the thematic services is a key activity and aligned with the objectives of the call. By the integration on the EOSC, the thematic services are expected to extend their functionality, increase their capacity or improve their performance and availability, reaching a wider and international user community. These improvements will be subject to a rigorous assessment based on a set of quantitative indicators and Key Performance Indicators.

The validation will be done in two main aspects: Software and service quality and user outreach and satisfaction. The software quality will be closely aligned with the WP3 Software Quality Assessment (SQA), which defines the quality assessment for the EOSC services (infrastructure and thematic services). The EOSC integration of the thematic services will be assessed through the WP3 SQA and WP3 FAIR recommendations and guidelines. The validation at the level of the users will evaluate outreach (new users, new countries involved, new datasets integrated) and the user satisfaction by means of standardized questionnaires.

We do not expect that all the thematic services to improve in all the dimensions, but this task will provide a coherent evaluation of such metrics. The baseline will be defined in T4.1, and the task will provide with the evidences for such improvement.

Deliverables

D4.1 – Best Practices Elicitation including Data Management Plans (R, PU) – M6 – Lead: UPV

This report will describe the methodology, the use cases elicited and the harmonisation of them so they can be used for both guiding the WP3 activities and for the Validation activity.
D4.2 - First prototype of the EOSC Thematic services (DEM, PU) – M16 – Lead: UPV *This deliverable will consist on a first prototype of the EOSC Thematic services and an associated report describing the architecture and roadmap.*
D4.3 - Final release of the EOSC Thematic services (OTHER , PU) – M24 – Lead: UPV *This deliverable will consist on the final release of the EOSC Thematic services and an associated report describing the service, access, training material and test validation.*
D4.4 - EOSC Thematic services validation report (R , PU) – M30 – Lead: UPV *This deliverable will consist on a report that describes the outcome of the validation of the thematic services.*
Milestones
M4.1. Inventory of Use Cases and EOSC Services. It will be verifiable through the release of D4.1 (M6). M4.2. Pioneer thematic services integrated in EOSC. A reduced set (at least three) of significant thematic services are integrated into EOSC. Although they will be in production mode, we do not expect to deliver the full functionality at that time. When reaching this milestone, the experience of these pioneer services will be used as basis for refining the plans for the rest of the services. A piece of evidence will be obtained to verify the achieving of the milestone. (M12) M4.3. All prototype services integrated. All the thematic services have been integrated in the EOSC and offer (but limited) functionality in production mode. Verifiable by the release of D4.2. (M16) M4.4. All thematic services in full production. Verifiable by the release of D4.3. (M24) M4.5. Feedback evaluated. All the services have been running for a period of time of not less than 6 months and feedback from users have been collected. Verifiable through the release of D4.4. (M30).

WP5 - Alignment of national policies and practices [M1 – M30]

Lead beneficiary: CESNET

Participant number	1	4	**5**	6	8	7	11	13	14
Short name	CSIC	PSNC	**CESNET**	IISAS	JISC	DANS	LNEC	FCT	EGI.eu
PMs per participant	10	5	**22**	5	4	5	5	15	15

Short description

This work package is dedicated to harmonization and development of policies and best practices for resource provisioning and allocation compatible with EOSC in the participating countries. WP5 is in charge of gathering national information and requirements, complement it with information from international stakeholders, conduct a gap analysis and develop recommendations that are validated across different countries, and promoted in the final stage of the project for adoption by national funding agencies and policy makers.

Strong emphasis will be given to the harmonization of transnational access policies, as the lack of shared European vision on provisioning resources between countries is perceived as one of the main obstacles to full EOSC implementation (where e.g. services may reside in countries with limited resources to cover EU-wide needs). In parallel, the work package will analyse the potential for harmonization of national policies that could allow better coordination between involved countries, could lead to joint procurement and service provisioning

organizational models, and thus provide resources in more efficient and coherent ways.

The alignment of policies would borrow key elements of the **European Interoperability Framework** (EIF) as a foundation for the work, by considering the overlap of national policy frameworks, particularly those applicable to the Legal, Organisational and Semantic interoperability layers.

It would also ensure policy alignment according to the twelve key principles of the EIF: (1) Subsidiarity and proportionality, (2) Reusability Inclusion and accessibility, (3) Administrative simplification, (4) Openness (5) Technological neutrality and data portability, (6) Security and privacy, (7) Preservation of information (8) Transparency (9) User-centricity, (10) Multilingualism, (11) Assessment of Effectiveness and (12) Efficiency

Objectives

- Ensuring that EOSC policies and practices are propagated to the national level, taking into account the requirements of national funding agencies, the needs of the national roadmaps and the constraints of national organizational and regulatory frameworks. Information on the national needs and constraints will be gathered and provided as input to the EOSC governance.
- Identify successful best practices and gaps so that recommendations are provided to help evolve the national policies and practices according to the needs of open science practices and international research collaborations.
- Increasing national and international organizational awareness and possible coordination in funding, procurement and provisioning of services in the EOSC scope.
- Providing input and advice to the EOSC governance on Rules of Participation.
- Facilitating - through compatible policies and operational conditions - the access, sharing and exploitation of nationally funded services and resources across international user communities in the context of EOSC.

Description of work.

T5.1 Landscape analysis (Lead partner: FCT [8PM], Contributors: CESNET [5PM], EGI.eu [5PM], CSIC [2PM], PSNC [2PM], IISAS [2PM], JISC [2PM], LNEC [2PM]) – **Task duration: M1-M20**

This task will conduct a landscape analysis by collecting information about the national policies, practices, roadmaps and strategies around funding, procuring, providing, accessing and sharing of services and resources in the EOSC scope. This goal will be achieved by liaising with funding agencies and policy makers in the countries covered by the project.

Special emphasis will be given to policies and practices that address the needs of different EOSC user groups (individual researchers, citizen scientists, research projects and collaborations, industries and SMEs) with a focus on the user communities that are already part of national research roadmaps.

The gathering of information will be organized by having dedicated meetings and special workshops for policy makers and funders. This activity will be coordinated with related activities conducted by ESFRI and e-IRG working groups.

The work will cover a wide range of policies, including those dealing with generic services for data production, processing and preservation, as well as those for thematic resources and services that target the needs of specific research communities. Transnational access to national resources will also be covered.

Existing policies and practices, where applicable, will be related to national and international regulatory frameworks (e.g. handling of privacy-sensitive data in compliance with GDPR).

The work will be done in close collaboration with task T5.3 to provide also the necessary wider background.

T5.2 Gap analysis and recommendations (Lead partner: CESNET [13PM], Contributors: FCT [4PM], EGI.eu [2PM], CSIC [4PM], PSNC [3PM], IISAS [3PM], JISC [2PM], LNEC [3PM]) – **Task duration: M1-M30**

Based on the input from T5.1 with the aid of information gathered through external collaboration (T5.3) this task will focus on the analysis of collected data. A gap analysis will be conducted by comparing the status of national organizational and regulatory frameworks (T5.1) against the need of integrated procurement and access provisioning, open science support and collaboration that are at the heart of the EOSC principles. Special emphasis will be given to the understanding of conditions, rules and policies that govern the transnational access to the national resources. To put the gap analysis into the national context, these requirements will be gathered from the user communities participating in the project, and those who are part of the national research roadmaps.

Recommendations will be identified by comparing the different landscapes at national levels, and by selecting the national best practices and strategies that will be considered more viable and more effectively applicable across the countries that are in the project scope.

The first part of the project will be devoted to the identification of the gaps, while the second phase will be focused on the definition of recommendations and their discussion, validation and promotion with funders and policy makers at national and international level.

Funders and policy makers will be involved in a series of national meeting and/or workshops. Where appropriate, international workshops (meetings) will be organized to deal with the cross-border collaboration and with conditions for transnational access. This stage of the T5.2 task will be concluded by an international workshop (may be collocated to some larger event) dedicated to the presentation of proposals and recommendations for the policy and strategy harmonization at the international level (T5.3).

The results of these discussions and work will be summarized in the final deliverable 5.3, where also proposals for the policy harmonization will be described.

T5.3 National and international liaisons (Lead partner: EGI.eu [8PM], Contributors: CESNET [4PM], DANS [5PM], FCT [3PM], CSIC [4PM]) – **Task duration: M1-M30**

T5.3 supports the other tasks of the project by focusing on the definition of the methodology and the implementation of an effective liaison structure and collaboration framework with policy makers and funders at the national and international levels. The liaison structure will involve:

- at the national level interaction with national funders and policy makers and potential national EOSC stakeholders;
- at the international level interaction with the EOSC Governance, EOSC-hub and other EOSC-related projects, especially those approved in this call in the subtopics (a) on the governance of EOSC and (c) on FAIR data uptake and compliance in all scientific communities, which includes data policy, practice and FAIR certification, and also relevant policy boards (e.g. ESFRI, e-IRG).

The task aims at achieving an effective coordination of policy-related activities at international level. This will be realized by establishing contacts with relevant national and international policy bodies, funding agencies and stakeholders.

The methodology will consist in:

- Identifying the relevant funders, policy makers and EOSC stakeholders that will be involved in the other WP tasks;
- Engaging such partners to establish an effective collaboration;
- Supporting the external partners in participating in the activities of WP5 through face-to-face meetings, workshops, webinars etc.
- Promoting the adoption of the WP5 recommendations within the countries in scope and beyond.

The task will ensure that the findings of this work package (landscape analysis, requirements, gap analysis and recommendations) are shared and promoted as widely as possible, and can be exploited by other initiatives with the aim of maximizing collaboration and synergies. One of the roles of this task is also to provide feedback to the EOSC governance on the Rules of Participation so they can be properly aligned with national priorities and

expectations.
Deliverables
D5.1 – National/International engagement plan with policy makers and funders – M6 (R, PU) – Lead: EGI.eu
D5.2 Landscape and gap analysis and recommendations – M19 (R, PU) – Lead: CESNET & FCT
D5.3 Feedback report on project policy recommendations – M28 (R, PU) – Lead: EGI.eu & CESNET
Milestones (mov = means of verification)
M5.1 – Landscape analysis finished – M9; mov: reports collected in internal website/shared space
M5.2 – Findings of D5.2 presented in a EOSC-related conference (e.g DI4R) – M21; mov: presentation
M5.3 – Series of the national workshops with national policy bodies and funding agencies – M24; mov: workshop (reports from meetings internally available)
M5.4 – International workshop on harmonization of national policies – M27; mov: workshop

WP6 - EOSC skills development [M1 – M30]

Lead beneficiary: PSNC

Participant number	1	2	3	**4**	5	7	8	10	15
Short name	CSIC	LIP	KIT	**PSNC**	CESNET	DANS	JISC	UPV	INCD
PMs per participant	15	3	3	**30**	3	10	24	3	3

Short description

WP6 is dedicated to support the professional development of capabilities and skills for EOSC stakeholders including research-performing organisations. In particular, we aim at expanding human capacity through a training programme and infrastructure that promotes training on the necessary skills for better utilisation of the EOSC services and supports the trainers involved.

WP6 builds on the skills & competences framework created by EOSCpilot and recommendations with regard to the training (EOSCpilot deliverables D7.1, D7.2 and D7.3).

Objectives

- Provide online platforms for content creation/hosting of training material and running hackathons
- Promote self-learning via a self-deployable training infrastructure, and the provision of cookbooks to deploy the most popular, services on demand
- Provide guidance and resources to support trainers in the best use of these platforms and to create quality resources
- Create a means of demonstrating quality through an accredited trainer program
- Work with national academic institutions to embed skills and resources
- Create guidelines to publicize and improve visibility of training materials.

For clarification, WP6 a) does not aim to be a catalogue of all EOSC materials, only those created on this platform; b) will not curate/update training materials; c) will rely on previous H2020 projects to identify a framework of skills; d) will focus only on developing materials to 'train the trainers' and selected materials of general interest.

Description of work.

T6.1 Deployment and customization of the training platform (Lead partner: PSNC 21PM, Contributors: CSIC 8PM, DANS 3PM, Jisc 3PM) – **Task duration: M1- M21**

This task will develop and deploy two online platforms to be used for MOOC training and Hackathon as a Service. (MOOC stands for Massive Open Online Course.)

MOOC platform

The central platform for MOOC training will be set up as-a-service to be included in EOSC Portal. The platform will be based on open source software to be chosen after evaluation of most commonly used platforms (e.g. Moodle, edX). It will be customised according to the identified service requirements (e.g. related with the taxonomy for discovery and filtering, requirements and methodology worked out in task 6.2) and extended with the new capabilities in terms of self-deployable training infrastructure and courses on the EOSC infrastructure.

Hackathon as a Service

This task will develop and deploy a platform to facilitate the organisation of hackathons on the EOSC infrastructure. This hackathon as a service will be built from different components evaluated, selected and customised for this purpose. The platform will be accessible through the EOSC Portal.

Using the hackathon as a service platform:

- Hackathon organisers can create and publish hackathons, describing the rules and technical requirements
- Users can find and register in hackathons.
- Infrastructure providers can offer their infrastructure to support hackathons.
- Dataset providers will provide datasets to be analysed.
- API providers can provide a set of functionalities to help participants in their development.

Best practices

In collaboration with task 6.2, the MOOC and hackathon platforms will be enriched with best practice guides, examples and tutorials on how to create EOSC-related tutorials using the new technical capabilities.

T6.2 Methodology and certification of the tutorials (Lead partner: Jisc 18PM, Contributors: DANS 7PM, PSNC 6PM, CSIC 3PM) – **Task duration: M1-M30**

In the past years a number of projects and initiatives have contributed to guidelines and best practices towards the development of skills in data science (e.g. Edison project, FOSTER, OpenScienceMOOC, Software Carpentry, Library Carpentry, EOSCpilot WP7).

We have identified a gap in what concerns the support and development of high-quality **online training material.** Task 6.2 will build on the work of previous initiatives to address this gap.

Based on the partner´s expertise we will create a training programme including support materials to equip trainers and educators to develop their own high quality online resources using the online platform developed in task 6.1.

In this way, we will contribute to build human capacity to support skills development in a sustainable way.

The main activities in this task will be:

1. Reviewing existing initiatives to identify lessons learned, materials for reuse and possible areas of collaboration

2. Developing quality criteria to ensure that the materials created are FAIR (as defined by the EOSCpilot deliverable D7.2). We will ensure that tutorials are described appropriately using the controlled vocabularies developed to map to existing frameworks (e.g. EOSCpilot skills and capability framework deliverable D7.2, Edison Data Science Competence Framework). We will develop guidelines to ensure that the training materials are pedagogically sound for all possible EOSC stakeholders. We will also explore the best ways to ensure

content is available in different languages.

3. A quality assurance and / or certification scheme specifically for online content. This task will explore the methodology for quality assurance for the provision of online content. This will involve consultation with various stakeholders, including training participants. We will explore opportunities to join forces with other EOSC training projects and develop a wider EOSC trainer accreditation.

4. Creation of a training programme for the development of teaching materials. This programme will include pedagogical approaches to online creation of tutorials adapted to different EOSC stakeholders. This may include face-to-face courses, online webinars and tutorials. Specific guidance will be included about how to use the online platforms (linked to task 6.1). We will also include guidance and recommendations for encouraging uptake and embedding, for example including guidelines in capstone projects, and guidance on making resources FAIR (building on EOSCpilot D7.2). We will develop a network / community of practice for trainers collaborating with other initiatives (linked to 6.3).

T6.3 Interaction with national educational programmes (Lead partner: CSIC 4PM, Contributors: LIP 3PM, PSNC 3PM, CESNET 3PM, UPV 3PM, Jisc 3PM, INCD 3PM, KIT 3PM) **– Task duration: M15-M30**

This task will establish contact and collaborate with Universities (in the participating countries) with programmes focussed on EOSC relevant topics. This collaboration will elevate the online EOSC training program to the quality necessary to be accepted as part of the universities' curricula.

The collaboration will benefit both sides: it will enrich university programmes with Open Science-focused technologies, will be an important channel to receive feedback on the methodology part and content of the tutorials, will allow to develop the EOSC-related skills and raise awareness among students.

Initial list of the universities to be contacted: Poznan University of Technology, Poznan University of Life Sciences, AGH University of Science and Technology, Adam Mickiewicz University, University Complutense of Madrid, International University Menendez Pelayo, University Polytechnic of Valencia, University of Zaragoza, University of Minho, University Institute of Lisbon, University of Karlsruhe, Newcastle University, University of Leeds, University of Leicester.

Deliverables

D6.1 – Training materials and quality assurance guidelines (R, PU) – M12 – Lead: Jisc

Two packages of materials, covering effective pedagogic design and development using the platform. A set of quality assurance guidelines would be produced.

D6.2 – Report on the final release of the customized training platform including the self-deployable tutorials capabilities, and Hackaton-as-a-Service platform (R, PU) – M21 – Lead: PSNC

This report will provide the feedback on the services delivered in the task T6.1

D6.3 – Final report about skills development support activities and related services (R, PU) – M30 – Lead: PSNC

This report summarises the methodology work done in task T6.2, as well as the university collaboration outcomes. In addition, the evolution of the services provided in the tasks T6.1 are described.

Milestones

M6.1 – PM3: Identification of tools to be used for platforms in T6.1

M6.2 – PM9: First release of the customized training platform including the self-deployable tutorials capabilities, and HaaS platform

Deliverable number and name	WP	Lead	Type, Dissem. level	Due
D1.1 – Dissemination and exploitation plan	WP1	CSIC	R, CO	M3
D1.2 – Dissemination and exploitation report	WP1	EGI.eu	R, CO	M15
D1.3 – Final dissemination and exploitation report	WP1	CSIC	R, CO	M30
D1.4 – Quality Plan	WP1	LIP	R, PU	M3
D2.1 – Roadmap for integration of national capacities into the EOSC	WP2	CYFRONET	R, PU	M5
D2.2 – Policy Gap Analysis	WP2	KIT	R, PU	M6
D2.3 – Intermediate report on integration efforts	WP2	CSIC	R+DEM, PU	M15
D2.4 - Final report on EOSC integration	WP2	IISAS	R+DEM, PU	M29
D3.1 – Software Maturity baseline	WP3	LIP	R, PU	M10
D3.2 – First prototype of Service Integration platform	WP3	CSIC	R, PU	M15
D3.3 – Intermediate report on technical framework for EOSC FAIR data principles implementation	WP3	DANS	R, PU	M12
D3.4 – Final release of Service Integration platform	WP3	CSIC	R, PU	M29
D3.5 – Final report on technical framework for EOSC FAIR data principles implementation	WP3	DANS	R, PU	M27
D4.1 - Best Practices Elicitation Report and Data Management Plans	WP4	UPV	R, PU	M6
D4.2 - First prototype of the EOSC Thematic services	WP4	UPV	DEM, PU	M16
D4.3 - Final release of the EOSC Thematic services	WP4	UPV	OTHER, PU	M24
D4.4 - EOSC Thematic services validation report	WP4	UPV	R, PU	M30
D5.1 – National/International engagement plan with policy makers and funders	WP5	EGI.eu	R, PU	M6
D5.2 – Landscape and gap analysis and recommendations	WP5	CESNET	R, PU	M19
D5.3 – Feedback report on project policy recommendations	WP5	CESNET	R, PU	M28
D6.1 – Training materials and quality assurance guidelines	WP6	JISC	R, PU	M12
D6.2 – Report on the final release of the customized training platforms	WP6	PSNC	R, PU	M21
D6.3 – Final report about skills development support activities and services provisioned	WP6	PSNC	R, PU	M30

3.2 Management structure, milestones and procedures

3.2.1 Project governance and management

The project is structured having a mind a clear separation of functions between Management and Governance. The Governance structure is defined to provide strategic guidelines to the management of the project. The responsibility of implementing the strategic goals in the day-by-day work of the project is on the managerial level.

The Management of the Project is the responsibility of the **Project Coordinator**. The Coordinator will ensure the timely delivery of project objectives and deliverables by continuously monitoring the project progress against the planned activity. The Coordinator identifies and tracks issues as well as proposes suitable corrective actions (i.e. resource reallocation, task force creation, etc.) that might require a formal decision by the consortium. The Coordinator will also act as the official point of contact between the Commission and the Beneficiaries.

To the effect of the Project Governance, a **Project Management Board (PMB)** will be created and chaired by the Project Coordinator in the first half of the project, and by the deputy coordinator in the second half. The project coordinator and WP1 leader is Isabel Campos (CSIC), the WP1 deputy coordinator is Jorge Gomes (LIP).

The PMB will be responsible for making decisions affecting project strategy, including risk management, and for ensuring that the activities are effective and carried out according to the agreed upon schedule. It is also responsible for the quality of the project outcome, including deliverables. The PMB is composed of the WP leaders (leader + deputy) and meets at least 3 times a year.

A **Collaboration Board (CB)** will be created, which will be composed by one representative per partner. The CB will be responsible for making decisions affecting the composition of the Consortium, resource allocation, for approving changes in the Description of Work (DoW), and for all other decisions having a direct legal or financial impact on project beneficiaries. The CB Chair should be elected by the CB among its members and appointed for a fixed time.

The **Work Package Leaders** are responsible for scientific and technical work of their respective Work Packages. This includes planning and control of all activities within the Work Package, preparation of deliverables and the collection of contributions from other partners participating in the respective Work Packages for internal and external reports. They are expected to raise critical issues to the PMB and provide support in coordinating cross-work package relationships within the appropriate activity area (e.g. through the participation on cross-work package task forces). Each WP will appoint a deputy coordinator, belonging to a different partner in the consortium. The rest of **WP leaders** are: WP2 - Marcus Hardt (KIT); WP3, Mario David (LIP); WP4 - Ignacio Blanquer (UPV); WP5 - Ludek Matyska (CESNET); WP6 - Marcin Plociennik (PSNC).

The project will create a **Technical Integration Board (TIB)** with the role of advising the PMB in the decision-making processes that affect integration and harmonization of infrastructures and services. The rationale for creating this board is having a cross-work-package technical body, with the remit of analysing the evolution of Cloud technologies in the lifetime of the project, and provide informed advising to the PMB when it comes to select integration technologies for EOSC-synergy resources.

The TIB has an import role in the project, as it will provide informed technical advice on how to optimize the integration solutions by adapting them to the fast changing technical environment of the Cloud ecosystems.

The members of the TIB will be selected inside the project consortium members. It will have a maximum of two technical experts, for each relevant integration area: computing, storage and data repositories. Nominations coming from the WP leaders will be analysed and appointed by the PMB as part of the project setup procedures (Month 1).

In EOSC-synergy, Management and Governance need to have in mind the strong collaborative needs of the present call. In particular the necessity to develop collaborative interactions with other projects approved in the same call, and in the EINFRA-12-2017 call. The Governance structure needs to reflect this goal, and thus be prepared to create the appropriate discussion channels to support such networking.

Interactions with the EOSC governance project in subtopic (a), and the FAIR data project in subtopic (c) requires the implementation of the formal communication channels to provide mutual feedback. The policy work package contains specific measures proposed in that direction. Following the need for a close-cooperation, EOSC-synergy will discuss with the other projects approved in subtopic (b) the setup of a coordination framework to establish the collaboration links to discuss the individual roadmaps of the projects from the technical and integration point of view, and the opportunities for collaboration and exchanging of mutual feedback.

3.2.2 Innovation Management

The innovation management will be coordinated by a dedicated task T1.3, and will rely on T1.4 for communication. The innovation management system will adopt open innovation as collaboration approach to ensure that productive interactions can take place across the various distributed actors of the consortium and with the external stakeholders. This activity will be extremely important to bridge the gap between the national levels and the European-wide coordination structure.

The project includes a range of innovative thematic services identified during the project preparation through the identification of novel mature services deployed at the national level. In the course of the project additional services with innovative potential will be considered for integration. The design of the service integration will be based on the continual improvement as designed in WP3 task T3.2.

3.2.3 Quality Management

The goal of Quality Management is to ensure that the stakeholder expectations are properly recognised and met, reduce mistakes, identify and recommend necessary changes for improvement and ensure participation of all members of the project team meet project objectives.

Regarding conflict resolution, at all levels within the project, i.e. in the Work Packages and the Governance Boards, decisions will be made in a consensual style as often as possible. Disputes within a work package that cannot be resolved by the work package leader should be referred to the Coordinator. The formal procedure to deal with such conflicts will be detailed in the Quality Plan deliverable. Additionally, the Consortium Agreement will describe the process for settlement of disputes

Any event that may jeopardize the overall completion date of the Project should be reported immediately to the Coordinator, which will call an emergency CB meeting or teleconference as required. Each party involved in the issue must present a short document describing their respective understanding of the conflict that includes at least one proposed solution. The CB reviews the conflict documents and following the procedures of the CB, each

member votes on one of the proposed solutions. The solution receiving the simple majority is implemented with the chairperson casting the tie-breaking vote as necessary.

The project is planning to manage quality through the following main two controls:

- Consortium Agreement with roles & responsibilities for mutual obligations.
- Quality & Risk Management Plan with roles, responsibilities and guidelines for communication, outputs, documentation, review process, reporting and risk management.

In addition, Quality Management will build on the following processes:

- Quality Assurance to assess if the guidelines defined in the Quality Plan are being followed and whether these are still appropriate for the project. Project outputs will be reviewed according to the review process for deliverables and milestones.
- Quality Control to collect and monitor the Key Performance Indicators (KPIs) and activity metrics, to identify improvements and suggest implementation actions, and to collect the lessons learned.

3.2.4 Strategy for knowledge management and protection

EOSC-synergy will provide an integration and operation layer of digital services for research and innovation that is based on open source licences and open specifications.

The approach to knowledge management is to ensure that all exploitable outputs are captured in a common catalogue of project results. Outputs will be assessed prior to disclosure for opportunities and benefits of IP protection, taking necessary steps in relation to ownership and management, and taking necessary action to secure the IP rights as needed.

All consortium partners have an active interest in establishing well-defined guidelines and policies for the management of knowledge, intellectual property and innovation from the early stages of the project. The Consortium Agreement will detail matters related to IPR regulation and legal issues and, in addition, guidelines will be included in the project quality management deliverable (D1.4) and will be continuously verified by the project management and amended if any improvements can be introduced. The project management will be responsible for the management of intellectual property and for the resolution of any IPR problems that occur.

The IPR policy will handle the following issues:

- *The Consortium Agreement* will establish a legal framework and provide regulations for issues related to work, IP ownership, access rights to background, results and any other matters of the consortium's interest.
- *Access rights to background and results*: the project partners will grant each other and their affiliated companies, royalty-free access rights to their background and results for the execution of the project.
- *IP ownership*: Pre-existing know-how and information related to the use of knowledge owned by individual partners from work carried independently of the EOSC-synergy project will be protected. The project partner carrying out the work leading to such results shall own results. If any results are created jointly by at least two project partners and it is not possible to distinguish between the contributions of each of the project partners, such results will be jointly owned by the contributing project partners.

- *Exploitation strategy of the obtained results*: Partners who own rights are encouraged to exploit these results and disseminate them to external bodies such as international journals and conferences and contributions to standards bodies. A more detailed joint exploitation plan will be implemented during the project as part of Task 1.3.
- *Contingency plan*: will be defined to ensure the access to knowledge crucial to the project development if a partner with specific IPRs leaves the consortium. Policies for the partial or full transfer of ownership of results between partners should be defined.
- *Open Access*: The consortium is fully committed to the European Commission requirements to support of open access for published articles. All scientific publications of project's results will be granted open access according to publisher and law regulations as set out in the grant agreement.

3.2.5 Risks Management

Likelihood	Impact			
	Minor	**Moderate**	**Major**	**Catastrophic**
Unlikely	Low	Low	Medium	Medium
Possible	Low	Medium	High	High
Likely	Medium	High	High	Extreme
Almost Certain	Medium	High	Extreme	Extreme

During the project, the project management will conduct planning, identification, analysis, response planning and control. The objective will be to detect threats and decrease their likelihood and impact by proper treatment as well as to collect lessons learned from risks occurrence to facilitate continuous learning of project management team.

The project will create a risk registry in partnership with the work package leaders and the Project Management Board. It will be kept confidential, with access restriction to PMB members only. A detailed risk management plan will be created at the start of the project and integrated within D1.4 (M3).

Table 3.2b: Critical Risks for Implementation

Only risks evaluated as Risk Level of at least "high" that could have a high adverse impact on the ability of the project to achieve its objectives are included (see Section 3.2.4 for the definition). All other risks will be managed by the Risk Management activity in WP1. (L=Likelihood, P=Possible, I=Impact, Maj=Major).

Description of Risk	L	I	WP	Proposed Risk mitigation
Fail to attract new resource providers to EOSC	**P**	**Maj**	**WP2**	Dissemination at regional/national level targeting digital infrastructures; bridging with the resource providers; training infrastructure managers to overcome integration issues.
Lack of opportunities or willingness to integrate common services; federation & collaboration services vertically or with added value and thematic services	**P**	**Maj**	**WP2, WP4**	Dissemination at national level; bridging with user communities and resource providers; facilitate integration processes.
Dependencies on current EOSC services being pursued in different	**P**	**Maj**	**ALL**	Close coordination with ongoing EOSC-related projects. As national infrastructures try to foster

projects				stability on the EOSC core functions.
New services fail to attract new users beyond the original target	**P**	**Maj**	**WP2, WP4**	Usage of services will be monitored so that it will be possible to better tune the training and dissemination activities to support adoption.
Lack of engagement with the proposed Software Quality procedures	**P**	**Maj**	**WP3**	Targeted dissemination towards open source communities and promoting software quality and best practices.
Lack of engagement towards FAIR	**P**	**Maj**	**WP2, WP5**	Dissemination of FAIR data practices; raising awareness on the methodology and the tools via the project communication channels.
Lack of engagement of National Authorities	**P**	**Maj**	**WP5**	Liaise with EOSC governance bodies and other countries already engaged, in a coordinated effort to reach out national authorities.
Low adoption of the training platform and self-training infrastructure	**P**	**Maj**	**WP6**	Improve the dissemination strategies towards concrete stakeholder groups; contribute to create more appealing training material.
Difficulties in establishing constructive interactions with EOSC-related projects	**P**	**Maj**	**WP1, WP5**	Organize concertation meetings with the support of the EOSC Executive Board and the EC.

3.3 Consortium as a whole

The EOSC-synergy consortium brings together key research performing organizations in Europe with a strong commitment in fostering multidisciplinary scientific and technological research, and in transferring this knowledge to society. Several of these organizations are also major providers of e-infrastructure resources that support their scientific communities at the national and international level.

The EOSC-synergy consortium has extensive and long-standing expertise in supporting the full lifecycle of advanced services acquired through state-of-the-art research and by providing support to national communities and international scientific collaborations at large scale. Provisioning of data and computing services to research projects at the frontier of knowledge requires the development of specialized knowledge at the research infrastructure centres. This expertise spans over domains such as resource federation management, system software adaptation, client tools development or complex service integration. Such advanced services are typically based on the latest technological developments, and need to be brought to production, by a process of service consolidation that usually takes place at the research infrastructure resource centres themselves.

Many of the most prominent institutions involved in the federation, deployment and operation of distributed computing and data management across Europe are engaged in EOSC-synergy. This includes very large public resource centres (specifically: CSIC, BSC, CESGA and CIEMAT in Spain; LIP and INCD in Portugal; KIT in Germany; PSNC and CYFRONET in Poland; CESNET in the Czech Republic; IISAS in Slovakia). This constituency secures a coherent approach in terms of resources for the EOSC Capacity expansion, and guarantees the capacity needed to support the activities proposed to develop our Capability building program.

An important challenge for the deployment of EOSC is the implementation of harmonized data access mechanisms, integrated with the EOSC core services, that respect EU directives in Cloud security and data

privacy. The consortium is very well positioned to address this challenge through the expertise of KIT, CYFRONET, LIP, CSIC, PSNC, CESNET and IISAS. As a result of this activity, major national data infrastructures such as the Helmholtz Data Federation (HDF) in Germany, the data repositories hosted by PSNC, BSC, Digital CSIC or INCD in Portugal, will be accessible via standard tools.

Also regarding integration of national data repositories in EOSC, a fundamental aspect to cover in this call is fostering the uptake of FAIR methodology. In this respect the consortium counts with key institutions in Europe such as KNAW-DANS or JISC. In the framework of EOSC-synergy, national agencies in charge of policy development for research data, such as FCT-FCCN (Portugal), Digital CSIC (Spain) and PSNC (Poland) will pioneer the integration of national repositories in EOSC, implementing the FAIR recommendations and guidelines of the project approved in the subtopic (c) of this call.

Relevant Research Communities internationally consolidated have been selected by the consortium to validate the service integration solutions provided in the course of the project.

The Earth Observation area is represented through a number of sub-areas strategically relevant at the country level. A good example is the Marine Coastal Observation and Modelling community from Portugal (represented by **LNEC**) and France (represented by **IRD/LEGOS**). Production ready technologies used as a service at ESA and several national programs for managing Earth Observation data will be integrated into EOSC and validated by INDRA.

The **Institute of Meteorology and Climate Research of KIT**[6] will validate the integration in EOSC of the service to measure Ozone concentration in the atmosphere, O3AS, which is a key service used by the **United Nations Environment Program (UNEP).** The "Sand and Dust Storm Warning Advisory and Assessment System" is a World Meteorological Organization (WMO) service of high impact for weather forecast and predictions in the Mediterranean area, which will be validated by the **Earth Sciences research group of BSC**[7].

Integration of services related to environmental analysis of water has a direct impact in the implantation of EU directives on water quality in Slovakia. In particular making existing data of water supply more transparent through appropriate metadata, as a method of implementation of the INSPIRE directive. Such technique can be applied as well in other countries such as the Czech Republic, where the **Research Centre for Toxic Compounds** will be in charge of validating the data service integrated in the context of EOSC-synergy.

The integration of data services for the area of Biomedicine has a big impact in several countries involved in the proposal. It aligns properly with the role of BSC in Spain to facilitate the access to Core Data Resources including EGA (European Genome-Phenome Archive) via the OpenEBench service. The Scipion service developed by CSIC, and expanded to the users of the Czech Republic in Masaryk University is also a very good example of cooperation between countries to serve the international community of 3D Electron Microscopy.

The Capacity and Capability expansion towards South-America builds on the background and collaboration history of **IBERGRID** in this geographical area. In particular **UPVLC** and **CIEMAT** will be involved in activities to develop interoperation mechanism with the Brazilian National Cloud of RNP, and enabling access through EOSC to non-European Open Data repositories.

[6] http://www.imk-asf.kit.edu

[7] https://www.bsc.es/discover-bsc/organisation/scientific-structure/earth-sciences

In the area of skills development, the consortium counts on highly reputed institutions in Europe, with links to the national Universities as indicated in WP6, to guarantee the activity of disseminating the work of the project in the respective countries, and building the human capabilities required to exploit EOSC. In particular, **JISC** in the UK, **DANS** in the Netherlands, **CSIC** in Spain, **FCT** and **INCD** in Portugal have mission to outreach and training national communities. **PSNC** has also a broad expertise in maintaining training services for the Polish universities, and develops an extensive program in training software developers in the country.

3.4 Resources to be committed

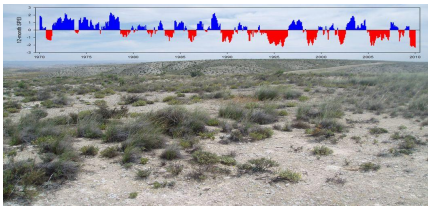
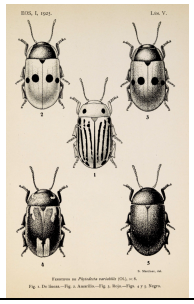
TABLE 3.4a - Table showing number of Person Months (PMs) required

Participant	WP1	WP2	WP3	WP4	WP5	WP6	TOTAL PMs /participant
1-CSIC (ES-JRU)	25	54	34	68	10	15	**206**
2-LIP	8	13	34	10		3	**68**
3-KIT		37	5	15		3	**60**
4-PSNC		25			5	30	**60**
5-CESNET		10		15	22	3	**50**
6-IISAS		20		10	5		**35**
7-DANS			25		5	10	**40**
8-JISC					4	24	**28**
9-IRD				20			**20**
10-UPV			10	30		3	**43**
11-LNEC			15	47	5		**67**
12-CYFRONET		30	15				**45**
13-FCT					15		**15**
14-EGI.eu	20				15		**35**
15-INCD		27				3	**30**
16 - INDRA				24			**24**
TOTAL PMs	**53**	**216**	**138**	**239**	**86**	**94**	**826**

Table 8. Expansion in terms of new national resources to be integrated into EOSC. Resources will be made available to EOSC-Synergy for service validation and to support integration of thematic services.

Type and size	Access model	Usage information
CZECH REPUBLIC: CESNET		
Cloud ~5,000 cores (end of 2019)	National and European collaboration	We expect part of this total capacity to serve the project thematic services
Online storage ~10 Petabytes	National	RECETOX and CryoEM, these are expected to get guaranteed capacity
GERMANY: KIT (Helmholtz Data Federation)		
Storage facility 10 Petabytes online	National; Available to KIT researchers, partners and the Helmholtz Association	Covers the earth and environment, matter, health, and energy scientific domains. Serves the Helmholtz Data Federation (HDF), EUDAT's B2SAFE and LSDF Online Storage service.
POLAND: PSNC and CYFRONET		
2 HPC systems ~ 4 Petaflops	National	Covers the physics, chemistry, biology, mechanics, engineering simulations and astrophysics domains.
Cloud ~ OpenStack 6,000 cores	National and EU collaboration	Demonstrators of R&D projects - National and H2020 Biology, bioinformatics, chemistry & astrophysics domains, and SMEs and industry
Online Storage ~ 10 Petabytes	National	National communities using the cloud infrastructure
Long term storage ~ 28 Petabytes	National	Available to biology, physics, chemistry, cultural heritage, universities, and researchers in the long tail of science
PORTUGAL: LIP and INCD		
Cloud Cluster ~ 5000 cores	National and European collaboration	Available to all areas of knowledge.
Online storage LUSTRE ~ 3PBytes	National and European collaboration	Available to all areas of knowledge.
SLOVAKIA: IISAS		
Cloud Cluster ~ 8,000 CPU/GPUs	National and Europeancollaboration	Operative end of 2019
Online storage ~5 Petabytes		Operative end of 2019
SPAIN: CSIC, CESGA, BIFI and CIEMAT		
HPC systems ~ 25.000 cores	National and institutional	Covers high energy physics, astrophysics, biology, medicine, engineering, climate.
Cloud ~ 24,000 cores	National and European collaborations	Dedicated support for Open Science projects at CSIC. Expected to become operational in 2019
Online storage (CSIC, CESGA) GPFS and LUSTRE 15 PBytes	National and European Collaborations	High energy physics, astrophysics, biology, climate, and materials science Hadoop-based BigData Infrastructure.
Online storage ~ CSIC, 5 Petabytes	National and European Collaborations	Dedicated support for Open Science projects at CSIC
Long term storage 3 Tape systems ~ 60 Petabytes	National and European Collaborations	High energy physics, astrophysics, humanities and climate science

Table 9. National Data Repositories to be used for FAIR integration as first pilots. More repositories coming from thematic communities and national catch-all repositories will be identified and integrated following the developments in Task T3.3

SPAIN - DIGITAL CSIC REPOSITORY the largest institutional repository of Spain. It was launched in January 2008 as a result of this institutional commitment with Open Access and to date, with nearly 165,000 works available.		
Collection Name and location	**Description**	**Access and status**
SPEIbase *Experimental Station of Aula Dei; Institute of Ecology of the Pyrenees.* Project website: http://spei.csic.es/ Latest Dataset https://digital.csic.es/handle/10261/153475 Associated software: https://digital.csic.es/handle/10261/10002	The SPEI (Standardised Precipitation-Evapotranspiration Index) is a multiscalar drought index based on climatic data since 1901. It can be used for determining the onset, duration and magnitude of drought conditions with respect to normal conditions in a variety of natural and managed systems such as crops, ecosystems, rivers, water resources, etc.	Updated versions of the datasets are being uploaded into DIGITAL.CSIC since 2010. All versions of datasets currently amount to 120-130 GB. 
Projects of Ethnoarchaeology in Southern Cone ***(Institución Mila i Fontanals - CSIC)*** https://digital.csic.es/handle/10261/155111 https://digital.csic.es/handle/10261/158444 https://digital.csic.es/handle/10261/164703 	From 1986-2005, a multidisciplinary team from Spain, Argentina and Chile carried out projects on the Isla Grande of Tierra del Fuego, Argentina. The objective was the evaluation and development of archaeological methodologies commonly used in the investigation of prehistoric societies.	Pilot of Open Science project as the authors are uploading all types of results generated through the whole research cycle. The collections already available are a mix of research results in different formats (e.g. text, audiovisual, image, databases) and each file ranges on average between 700kB and 6MB.
EOS, Spanish Journal of Entomology, (National Museum of Natural Sciences) *https://digital.csic.es/handle/10261/134880*	EOS was created in 1925, and soon after its launch the journal became well known amongst main research centers and entomologists in the world and contributed to giving it a marked international profile. It is the oldest entomological journal in Spain and it is still a reference resource for global community of entomologists, particularly those devoted to zoological taxonomy.	Researchers, students, cultural heritage community, citizens curious about zoology in general and insects in particular. 

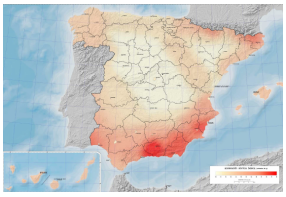
Collection of Earth Sciences datasets (Instituto de Ciencias de la Tierra Jaume Almera) *https://digital.csic.es/handle/10261/101879* 	Datasets about seismic activity in the Iberian Peninsula and data from research projects about Earth's Structure and Dynamics and Crystallography, and Environmental Geology and Geohazards.	Useful Information for researchers, policy makers, developers, R&I sector. Ongoing initiative. Files sizes greatly vary, with some in the range of kB and others exceeding 20GB.
Poland		
Collection Name	**Description**	**Access and status**
AMUNATCOLL ***"AMU Nature Collections"*** *AMUNATCOLL - digitization and sharing of natural data resources of the Faculty of Biology of of Adam Mickiewicz University-Poznań* AMU website: **http://international.amu.edu.pl/** Associated software: **https://dingo.psnc.pl/en/dlibra-en/**	The scientific value of the natural collections collected at the Faculty of Biology of Adam Mickiewicz University is at the core of the AMUNATCOLL project implementation. For decades, the activities of WB UAM (and earlier UAM structures conducting research and educators in the field of natural sciences) collected one of the largest in the country, and in some areas in Europe, a collection of botanical, mycological and zoological specimens. These collections are not only archival, but also have an invaluable scientific value, giving great potential for research and didactic work, as well as for use in cultural, socio-economic and legal contexts.	Digitalization of collections of plants, fungi and animals giving 2 million unique records. The plan assumes the digitization of size resources 960 TB. Database compatible with standard GBIF (The Global Biodiversity Information Facility), dedicated API interface The project is in its initial phase of digitisation 

Table 10. International Data Repositories proposed to be integrated with EOSC services

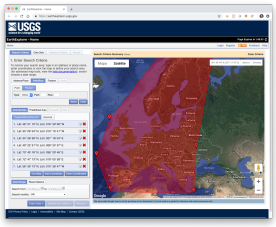
Non-European-based Data repositories		
Collection Name	**Description**	**Access and status**
Latin American Giant Observatory (LAGO) *http://lagoproject.org*	The LAGO is an extended cosmic ray observatory composed of a network of water Cherenkov detectors. This detection network is designed to make detailed measurements of the temporal evolution of the radiation flux coming from outer space at ground level. LAGO is mainly oriented to perform basic research in three areas: high energy phenomena and the extreme universe, space weather and atmospheric radiation at ground level	The repository uses data management standard tools in Distributed Computing to store both experimental measurements and simulated data. In the framework of EOSC-synergy the repository will be made accessible via the EOSC portal. 
United States Geological Survey (USGS) - LANDSAT data *https://landsat.usgs.gov/*	At over 40 years, the Landsat series of satellites provides the longest temporal record of moderate resolution multispectral data of the Earth's surface on a global basis: the world's longest continuously acquired collection of space-based moderate-resolution land remote sensing data. It is a unique resource for those who work in agriculture, geology, forestry, regional planning, education, mapping, and global change research. Landsat images are also invaluable for emergency response and disaster relief.	Landsat Level-1 data products can be downloaded at no charge from different applications. Bulk requests can be performed freely by means of authenticated users. 

Table 11: Summary of the thematic services that EOSC-synergy will bring to the EOSC ecosystem

Service	Domain	Added value for European researchers
WorSiCa: European-wide service for the detection of the coastlines changes, coastal inundation areas and inland waterbodies	Earth Observation	Will enable researchers to generate maps of coastal and inland regions. These products can be useful for emergency and prevention methodologies in case of inundations or reservoir leaks. Societal impacts of the service include: preservation of lives during an emergency or efficient management of water resources targeting water saving in drought-prone areas
G-Core as a data manager allows for developing added-value products for Earth Observation (EO)	Earth Observation	Will democratize the usage of EO data out of the scope of nominal fields. EO data has the potential to become the starting point for a great variety of added value services. Will help defining new products and services mixing Earth Observation data with other type of data for scientific and social environments.
SAPS-as-a-service to compute the Surface Energy Balance Algorithm for Land	Earth Observation, Agricultural Sciences	Will provide wider access to knowledge on the impact of human and environmental actions on vegetations, leading better forest management and analysis of risks.
OpenEBench: a platform to support communities-led benchmarking efforts across Life Sciences.	Genomics	Scientists will be able to have up-to-date collections of analytical -omics workflows, which can be deployed across heterogeneous systems.
Scipion: on demand deployment and customisation of processing tools in the EOSC Portal.	Structural Biology	Researchers can have all their data pre-loaded loaded on a cloud instance powered by EOSC compute resources on the back-end. This means that scientists with minimal computational background (or compute resources of their own) can access the latest tools as well as powerful computational resources.
SDS-WAS: a service for modelling dust transport in the environment	Climate and Weather Forecast	This service is already used by industrial users (e.g. Gas Natural Fenosa; NCAR; Meteoplay; the Airport of Dubai). The integration in EOSC will make the service available to European researchers.
MSWSS for analysis of water network distribution integrated with **EPANET** for on-line analysis of toxics in drinking water supply	Environment	The MSWSS and EPANET integration in EOSC will transform technology into an affordable and easy-to-use package which could be deployed on a nation-wide scale. This will contribute to the unification of crisis management methods.
O3AS: a workflow comprising novel analysis tools that access Climate Data Archives	Climate and Environment	Researchers to study, for example, how stratospheric ozone protects life on Earth from harmful UV radiation by providing robust results of ozone projections for impact studies to gauge potential damage and avoided damage.
RECETOX integrated with a Federated Cloud back end	Environment	Researchers will be able to use RECETOX – an mass-spectrometry analysis tool used in toxicity studies – which is currently offered offline to a small number of scientists.. By integrating it in EOSC following the INSPIRE directive, the service will become open to all researchers at a larger scale.

Table 12: Data and Processing services that EOSC-synergy will bring to the EOSC ecosystem

Area	Data Services	Processing services
Earth Observation	• **Copernicus Missions Data**: INDRA to identify which data services from the https://spacedata.copernicus.eu/web/cscda/missions . • **United States Geological Survey** (USGS, https://www.usgs.gov/): Landsat images (SAPS). • INTA Data: **PAZ satellite** https://directory.eoportal.org/web/eoportal/satellite-missions/p/paz • WorSiCa Data: Copernicus satellite data, **EMODNET-Physics, EMODNET- Bathymetry**	**WorSiCa**: detection of the coastline, coastal inundation areas and inland water bodies water detection using remote sensing (satellite and UAVs) and in situ data (from field surveys). **G-Core** technology, for managing the Earth Observation data in terms of the acquisition, storage, cataloguing and processing data. **SAPS** processing of SEBAL and similar algorithms in the cloud with providing mechanisms to allow the sharing of the processed data following the FAIR principles.
Environment	• The **SDS-WAS Data Repository** is made by a folder tree organized by years and months with daily forecast data files in netCDF format for each model part of the ensemble. Accessible by direct manually download and http API automatic download, for single files and for entire months or years. Two users profiles are available (both by registration): the public profile is able to download data with two days of delay and the restricted profile can download in real-time. • **Data Repository from the Association of Water Companies** (Slovakia) and its partners and from the DHI Slovakia as a water supply consulting company and its clients.	• Several modelling techniques for dust forecast have been applied and offered through national services (e.g. **https://sds-was.aemet.es/forecast-products/dust-forecasts**). • Several hydraulic models for water supplement system have been applied, one of them is also open source model EPANET (**https://www.epa.gov/water-research/epanet**) By the integration of these services in the EOSC more complex derived services can be built.
Cimate	• HDF (**Helmholtz Data Federation**) and LSDF (Large Scale Data Federation) at KIT for replication, caching and storing repository data, as well as storage of additional simulations • The **ESGF** consists of federated data centres (including CEDA and DKRZ) that enable access to the largest archive of climate data world-wide. • Collaboration with **CEDA for accessing chemistry-climate project data.**	ForHLR-II and JUWELS are HPC infrastructures within the Helmholtz Association of German research centres (HGF). Essential requirements on the hardware are fast I/O and tiered storage solutions. The use of GPUs will be evaluated in the future.

Life Sciences and Biomedicine scientific communities	• **ELIXIR Core Data Resources including EGA** (*European Genome-Phenome Archive* - https://ega-archive.org) at both data and metadata levels for integrative analysis. • **Instruct-cloud CRYO-EM Database** (output of the project EOSC-Life).	**Scipion** is already a very matured software, with thousands of downloads per year, has been the subject of one of the EOSC Pilot Science Demonstrators, and is already prepared for automatic cloud deployment. **OpenEBench** is a platform for supporting scientific communities-led benchmarking efforts across different domains in Life Sciences.
Thematic services for the Astrophysics scientific community.	• **LAGO Measurements and Simulations Database**	LAGO is oriented to perform basic research in three areas: high energy phenomena and the extreme universe, space weather and atmospheric radiation at ground level. It is also producing interesting results in determining radiation fluxes received by airplane cabin crew or volcanology forecasting from muology determination.

EOSC-synergy

Full name

European Open Science Cloud – Expanding Capacities by building Capabilities

PROPOSAL – Technical Annex

Sections 4-5: Members of the consortium & ethics and security

Table of Contents

4. Members of the consortium

The EOSC-synergy proposal has been developed under the coordination of **IBERGRID**, the Iberian scientific and technological cooperation framework signed by Spain and Portugal in 2003.

The so called *Plan Ibérico Común* adresses the implementation of a common plan for the interconnection of Research Networks and the coordination of Scientific applications, distributed computing infrastructures, supercomputing infrastructures and data repositories[1]. Therefore this call is an optimal framework to exploit the cooperation model built at the Iberian level.

The consortium counts with the participation and expertise of the main research infrastructure and data providers in Germany, Poland, Czech Republic, Slovakia, experts in policy development in the Netherlands and the United Kingdom. The validation of services at the level of user communities extends as well to France via the research community of Marine Earth Observation.

The partners share a long history and collaboration in development and provisioning of IT services to research, as described in the section dedicated to the "Consortium as a whole".

- *Cloud Expansion towards Brazil*

As described in previous sections, the consortium extends activities to Brazil, in particular to the **Universidade Federal de Campina Grande (UFCG)**, which will act as a link to develop an interoperability program with the **Brasilian Cloud of RNP**. UFCG is the Brasilian coordinator of the H2020 project **ATMOSPHERE** ("Adaptive, Trustworthy, Manageable, Orchestrated, Secure, Privacy-assuring Hybrid, Ecosystem for Resilient Cloud Computing", one of the flagship projects of the EU-Brazil cooperation framework in H2020.

EOSC-synergy will support UFCG in joining the meetings of the consortium with travel budget, which has been allocated to the coordinator.

CIEMAT will act as link with the LAGOS experiment, in order to make available thorugh EOSC the data repositories of the experiment.

- *Letters of Support*

The project has received letters of support from the LAGO Experiment and from the project EOSC-hub. The letters are attached at the end of this document.

[1] https://wibergrid.lip.pt/site/institutional-information

4.1 Participants

4.1.1 CSIC

CSIC is the legal entity coordinating EOSC-synergy. As institution it participates with three research institutes: IFCA, CNB and Digital CSIC.

The department of *Advanced Computing and e-Science* of the **Institute of Physics of Cantabria (IFCA) will provide the coordination and management effort.** Besides, IFCA is involved in Fostering the Adoption of Services (WP3) for EOSC, and Harmonization of Technical Standards (WP2). IFCA will also articulate the input of the Spanish institutions in the project to the area of national policy harmonization.

DIGITAL.CSIC will participate in the WP3 and WP6 activities regarding best practices and harmonization of adoption of FAIR data in national repositories.

DIGITAL.CSIC is the reflect of the institutional commitment of CSIC with Open Access. It was **created in January 2008 and to date, with nearly 165,000 works available, it is the largest institutional repository in Spain**. The repository is a project under the coordination of the Unit of Scientific Information for Research. DIGITAL.CSIC got awarded with the **Data Seal of Approval in 2015**.

DIGITAL.CSIC organises, preserves and enables open access to a wide variety of research outputs (ranging from publications and conference contributions to non-traditional outputs such as datasets, research software and notebooks). DIGITAL.CSIC is actively involved in a number of international and national initiatives, including membership to several Open Access Repositories working groups under the coordination of REBIUN (National Network of University Libraries) and FECYT in order to promote common policies and standards; FESABID (Spanish Federation of Archives, Libraries and Museums Associations) Copyright National Strategy Working Group, DataCite Metadata Working Group, COAR Controlled Vocabularies Working Group and participation in Open Science initiatives initiated by OpenAire, LIBER Europe, RDA, JISC, Science Europe and Spain's National Library.

The **National Center of Biotechnoly (CNB)** will participate via the research unit on Three-dimensional Electron Microscopy under cryogenic conditions (cryo-EM). In particular in the deployment and validation at the EOSC level of the data analysis service Scipion (WP4). The specific version of Scipion developed for Cryo-EM facilities is currently being at the heart of some key EU resources, such as Diamond/eBIC –probably the largest cryo-EM facility in the world-, in Oxford, SciLab, in Stockholm and the cryo-EM line of the European Synchrotron (ESRF).This work will take place in the Work Package 4.

Key personnel

Dr. Isabel Campos Plasencia *[F], is the coordinator of EOSC-synergy. She is staff researcher at CSIC since 2008. She is the Coordinator of the Spanish distributed computing network (es-NGI), representative of Spain in the EGI Council, and coordinator of IBERGRID in Spain. As a member of the 2nd EOSC High Level Expert Group appointed by the EC, she has played an important and very active role in providing recommendations and policy baselines to support the launching of the EOSC initiave.*

As scientific background, she holds PhD in Theoretical Physics by the University of Zaragoza in 1998 in the area of computing simulations for Lattice Quantum Chromodynamics (Lattice QCD). She has held positions as research associate at DESY in Hamburg, Brookhaven National Laboratory (USA), and at the Leibniz Computer Center in Munich. She has a deep experience in parallel computing simulations for Lattice QCD in HPC systems, and has been involved in a number of projects to develop special purpose computers. She has an extensive experience in integration of applications and data management in the application areas of HEP, Environment, Nuclear Fusion and Complex Systems. She has over 65 publications in peer-reviewed journals and has participated or presented more than 140 communications to international conferences.

Dr. Jesus Marco de Lucas *[M], is a CSIC research professor at IFCA and current Vicepresident of Scientific Research at CSIC and responsible for the Open Science program. In EOSC-synergy he will play a fundamental role in all the policy activities regarding harmonization of procedures and uptaking of the Open Data paradigms at the National level in Spain. As a physicist he has published more than 400 papers in HEP and Computing. He has been director of IFCA (2003-2007) and coordinator in the area of Physics and Technology within CSIC (2008-2010). He has been proposer and coordinator of the European project Interactive European Grid (2006-2008), and has participated in several more EU projects within FP5, FP6 & FP7. He leads the Advanced Computing research line at IFCA. He is associate editor on Tools & Method section of EPJC journal (Springer), and participates in several advisory boards at national and international level. He has wide experience in collaboration with the industry, including simulation, or installation and operation of remote sensors, or advance database management. He is currently involved in several multidisciplinary initiatives involving advanced computing, and in particular LIFEWATCH, where he is member of the core-ICT board.*

Agnès Ponsati *[F], is the head of CSIC Unit of Information Resources for Research. This Unit manages the most important scientific library networks in Spain. She has been involved in promoting Open Access among CSIC scientific community through CSIC Institutional Repository Digital. CSIC as a useful initiative to re-situate scientific libraries within the scholarly workflow. In EOSC-synergy Digital CSIC will play a fundamental role in fostering the implementation of FAIR data principles on the data collections(WP3), and definition of best practices in the area of skills development for research librarians (WP6). She holds a degree in Hispanic Philology (University of Barcelona) and a Diploma in Librarianship and Documentation from the same university. Since 1994 she has been head of the CSIC Unit of Information Resources for Research till August 2017. From September 2017 she has been Director of the Division of Processes and Digital Services at Spanish National Library till November 20018. Since then she is back to the former position at CSIC. Previously she worked at the Technical*

Department of Barcelona's University Library (1987-1990) and as automation manager at Catalonian CSIC libraries branch (1990-1993). Specialist in automated library systems in distributed networks and in management and coordination of library services and collections at research hybrid-digital libraries. Responsible for the set up of CSIC Virtual Library and the management of the digital collections and all technological supporting tools. She has published papers on technical management of library systems and collections in distributed environments and management of library services in the context of scientific and research libraries. She has taken part in several national and european projects on automation of document collections, catalogue conversion systems, bibliographic protocols, meta-search engines and reuse of heritage collections by applying new technologies. She also has a wide expertise in publishing negotiations and licensing digital contents and has been member of several publishers' library advisory boards (ELSEVIER, SPRINGER, WILEY, IOP and BRILL). She is also member of the SCOAP3 Global Council representing the Spanish library consortia' and member of the LIBER executive Board.

***Dr. José María Carazo Garcia** [M], born in 1959, studied Physics at University of Granada, Spain, and obtained his PhD in Structural Biology at University Autonoma of Madrid (UAM) in 1984. After a post-doc period at the Wadsworth Center of the NYS Department of Health, (USA) with Nobel Laureate Joachim Frank, he joined the National Center of Biotechnology as head of the Biocomputing Unit in 1989, where he was appointed full professor of the Spanish National Research Council (CSIC) in 2005. He is also Honorary Professor of the UAM since 2000, first in Computer Architecture and then in Structural Biology. Professor Carazo has a sustained experience in the field of Three-dimensional Electron Microscopy under cryogenic conditions (cryo-EM), especially in the methods development area. His laboratory has opened whole new areas in the field, naming just as examples the recent successful family of Maximum Likelihood algorithms (developed in Madrid from 2007 to 2011) or the very much used EMDataBank (started from the European Union Bioimage project, that he Coordinated from 1996 to 1999). At the level of software developments, he develops and supports software packages like Xmipp and Scipion, who have been downloaded from close to 1020 individual IP addresses just in the last year, as well as web services, like 3DBionotes, with close to 3486 different users counting from March 2018. Note that the specific version of Scipion developed for Cryo-EM facilities is currently being at the heart of some key EU resources, such as Diamond/eBIC –probably the largest cryo-EM facility in the world-, in Oxford, SciLab, in Stockholm and the cryo-EM line of the European Synchrotron (ESRF). The research record of Carazo's laboratory can be inferred from the h-index, the close to 8000 citations (as Scopus on November 13th, 2018), and the granting of a ERC Syng. Grant starting 2019. As recognition to both his contributions to the EM field and also to the way of forming successful teams around him, and in the context of an international competition in Europe, his laboratory has been selected as the Instruct Center for Image Processing in Structural Biology, which is part of the Strategic European Research Infrastructure INSTRUCT, starting in 2011. This central role certainly will help in assuring the dissemination of his "fresh" developments in a global manner.*

General description of CSIC as national institution

The CSIC (Spanish National Research Council) is Spain's largest public research institution, and ranks third among Europe's largest research organization. The CSIC is attached to the Spanish Ministry of Science, Innovation and Universities through the State Secretariat for Research, Development and Innovation, and plays a key role in scientific and technological policy in Spain and worldwide. According to its Statute (Article 4), CSIC has 4 main missions:

- to foster multidisciplinary scientific and technological research
- Knowledge transfer to industry and society
- Education and training of scientific and technical staff
- Creation of Technology Based Companies

CSIC has 10.940 employees, including 3.764 researchers. CSIC has 120 Institutes spread across the country and covering different areas of Science and Technology. 67 of them are fully-owned institutes and 53 are Joint Research Units in partnership with other Spanish universities or research institutions. CSIC has also a delegation in Brussels.

CSIC supports research and training across a wide range of knowledge, from the most basic or fundamental aspects of science to the most complex technological developments; from human and social sciences to food science and technology, including biology, biomedicine, physics, chemistry and materials, natural resources and agricultural sciences. As the third largest research organization in Europe, CSIC carries out research in all fields of knowledge, throughout its 123 Institutes distributed in eight areas:

- Humanities and Social Sciences
- Biology and Biomedicine
- Natural Resources
- Agricultural Sciences
- Physical Science and Technologies
- Materials Science and Technology
- Food Science and Technology
- Chemical Science and Technology

CSIC produces 20% of the national scientific output (more than 12.000 ISI paper in 2014). CSIC remains the leading patent filer among research bodies in Spain with more than 180 patent requests published in 2014.

CSIC has a broad experience in managing large and singular infrastructures. For instance, CSIC provides services to the entire scientific community through management of the Singular Scientific and Technological Infrastructures (ICTS) such as Calar Alto Astronomical Observatory (with MaxPlanck), Doñana Biological Station, European Synchrotron Radiation Facility, Hesperides Ocean Research Vessel, Integrated Micro and Nanoelectronics Clean Room, Juan Carlos I Antarctic Base, Max Von Laue-Paul Langevin Institute and Sarmiento de Gamboa Ocean Research Vessel.

In addition, CSIC has a broad experience in conducting R&D projects funded by national and international public agencies and industry.

CSIC is a major player in the development of the European research area and therefore a significant contributor to the European integration process. Within the 7th Framework Programme CSIC has signed 726 actions (including 70 coordinated by CSIC). As to the number of projects, CSIC is listed the 1st organisation in Spain and the 4th in Europe within the research organizations, with a total FP7 contribution of over 264 million euros (E-CORDA).

As to the funding obtained by CSIC within each programme, the distribution is People 20%, Cooperation 47%, Capacities 8% and Ideas 25%. Taking into account the research areas, the most relevant ones in terms of funding have been Physical Science and Technology and Biology and Biomedicine.

In H2020 (2014-2018) CSIC has obtained 446 projects with a total EU financial contribution of 181 million euros. As E-CORDA points out CSIC is listed the 1st organisation in Spain and the 3rd participant within the research organization by number of projects.

CSIC is a major player in the ERC programme with a total of 87 projects signed as Host Institution in all areas of knowledge. CSIC is also an active member in Knowledge and Innovation Communities (KIC), such as Raw Materials and Food of the European Institute for Innovation and Technology (EIT).

4.1.2 LIP

LIP is the institution coordinating IBERGRID in Portugal. In EOSC-synergy LIP is the deputy coordinating institution (WP1) responsible for putting in place the quality management procedures. It will also support the activity of implementing software maturity, and coordinate the WP3. LIP supports the deployment of services in WP4, and will also take care of outreaching towards the Universities in Lisbon for the developments in WP6.

Description of the partner

Laboratório de Instrumentação e Física Experimental de partículas (LIP) is a Portuguese scientific organization, aiming at research in experimental particle physics and related detector technology. LIP was created in 1986 as the Portuguese Laboratory for collaboration with the European Organisation for Particle Physics (CERN). LIP develops its R&D activities mainly in the framework of experiments at CERN, but also with the European Space

Agency (ESA). LIP also uses and shares other international scientific facilities such as GSI, SNOLAB, SURF (LUX) or the Pierre Auger Observatory. The LIP scientific activity is organized in research groups focused in three main areas:

- Experimental particle physics and astroparticle physics

• Development of new instruments and methods (detectors for nuclear and particle physics, medical applications, radiation environment studies and applications for space missions)

• Advanced computing

LIP is member of the Portuguese National Distributed Computing Infrastructure (INCD) a digital research infrastructure that delivers compute and data services to the Portuguese scientific and academic communities. INCD is part of the Portuguese Science Foundation (FCT) roadmap of research infrastructures. The LIP Distributed Computing and Digital Infrastructures team is responsible for the development and coordination of the INCD infrastructure services including cloud, HPC, HTC and other higher level added value services. LIP is also member and Portuguese coordinator of IBERGRID the Iberian distributed computing infrastructure that joins the Portuguese (INCD) and Spanish (NGI-ES) infrastructures.

LIP has extensive experience in the deployment and operating of complex distributing computing infrastructures. Since 2001 continuously participates in international research projects in this domain. In the context of the Portuguese distributed computing infrastructure, LIP provides support to a wide range of scientific communities.

Key personnel

Jorge Gomes *[M], is EOSC-synergy deputy coordinator. He is a computing researcher at LIP, he was technical coordinator of the Portuguese National Grid Initiative and is currently member of the board of directors of the Portuguese National Distributed Computing Infrastructure (INCD), he is also researcher at LIP. He worked in the development of data acquisition systems for High Energy Physics experiments at CERN, and participated in pioneer projects in the domains of digital satellite data communications, IP over ATM, and advanced videoconference over IP networks. Since 2001 he participated in international projects on grid computing including CrossGrid, Enabling Grids for E-Science (EGEE) I, II and III, E-Infrastructure Shared Between Europe and Latin America (EELA), int.eu.grid, EGI-INSPIRE, EGI-ENGAGE and INDIGO-DataCloud. Since 2008 he is coordinator of the Portuguese Tier-2 computing facility in the Worldwide LHC Computing Grid (WLCG) supporting the CERN Large Hadron Collider experiments. He participated in the setup and coordination of the EUgridPMA trust fabric for research e-Infrastructures in Europe, Middle-East, and Africa. He is responsible for the Portuguese participation in the IBERGRID initiative that joins the Portuguese and Spanish distributed computing infrastructures. He coordinates the Portuguese participation in the European Grid Initiative (EGI), and is the Portuguese representative at the EGI council. He is currently participating in the European projects DEEP-HybridDataCloud and EOSC-hub.*

Mario David *[M], is a research associate at LIP. He holds a PhD in Experimental Particle Physics from the University of Lisbon. He held a research associate position at Institut de Physique du Globe de Paris (IPGP/CNRS) as Scientific Software Developer for the VERCE project, in particular in the data intensive use cases for seismology. Previously he held positions as Post-Doc and research associate at LIP participating in FP6 and FP7 projects such as DataGRID, CROSSRID, EGEE I-III, Int.EU.GRID, and*

EGI-Inspire. More recently he participated in the H2020 projects EGI-ENGAGE and INDIGO-DataCloud where he coordinated the pilot infrastructure services. He is also a collaborator of the Worldwide LHC Computing Grid and Ibergrid infrastructures. Throughout those projects, he was actively involved in the Validation, Quality Assurance and testing of middleware, in regional and global operations and user support. He has coordinator of the Service Deployment and Validation in EGI. Currently he is participating in the DEEP-HybridDataCloud project where he coordinates the testbed and integration with EOSC services. He has over 20 publications in peer review journals and over 35 communications in international conferences.

4.1.3 KIT

KIT plays the coordinating role in the Work Package of Expanding Capabilities at the Infrastructure level (WP2), where it will also drive the integration of the Helmholtz Data Federation data services into EOSC.

The researchers from the Clima center of KIT participate in the validation of services for the environmental communities.

Description of the partner

Karlsruhe Institute of Technology (KIT) is the combination of a science and research institution with a University in Germany. Funded by the Federal Republic of Germany and the state of Baden-Württemberg KIT is leading in innovation, research and teaching with a staff of 10000 employees and 25000 students. KIT is a member of the Helmholtz Association of German Research Centres (HGF).

The Steinbuch Centre for Computing (SCC) is the computing centre of KIT and supports the IT-demands of the multi-disciplinary research at KIT and within HGF. It has solid experience in distributed computing infrastructures as well as their federated identity management, authentication, authorisation and secure operation of infrastructures.

SCC successfully participated in many European projects and is currently contributing to AARC2, ELI-Trans, EOSC and Deep Hybrid Data Cloud, Human Brain Project.

SCC also developed and runs the largest German university cloud storage for sync-and-share, providing access to more than 350,000 students and 100,000 researchers. SCC provides 28 PB disk and over 30 PB tape storage to a diverse range of scientific experiments from more than 130 institutes on site and internationally.

The Institute of Meteorology and Climate Research (IMK) – Atmospheric Trace Gases and Remote Sensing (ASF) at KIT investigates dynamical, microphysical and chemical processes in the Earth's atmosphere, with the goal to understand, quantify and predict natural variability and long-term changes. With accurate measurements of atmospheric trace gases from various observational platforms (ground-

based stations, air craft, balloons, satellites) and comprehensive modelling of the atmosphere and its composition we unravel complex links and feedbacks in the climate system.

Key personnel

Dr. Marcus Hardt *[M], is the coordinator of WP2 in EOSC-synergy. He earned his PhD with precise and distributed simulations of Ultrasound waves in human tissue within the KIT project "Ultrasound Computer Tomography", where he contributed simulations for ultrasound waves in medical devices. He worked as founding member of WebSmart Technology GmbH and as IT freelancer between 1999 and 2002. Since 2002 he is working as a scientist at the Research Group Cloud Computing at SCC. As a member of the CrossGrid and int.eu.grid integration teams he managed the fully automated software build, deployment and configuration on the Europe-wide DCI. He was furthermore responsible for the KIT activities in the EU-Project EUFORIA. In his present position Marcus Hardt is the technical coordinator of the Large Scale Data Management and Analysis project LSDMA, a Helmholtz Portfolio extension, targeted at developing uniform interfaces for Germany's data intensive research projects.*

Prof. Dr. Achim Streit *[M], is the director of the Steinbuch Centre for Computing (SCC) and at the same time professor for computer science at KIT since mid-2010. In this capacity he is responsible for the HPC and Big Data activities at SCC – both hardware and systems as well as R&D activities. He is the lead PI of the German Helmholtz Association' project Large Scale Data Management and Analysis (LSDMA), which is about fostering data-intensive science in Germany through Data Life Cycle Labs and generic methods research and which involves in total 11 partners and an annual budget of 3 M€ for 5 years. Prior to KIT, he was leading the Grid and Data activities at the Jülich Supercomputing Centre (JSC) of Forschungszentrum Jülich, Germany. In this role he was very active in several EU infrastructure projects such as DEISA, PRACE, EGEE-2, OMII-Europe, EGI-InSPIRE and ETICS-2 on the managerial level and initiated the EMI and EUDAT projects. He holds a PhD in computer science specialising on scheduling for cluster and distributed systems. He is co-author of more than 70 peer-reviewed publications in peer-reviewed conferences, workshops and journals.*

Peter Braesicke *[M], is professor for theoretical atmospheric physics and section head for modelling at KIT's IMK-ASF. After finishing his PhD in Berlin on equatorial waves and transport in climate models he spent 14 years in various roles at the University of Cambridge (UK). There, he contributed to the development of the UK's composition-climate model and lead aspects of the model's development. As a senior) research associate he (co-)supervised a number of students and PhD projects. He is working on variability and trends in the composition-climate system, including the interplay of stratospheric ozone and climate change. He is currently lead author of the Scientific Assessment of Ozone Depletion (Chapter on Global Ozone), chair of the European Climate Research Alliance (ECRA) and scientific coordinator of the (regional) Helmholtz Climate Initiative (REKLIM).*

Publications

Schröter, J., Rieger, D., Stassen, C., Vogel, H., Weimer, M., Werchner, S., Förstner, J., Prill, F., Reinert, D., Zängl, G., Giorgetta, M., Ruhnke, R., Vogel, B., and Braesicke, P.: ICON-ART 2.1: a flexible tracer

framework and its application for composition studies in numerical weather forecasting and climate simulations, Geosci. Model Dev., 11, 4043-4068, https://doi.org/10.5194/gmd-11-4043-2018, 2018.

Langematz, U., Schmidt, F., Kunze, M., Bodeker, G. E., and Braesicke, P.: Antarctic ozone depletion between 1960 and 1980 in observations and chemistry–climate model simulations, Atmos. Chem. Phys., 16, 15619-15627, https://doi.org/10.5194/acp-16-15619-2016, 2016.

Braesicke, P., Keeble, J., Yang, X., Stiller, G., Kellmann, S., Abraham, N. L., Archibald, A., Telford, P., and Pyle, J. A.: Circulation anomalies in the Southern Hemisphere and ozone changes, Atmos. Chem. Phys., 13, 10677-10688, https://doi.org/10.5194/acp-13-10677-2013, 2013.

4.1.4 PSNC

The Poznan Supercomputing and Networking Center will be involved in expanding the capacity of EOSC with Polish resources (WP2), providing input to National policies (WP5) and coordinating the deployment of the training platform (WP6).

Description of the partner

Poznan Supercomputing and Networking Center, affiliated to the Institute of Bioorganic Chemistry of the Polish Academy of Sciences, was established in 1993 by the State Committee for Scientific Research. PSNC is responsible for the management of the countrywide Polish National Research and Educational Network (NREN) PIONIER, which is built on PSNC's own fibres and is connected to the GEANT2 network. From the very beginning PSNC has been a key player in the development of the polish e-Infrastructure, currently serving as an HPC service provider, network services provider, and operator of the Poznan Metropolitan Area Network (POZMAN), the Polish national scientific TV platform (PlatonTV), and the national digital library federation.

PSNC disposes of extensive ICT research infrastructure, including 1.4+ PFLOPS of computing power and 47 PB storage and data management infrastructure, providing direct support for the scientific communities in Poland as well as in Europe. PSNC has built a co-working space and is coordinating living lab communities. Via its co-working space, PSNC supports 15+ user communities, gathered around ICT technologies and application areas. PSNC's Poznan Living Lab, as effective member of ENoLL, runs city challenge workshops, hackathons and design thinking interdisciplinary courses. PSNC coordinates HPC4Poland Digital Innovation Hub, focused on HPC-High Performance, as well as the leading Polish ICT research driven Wielkopolska ICT Cluster with 100+ members.

As one of the most recognised applied research centers in Poland, PSNC mission has been the integration and implementation of scientific research results in the form of fully-fledged services for scientists, the public administration, and different sectors including agri-food and others, as well as the deployment and maintenance of such services in the polish e-infrastructure. In particular, regarding the agri-food sector, PSNC has focused part of its efforts during the last years towards the development of a smart agriculture

infrastructure in Wielkopolska region. PSNC employs about 300 people in four divisions and has coordinated 20+ international EU projects and participated in 130+ European and 70+ national projects.

Key personnel

Dr. Marcin Płóciennik *[M], is the coordinator of WP6 in EOSC-synergy. He is head of the IoT Systems Departament in PSNC. Since 1998 he was working in a number of projects focused mainly on researches concerning distributed computing, scientific workflows, sensor networks and remote instrumentation and Internet of Things, participating in project of 5,6 and 7 FP, H2020: CrossGrid, BalticGrid I/II,, EGEE, Euforia, int.eu.grid, DORII(deputy project coordinator), EFDA ITM ISIP, EUROFusion WPISA CPT(Core programming team deputy leader), EGI_Inspire, national projects: Future ICT, PLGrid/PLGrid+/PLGrid NG). symbIoTe, INDIGO-DataCloud (WP Leader). Currently he is involved as well in the EOSC-hub project. He was leading also international OGF research and standardization bodies related to remote instrumentation and sensors (RISGE-WG, ARIWG). He is the author and co-author of 40+ conference papers and articles in international journals, member of programme committees of international ICT conferences.*

Dr. Norbert Meyer *[M], is currently the head of the Supercomputing Department of PSNC. His research interests concern resource management in distributed environments, accounting, data management, and network security. Norbert Meyer is the member of the e-IRG (e-Infrastructure Reflection Group), chair of the data management task force, co-author of several white papers, member of STRATOS group. He coordinated EU projects: DORII and RINGRID., has been participating in many projects and project proposals of 5, 6 and 7 FP: CrossGrid, EGEE, EGEE2, CoreGRID, BalticGRID, Expres, RinGRID, Phosphorus, int.eu.grid, DORII, Expres, NEXPRES, PRACE, PRACE-1IP, PRACE-2IP, EUDAT, IGE (total 35+ international proposals, 2000-2014).*

Radosław Januszewski *[M], received the M.Sc degree in Computer Science from the Poznań University of Technology in 2001 and is writing his Ph. D. disseration at the same university. Currently he is the employee of the Poznan Supercomputing and Networking Centre (http://www.man.poznan.pl), the Supercomputing Department where he works the leader of HPC and Data Centre Technologies department responsible of on HPC solutions and Data Centre design and operations. He took part in the preparations of some project proposals (e.g. National Data Store, accepted by the Polish Government) 2005-2007 he worked on the integration of low-level checkpointing services with the Grid environment for the fault-tolerance purposes. He is co-author of several papers related to fault tolerance in the Grids. In 2008-2015 he was involved in the PRACE project focusing on HPC energy efficient systems and warm-water, directly liquid cooled HPC solutions. Some of the results were published in paper "Evaluation of the impact of direct warm-water cooling of the HPC servers on the data centre ecosystem" in the proceedings of the ISC 2014. In the PL- GRID project he is involved in deploying Cloud technologies (virtualization of the computing resources) for increasing reliability of the HPC computations. Since 2015 he is the leader of the HPC team in PSNC*

Publications

- Marcin Plociennik, Sandro Fiore, Giacinto Donvito, Michal Owsiak, Marco Fargetta, Roberto Barbera, Riccardo Bruno, Emidio Giorgio, Dean N. Williams, Giovanni Aloisio, Two-level Dynamic Workflow Orchestration in the INDIGO DataCloud for Large-scale, Climate Change Data Analytics Experiments, Procedia Computer Science, Volume 80, 2016, Pages 722-733, ISSN 1877-0509
- Sandro Fiore, Marcin Plociennik, et al. Distributed and cloud-based multi-model analytics experiments on large volumes of climate change data in the Earth System Grid Federation eco-system. Proceedings of the IEEE Big Data 2016
- Michal Owsiak, Marcin Plociennik, Bartek Palak, Tomasz Zok, Cedric Reux, Luc Di Gallo, Denis Kalupin, Thomas Johnson, Mireille Schneider, Running simultaneous Kepler sessions for the parallelization of parametric scans and optimization studies applied to complex workflows, Journal of Computational Science, Available online 19 *December 2016, ISSN 1877-7503,* *http://dx.doi.org/10.1016/j.jocs.2016.12.005*
- Plociennik M., Zok T., Altintas I., Wang J., Crawl D., Abramson D., Imbeaux F., Guillerminet B., Lopez-Caniego M., Campos Plasencia I., Pych W., Ciecielag P., Palak B., Owsiak M., Frauel Y., Approaches to Distributed Execution of Scientific Workflows in Kepler, Fundamenta Informaticae, Volume 128, Issue 3, p. 281- 302
- Plociennik M., Zok T., Owsiak M., Palak B., Guillerminet B., Frauel Y., Imbeaux F., Scott B.: High level tools for fusion simulations workflows in distributed computing environment. In Proceedings of HPCS 2012: 602-608

4.1.5 CESNET

CESNET will be involved in expanding the capacity of EOSC with the resources of the Czech Republic (WP2), providing input to National policies and coordinating the national policy work package (WP5) and supporting the inclusion of thematic services of national interest into EOSC (WP4). Furthermore it will take care of outreach with the local Universities in what regarding WP6 training developments.

Description of the partner

CESNET is an association of universities of the Czech Republic and the Czech Academy of Sciences. It operates and develops the national e-infrastructure for science, research and education, which encompasses a computer network, computational grids, data storage and collaborative environment. It offers a rich set of services to connected organizations. CESNET also serves as a body coordinating the national grid and cloud activities through its MetaCentrum department; as Czech NGI it represents Czech Republic in the EGI Council. CESNET operates the national computer network backbone and also operates a substantial computing and storage resources, available to all bona fide researchers in the Czech

Republic and abroad (esp. for those collaborating with Czech researchers). CESNET runs national academic CSIRT team and involves in the cyber-forensic activities.

CESNET is involved in a large number of national and international projects, especially those related to the building and development of large scale e-infrastructures, covering both network (though GEANT series of projects) and grid and cloud environment (through the EGI related projects). Its representatives play an important role at the European level, serving as WP leaders and PMB members in several projects.
CESNET has a long history of R&D in the area of grid and recently cloud middleware systems, together with CERIT-SC (Masaryk University) the Authentication and Authorization systems, namely Perun used in EGI FedCLoud and also in ELIXIR and recently also BBMRI AAIs. Additional development covers also Software Defined Networks (SDN), bridging computer network and cloud expertise.

Relevant Projects

1. EGI ENGAGE – Engaging the EGI Community towards an Open Science Commons
1. DEEP-Hybrid DataCloud – Designing and Enabling E-Infrastructures for intensive Processing ina Hybrid DataCloud
2. AARC2 – Authentication and Authorization for Research and Collaboration
3. GN4-2 – SGA2-GEANT2020 – Framework Partnership Agreement
4. ELIXIR EXCELERATE – Fast-track ELIXIR implementation and drive early user exploitation across the life-sciences.

Relevant Publications

1. M. Kimle, L. Kosaristan, B. Parak, Z. Sustr: Managing Virtual Appliance Lifecycle in IaaS and PaaS Clouds. In International Symposium on Girds and Clouds, Taipei, Taiwan, 2016, PoS (ISGC 2016) 034, https://pos.sissa.it/archive/conferences/270/034/ISGC%202016_034.pdf
2. D. Kouril, M. Poul, M. Prochazka: Using PKI to Provide Credential Delegation in non Web-based Federations. In Proceedings of the International Conference on Information Science and Applications, Lecture Notes in Electrical Engineering. Pattaya, Thailand: Springer Berlin Heidelberg, 2015. s. 525-532. doi:10.1007/978-3-662-46578-3_62.
3. S. Toth. M. Ruda: Distributed job scheduling in MetaCentrum. Journal of Physics: Conference Series, Institute of Physics, 2015, Vol. 608 (1), pp. 012025-12029. doi:10.1088/1742-6596/608/1/012025.
4. J. Filipovic, M. Madzin, J. Fousek, L. Matyska. Optimizing CUDA code by kernel fusion: application on BLAS. The Journal of Supercomputing, Springer US, 2015, Vol. 71 (10), pp. 3934-3957. doi:10.1007/s11227-015-1483-z.

Key personnel

Luděk Matyska *[M], is a full professor at Masaryk University (MU) as well as a senior researcher at CESNET. Since 2013, he serves as a director of the Institute of Computer Science at Masaryk University and he is also the director of CERIT Scientific Cloud (CERIT-SC) at MU, a national e-infrastructure dedicated to promotion and building of national cloud and integrated grid-cloud infrastructures; complementing the CESNET e-infrastructure. He works for CESNET since 1998, serving as principal co-investigator of its research programs. He is and has been involved in many national (e.g. ELIXIR CZ, CERIT-SC, MetaCentrum) and international projects (e.g. EOSC-Hub, ELIXIR EXCELERATE, DEEP, EGI ENGAGE, INDIGO-DataCloud, SDI4Apps) as a head of the CESNET or MU team. He was a chair and currently serves as a member of EGI Executive Board, is member of EOSC Hub PMB, co-chair of ELIXIR Compute Platform Executive Committee and co-chair of ELIXIR EXCELERATE WP4. His research interest lies in security in large scale distributed systems, e-infrastructure (network, computing and data) architecture and policy and generally in cloud and grid systems. He authored or co-authored more than 100 papers and conference contributions. He will serve as WP5 leader in the project and will be involved in other WP (esp. WP4 and WP6).*

Miroslav Ruda *[M], works for CESNET, where he is a head of the MetaCentrum department and the official Czech representative in the EGI Council; he is responsible for the compute and cloud part of CESNET national e-infrastructure. Since late nineties he has been involved in a long series of grid and recently cloud infrastructure related projects both at the national (esp. national research e-infrastructure projects) and international (recently EOSC-Hub, ELIXIR EXCELERATE and DEEP projects), levels. He was EGI ENGAGE PMB member. He is one of the leading persons in the EGI FedCloud and generally in the EGI cloud activities, with a special focus on management and operation, including resource allocation, scheduling and planning. In EOSC SYNERGY he will be primary responsible for WP2 and WP3 activities*

Aleš Křenek *[M], works as senior researcher at CESNET and he is also leading the capacity building project dedicated to the capacity building of national e-infrastructure CERIT-SC at Masaryk University. In the past, he has been involved in the development of grid middleware in the EDG, EGEE, and EMI projects, being responsible for one of development teams. He was also leader of the international software support team of EGI and recently he was leader of an infrastructure workpackage in the H2020 West-Life project; he is noe in a similar position in the EU funded EuroPDX project. In the EOSC SYNERGY he will be responsible for overseeing and thematic services on CryoEM and UMSA in WP4.*

4.1.6 IISAS

The Institute of Informatics of the Slovak Academy of Sciences will be involved in expanding the capacity of EOSC with Slovak resources (WP2), where they also coordinate the task fo integration of national repositories. They will also be providing input to National policies (WP5). Users from Slovakia will participate in the validation of services for the environmental community.

Description of the partner

Institute of Informatics of the Slovak Academy of Sciences (UISAV) is one of more than 50 scientific and research institutes of the Slovak Academy of Sciences in Bratislava, Slovakia. UISAV, established in 1956, is a leading research Institute in informatics and information technology in Slovakia. The Institute employs around 70 researchers; almost half of them are women. The scope of researched development activities includes informatics, information technology, robotics, control theory and artificial intelligence. The Institute is very active in EU-wide and national research projects, mainly in the areas of distributed, grid and cloud computing, knowledge management, information security and data processing. It is structured into eight departments covering a wide range of technologies: Parallel and Distributed Information Processing, Design and Diagnostics of Digital Systems, Numerical Methods and Algorithms, Speech Analysis and Synthesis, Discrete Process Modelling and Control, Sensor Systems, PCB Design and Production and Electron Beam Lithography.

The principal partner is the Department of Parallel and Distributed Information Processing, with an excellent track record in solving EU and national research projects and in cooperating with IT industry. The department started to be engaged in European research in the 4th Framework Programme, and since it has participated in over 60 research projects, steadily moving from distributed and HP computing, through Grid to Cloud Computing as well as Big Data and Artificial Intelligence research.

UISAV has strong experience in both research and commercial applications in the Information retrieval and processing field. UISAV has cooperated with different types of business sectors e.g.: advertising (Magnetic Media - US based company), social media (Connectik – Switzerland based company) and media monitoring (SCS - Switzerland based company) as well as security field of mobile device and Big Data transfer and preparation for automated data privacy tools (IBM Slovakia). In the research area of information polling and processing UISAV participated in SECRICOM and REDIRNET project.

UISAV is also publishing an academic journal, Computing and Informatics (CAI; http://www.cai.sk). The journal is published under a delayed open access model (all publication are available in golden open access status six months after the publication).

Relevant projects or activities

1. PROCESS - PROviding Computing solutions for ExaScaleChallengeS.H2020-777533, 2017-2020, http://www.process-project.eu/. The project will pave the way towards exascale data services that will accelerate innovation and maximise the benefits of these emerging data solutions. To providing the service prototypes for very large data, the project addresses the work programme goals by using the tools and services with heterogeneous use cases, including open data for global disaster risk reduction.
2. DEEP-HybridDataCloud - Designing and Enabling E-infrastructures for intensive Processing in a Hybrid DataCloud.H2020, 2017-2020, http://deep-hybrid-datacloud.eu/. The key concept proposed in the DEEP Hybrid DataCloud project is the need to support intensive computing techniques that require specialized HPC hardware, like GPUs or low latency interconnects, to explore very large datasets. A Hybrid Cloud approach enables the access to such resources that are not easily reachable by the researchers at the scale needed in the current EU e-infrastructure.

3. EOSC-hub - Integrating and managing services for the European Open Science Cloud.H2020-777536, 2018-2020, http://eosc-hub.eu/. The EOSC-hub project creates the integration and management system of the future European Open Science Cloud that delivers a catalogue of services, software and data from the EGI Federation, EUDAT, CDI, INDIGO-DataCloud and major research e-infrastructures. This integration and management system (the Hub) builds on mature processes, policies and tools from the leading European federated e-Infrastructures to cover the whole life-cycle of services, from planning to delivery.
4. EGI-Engage - Engaging the Research Community towards an Open Science Commons. H2020-654142, 2015-2017, http://www.egi.eu/about/egi-engage/. EGI-Engage accelerates advancements within the European Grid Infrastructure (EGI) in strategy, policy, business and technical innovation, user engagement towards researchers within the long-tail of science, domain-specific research communities, Research Infrastructures (RIs) within the ESFRI roadmap, as well as SMEs and industry at large.
5. REDIRNET - Emergency Responder Data Interoperability Network.EU FP7-607768,2014-2016, http://www.redirnet.eu/. Provided a decentralized framework for interoperability for first responders' systems based on a public metadata gateway controlled by the agencies themselves via a REDIRNET socio-professional web. UISAV has developed a Poll Management and Aggregation Service (PMAS), which is able to request and reserve resources for crisis mitigation in a very secure, reliable and fast manner. UISAV has contributed also to the architecture design, requirements elicitations from Slovak first responders as well as from the Ministry of interior of the SR, development of ontologies for crisis response, the Secure Metadata Gateway and the Socio-professional Web.
6. VENIS - Virtual Enterprises by Networked Interoperability Services. FP7-284984, 2011-2015, http://www.venis-project.eu/. VENIS is a FInES Cluster FP7-FoF (Factory of the Future) Project. The VENIS project is aimed at providing a new level of interoperability between Large and Small Enterprises, according to "Virtual Enterprise" paradigm.
7. EUSAS - European Urban Simulation for Asymmetric Scenarios.EDA A-0938-RT-GC, 2010-2012, https://www.eda.europa.eu/. European Defense Agency project involved generic and modular agent architecture representing simulated people under various behavior-affecting perceptions as well as behavior moderators with psychological backgrounds.
8. ADMIRE - Advanced Data Mining and Integration Research for Europe.FP7-215024, 2008-2011, http://www.admire-project.eu/. The project is pioneering architectures and models that deliver a coherent, extensible and flexible framework to make the best use of a wide range of distributed data resources. The new architecture improves the accessibility and exploitation of data for domain experts and provides a workbench for data mining and integration experts, who will, in turn, improve their productivity and the rate of deployment of new methods and applications.
9. Secricom - Seamless Communication for Crisis Management.Best Project Award in the SECURITY Call, FP7-218123 IP, 2008-2011, http://www.secricom.eu/. The aim of the project was to create a seamless communication infrastructure with advanced intelligent functionality for crisis management with participation of multiple agencies and stakeholders. UISAV has developed a novel Secure Agent Infrastructure (SAI) for secure collection and communication of data. UISAV has also contributed to System analysis and design and Integration of research results.
10. COMMIUS - Community-based Interoperability Utility for SMEs. FP7-213876, 2008-2011, http://www.commius.eu/.Commius delivers an adaptable and customizable software prototype, providing SMEs with "zero-cost entry" into interoperability and allowing them to reuse existing and familiar applications for electronic communication. This is made possible by a number of innovative scientific, technical and business advances over the existing state-of-art.
11. GAMMA - Global ATM security management. FP7-SEC-2012-1, 2013 – 2017, http://www.gamma-project.eu/. In this project UI SAV is responsible for using speech analysis in

voice communication as an add-on feature to improve air traffic management security. The same approach can be applied to airborne, maritime and land communication systems.

12. Relationship between trust and entrainment in speech. Air Force Office of Scientific Research of the USA, FA9550-15-1-0055, 2015 – 2017, https://www.wpafb.af.mil/afrl/afosr/. This project investigates various aspects of human-machine communication. This topic is very important for the optimization of human-machine interfaces in communication infrastructures.

Infrastructure or technical equipment

UISAV HPC cluster: IBM system: 52x IBM dx360 M3, 8x IBM dx360 M4 2x NVIDIA Tesla K20, 2x IBM dx360 M3 NVIDIA Tesla M2070, 2x x3650 M3 managing servers, 4x x3650 M3 data managing servers, x3550 M4 server, InfiniBand 2x 40 Gbps, 2x DS3512 with 72TB disk. The cluster is already a part of international test-beds of European infrastructure in international IST, FP and H2020 projects

UISAV Hadoop cluster: Scalable dedicated computing infrastructure for data processing and analytics with 1x server and 14x client. Node specification: 2x Intel® Xeon® Processor E5-2620 (15M Cache, 2.00 GHz, 7.20 GT/s Intel® QPI, 6x cores, 12x threads) + HyperX threading (24 simultaneous tasks per client), 32GB RAM (2 nodes 48GB), 1TB HDD (2 nodes 500GB). Total storage capacity of the cluster is 13TB.

Key personnel

Assoc. Prof. Dr. Ladislav Hluchý *[M], is the head of the Department of Parallel and Distributed Information Processing. He has been the director of the Institute of Informatics, Slovak Academy of Sciences (UISAV) for more than 20 years. He received his MSc and PhD degrees both in Computer Science. He has rich experience as a R&D project manager, work-package leader and coordinator in projects of the 4th, 5th, 6th, 7th Framework Programmes and H2020 program; e.g. PROCESS (PROviding Computing solutions for ExaScaleChallengeS), DEEP - Hybrid Data Cloud (Designing and Enabling E-infrastructures for intensive Processing in a Hybrid DataCloud), EOSC-hub (Integrating and managing services for the European Open Science Cloud), EUSAS (European Urban Simulation for Asymmetric Scenarios EDA A-0938-RT-GC), VENIS (Virtual Enterprises by Networked Interoperability Services FP7-284984), ADMIRE (Advanced Data Mining and Integration Research for Europe FP7-215024), REDIRNET (Emergency Responder Data Interoperability Network FP7-607768), Pellucid (A Platform for Organisationally Mobile Public Employees IST-2001-34519), ANFAS (Data Fusion for Flood Analysis and Decision Support IST-1999-11676), as well as multiple Slovak national R&D projects. He is a member of IEEE, e-IRG, EGI Council, the Editor-in-chief of the CC journal Computing and Informatics. He is also (co-)author of scientific books and numerous scientific papers (more than 570), contributions and invited lectures at international scientific conferences and workshops. He is a supervisor and consultant for PhD study at the Slovak University of Technology in Bratislava.*

Dr. Viet Tran *[M], is an experienced researcher in the field of computer sciences and informatics. His primary research fields are distributed systems, grid computing and cloud computing. He actively participates on writing project proposal and solving European project in 6th, 7th framework programme and H2020. He participated on R&D*

projects EGI-Engage (H2020-654142), EGI-InSPIRE (FP7-261323), ADMIRE (FP7-215024), MEDIGRID (FP6-004044), EGEE (INFSO-RI-508833), K-Wf Grid (FP-511385), ANFAS(IST-1999-11676), CrossGrid (IST-2001-32243), DEGREE (FP6-034619), int.eu.grid (RI-031857) and several 4FP projects. Currently he is working as workpackage leader in DEEP-HybridDataCloud (H2020-777435) and PROCESS (H2020-777533) and also as activity leader in EOSC-Hub (H2020-777536). He was a committee member in many significant conferences and workshops. He is the author and co-author of more than 100 scientific papers.

***Dr. Tomáš Gibala** [M], is an experienced water modeller and researcher in the field of water management. His primary research fields are mathematical modelling and implementation of DSS based on AI into water management. He actively participates on writing project proposals and implementing research project on local basis. He participated on R&D projects as the key research manager or key expert in "Preparation of Regional Master Plans for Water and Wastewater in Western Region of Republic of Bulgaria", MIDP-MP-QCBS3, Loan No: 7834-BG; "Risk research center for water supply in big cities, Operational programme Research and Innovation", ITMS project code: 2624011082; "Implementation of water loss monitoring system, maintenance and servicing of installed equipment for SPaP, a.s."; "Implementation of on-line water supply system model for early warning in BVS, a.s."; "VESNA Project - Flood in Town- Decision Support Tool", APVV-0234-07; "Flood hazard and flood risk maps for Slovakia, Slovak water management enterprise" and other.*

4.1.7 KNAW-DANS

The Data Archiving and Networked Services (DANS) participate in the project in the framework of providing guidelines to the practical implementation of FAIR principles (WP3) where DANS also coordinates the related task. DANS will provide feedback at the national policy level (WP5) and outreach towards the Dutch Universities in WP6, as well as participating in the development of best practices and guidelines for training material.

Description of the partner

Data Archiving and Networked Services (DANS), an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW), encourages researchers to make their digital research data Findable, Accessible, Interoperable and Reusable (FAIR). Relying on a team of more than 50 staff members, KNAW-DANS provides expert advice, training and certified services. Managing, storing and finding scientific data and information is facilitated by the three technical core services of KNAW-DANS: the DataverseNL network for data management in ongoing research, EASY for long term sustainable data archiving and NARCIS, the national portal for scholarly information, providing access to scientific datasets, (open access) publications, projects and other types of scholarly information from the Dutch research community. KNAW-DANS furthermore provides training and advice on data management issues and carries out research on sustained access to digital information. KNAW-DANS is supported by both the Royal

Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO), the two major research funding organisations of the Netherlands. As part of its mission, KNAW-DANS supports the Open Access principle, while being aware of the fact that not all data can always be freely available without constraints.

The organisation and development of elaborate training sessions on the different aspects of the Data Life Cycle is one of the primary services of DANS. The focus of most of the training is on Research Data Management (RDM) but also related subjects, such as good data stewardship or the certification and accreditation of (long term) repositories, are captured in conventional and/or online training sessions. The course "Essentials 4 Data Support" (http://datasupport.researchdata.nl/en/) is one of the successful products of the DANS training team and well known and received in the Netherlands and abroad. In 2016 the course was nominated for the "Digital Preservation Award for Teaching and Communications" by the UK's Digital Preservation Coalition.

All activities of KNAW-DANS are fed by the in-house expertise on data management and data archiving. The KNAW-DANS Electronic Archive SYstem (EASY) is a trustworthy long term repository holding DSA, WDS and CoreTrustSeal certifications and is the first to have been awarded the Nestor Seal for Trustworthy Digital Archives. KNAW-DANS is one of the founders of the Data Seal of Approval (DSA), is active in the Word Data System (ICSU/WDS) certification and is still heavily involved in their successor the CoreTrustSeal. As host and developer of the National Academic Research and Collaborations Information System (NARCIS), KNAW-DANS is an expert in the field of metadata and metadata standards and the connection and interoperability of different scientific data sources. KNAW-DANS is a member of DataCite and involved in several initiatives and projects around Persistent Identifiers, the crucial identification elements in the ever increasing scientific data landscape. Science driven technical developments within KNAW-DANS focus on mechanisms for automated data ingestion, automatic enrichment of data sets with Linked Open Data (LOD) and the storage of rich and more complex data packages. Next to the development of improved archiving services, KNAW-DANS is optimising its repository functionalities to improve the findability, interoperability and usability of the stored data, turning the repository services into a fundament for new research and the proper replication of completed and archived research.

Driven by data, KNAW-DANS ensures the further improvement of access to digital research data with its services and participation in (inter)national projects and networks. KNAW-DANS is involved in various European data infrastructures and projects, e.g. DARIAH, CESSDA, EHRI, EUDAT, OpenAIRE, EOSCpilot, EOSC-hub, K-PLEX, FREYA, and RDA 4.0, contributing with services and expertise in data curation, certification, Research Data Management (RDM) support and training. Based on this vast amount of project experience, KNAW-DANS is now also leading community engagement activities in several projects. The KNAW-DANS expertise crosses disciplinary boundaries although the main focus is on humanities, social sciences and life sciences. Novel developments within KNAW-DANS focus on domain specific data services and Knowledge Organisation Systems and the operationalisation of the FAIR-data principles.

Publications and services

- Publication: Wilkinson, M.D., Dillo I.G. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3. https://dx.doi.org/10.1038/sdata.2016.18
- Publication: Science Europe (2018). Science Europe Guidance Document Presenting a Framework for Discipline-specific Research Data Management. Edited by Peter K. Doorn, Chair of the Science Europe Working Group on Research Data. D/2018/13.324/1
- Publication: Doorn, P.K., & Aerts, P. (2016). A Conceptual Approach To Data Stewardship and Software Sustainability: Scientists in charge, with a little help from their friends. DANS Working Paper. urn:nbn:nl:ui:17-59c24848-9cf7-437c-b2d5-e943e9e4a35e
- Service: KNAW-DANS is involved in training courses on data management and preservation such as the RDNL course "Essentials 4 Data Support" (http://datasupport.researchdata.nl/) and APARSEN course "Sustainability: securing the value of digital data assets" (http://www.alliancepermanentaccess.org).
- Service: KNAW-DANS is the National Open Access Desk (NOAD) within the OpenAIRE network, providing training, support and expertise on Open Science data, publications and expertise (https://www.openaire.eu/contact-noads).

Relevant previous projects or activities

FAIRsFAIR (2019-2021): KNAW-DANS will coordinate the FAIRsFAIR project which addresses the development and concrete realisation of an overall knowledge infrastructure on academic quality data management, procedures, standards, metrics and related matters, based on the FAIR principles. The task of FAIRsFAIR is to deliver essential Rules of Participation (RoP) and regulatory compliance for participation in the European Open Science Cloud (EOSC). The emerging EOSC governance structure will use these RoPs to establish whether components of the infrastructure function in a FAIR manner.

FAIR Software Route (2018-2019): KNAW-DANS and The Netherlands eScience Center (NLeSC) have jointly launched the initiative to develop a FAIR Software Route. This initiative should support researchers and scientific developers to find the most suitable platform for the FAIR deposition of their research software and code. The FAIR Software Route shall be embedded in the Dutch Research Software Directory: https://www.research-software.nl

The FAIR Software Route explicitly serves as a trial for further developments at the European level and covers important more generic subjects such as metadata standards for software, landscaping software preservation initiatives and domain specific software developments and facilities.

Development of the European Open Science Cloud (EOSC): KNAW-DANS is participating in the projects EOSCpilot and EOSC-hub.

EOSCpilot (2017-2018), is laying the fundaments for the development of EOSC, building on and leverage already available resources and capabilities from research infrastructure and e-infrastructure organisations to maximise their use across the research community. KNAW-DANS is involved in work packages on Policy, Skills and Engagement.

EOSC-hub (2018-2020), focuses on the construction of the training and integrated service catalogues that will underpin the EOSC infrastructure. Within EOSC-hub, KNAW-DANS is involved in the work packages concerning training activities as well as the establishment of long term data preservation services as part of the EOSC service catalogue. http://eoscpilot.eu/ & http://eosc-hub.eu

OpenAIRE (2010-2020): the "Open Access Infrastructure for Research in Europe" (OpenAIRE) is the open science backbone of Europe, promoting open scholarship and improving the discoverability and reusability of research output. OpenAIRE was built in steps (OpenAIREplus, OpenAIRE2020, OpenAIRE-Advance) and KNAW-DANS has been participating from the early days on. Within the currently running OpenAIRE Advance project, well embedded in the EOSC developments, KNAW-DANS is responsible for the role of National Open Access Desk (NOAD), as it was in previous OpenAIRE projects. Apart from acting as a NOAD, KNAW-DANS is also responsible for training on Research Data Management (RDM) and for trials around the use of new metadata standards in the OpenAIRE distributed open science infrastructure.

Within the context of OpenAIRE, KNAW-DANS has developed a proto-type of a tool to assess the FAIRness of data and / or data repositories (http://blog.ukdataservice.ac.uk/fair-data-assessment-tool/), and actively promoted this approach through webinars (https://www.eudat.eu/events/webinar/fair-data-in-trustworthy-data-repositories-webinar). https://www.openaire.eu

CoreTrustSeal (2009 - present): KNAW-DANS was one of the founders of the Data Seal of Approval (DSA) international board. DSA aligned itself with ICSU-WDS through the RDA Repository Audit and Certification DSA–WDS Partnership WG in 2016. This cooperation lead to the the launch of a new certification organization in 2017: CoreTrustSeal. KNAW-DANS has had a presence in the boards of DSA and CoreTrustSeal throughout its lifespan. https://www.coretrustseal.org

Significant infrastructure and/or any major items of technical equipment, relevant to the proposed work

KNAW-DANS is maintaining 3 technical core-services relevant to the EOSC-synergy proposal:

EASY (Electronic Archiving SYstem) is the digital archiving system of KNAW-DANS. EASY offers public access to thousands of datasets in the humanities, the social sciences and many other disciplines. EASY can also be used for the online depositing of research data and EASY contains a non-public section serving as a so called Dark Archive for several (large) data owning organisations. https://easy.dans.knaw.nl/ui/home

NARCIS (National Academic Research and Collaborations Information System) is the national gateway to scholarly information in the Netherlands, providing access to over 260.000 datasets and over 1.600.000 (open access) publications in institutional repositories, and information about research, researchers and projects. KNAW-DANS is hosting as well as developing NARCIS. http://www.narcis.nl/

KNAW-DANS manages DataverseNL, a Virtual Research Environment used by Dutch universities and other research organisations for managing, sharing and storing research data. https://dataverse.nl

Key personnel

Ingrid Dillo *[F], is deputy director at KNAW-DANS and holds a PhD in history. Her research focused on the history of Dutch 18th century trading companies, shipping and maritime history. She has worked in the field of policy development for the last 25 years, including as senior policy advisor at the Dutch Ministry of Education, Culture and Science. Among her areas of expertise are research data management and the certification of digital repositories. She is a member of the International Board of the Data Seal of Approval (DSA), the Technical Advisory Board of Research Data Alliance (RDA), and the Board of Directors of the DRYAD repository. For the RDA she acted as the Interim Secretary General in the second half of 2017 till the beginning of 2018. She also is vice chair of the Scientific Committee of the ICSU/World Data System (WDS), co-chairs the RDA/WDS Interest Group on Certification of Digital Repositories and the RDA/WDS Interest Group on Cost Recovery for Data Centres, was an active member of the former RDA/WDS Repository Audit and Certification DSA-WDS Partnership, and participates in the Research Data Expert Group of the Knowledge Exchange. She will be the coordinator of FAIRsFAIR, the project awarded under the EOSC related H2020 call: INFRAEOSC-05C (see relevant previous project description).*

Eliane Fankhauser *[F], obtained her PhD about polyphonic music in the Netherlands in the late Middle Ages at Utrecht University. At KNAW-DANS she is a project manager and policy officer in the role of which she currently works on the external communication for the FREYA project and is the coordinator of the Oral History collection in the EASY repository. Together with a small team, moreover, she is working on the establishment of a tool (FAIRdat) for researchers to measure and evaluate the FAIRness of data.*

Ellen Leenarts *[F], is a project leader at DANS and has been involved in European projects as CESSDA SAW, EUDAT and OpenAIRE. She is part of the project teams of OpenAIRE Advance, EOSC hub and EOSCpilot because of her experience in Data Management and Education & Training. In 2017 she led the CESSDA RDM Training project that produced the widely used online CESSDA Expert guide on Data Management. Ellen leads work package 10 and work package 13 in the EHRI project since August 2017.*

Cees Hof *[M], received his MSc degree for research in the field of aquatic ecology and ecotoxicology and moved into animal systematics, taxonomy, palaeontology and geochemistry for his PhD research at the University of Amsterdam and postdoctoral research in the UK (Bristol). He was the coordinator of the European Network for Biodiversity Information (ENBI) and for more than 10 years developing the Dutch branch of the Global Biodiversity Information Facility (GBIF). Within the GBIF network Cees coordinated the activities of the European GBIF nodes from 2011 till 2013 and consolidated the network of Dutch biodiversity data-nodes. Nowadays Cees Hof works at KNAW-DANS as a project manager and project acquisitioner and is responsible for the KNAW-DANS interaction with the life science community. Current project responsibilities include the development of the (national) FAIR Software Route, a joint project with the Netherlands eScience Center (NLeSC), and communication and dissemination activities*

in work packages of EOSCpilot and EOSC-hub. From 2019 onwards, Cees will be involved in the coaching of the Essentials 4 Data Support training module, the successful (inter)national RDM course that KNAW-DANS organises jointly with SURFsara and the technical University Delft.

4.1.8 JISC

JISC will provide feedback at the national policy level (WP5) and outreach towards the UK Universities in WP6. In WP6 JISC will take care of coordinating the task for the development of best practices and guidelines for training material.

Description of the partner

Jisc is a not-for-profit organisation for digital services and solutions. Jisc is the UK's expert on digital technologies for research and education, serving the research community with a robust suite of national and international services including a world-class research support framework for repository services; the Janet network; cybersecurity services; access and identity management services; and key preservation, discovery and digital scholarship services. Jisc also provides analytics services, and technical, information and professional services to support research data management and open science.

Jisc is a responsive membership organisation, delivering economies and benefits for its members in line with their expressed needs. It serves a large community of higher and further education institutions in the UK and is experienced in reaching out to and engaging with its stakeholders to understand and respond to their requirements.

Jisc is an internationally recognised leader in open science, for example playing active roles in drafting the 2016 Amsterdam call for action (that informed the EU Competitiveness Council's outcomes on open science), in the Research Data Alliance, and in the global Open Scholarship Initiative launched in 2016. Jisc is the national NREN and a member of GÉANT; is the UK National Open Access Desk within OpenAIRE; is a member of the EUDAT CDI Council; alongside STFC, Jisc forms the UK NGI representation and membership of the EGI; is a founding member of the Knowledge Exchange, an international partnership of organisations in six European countries, focusing on open science; and has been the national representative to e-IRG, the e-Infrastructure Policy Forum and Digital ERA Forum. Nationally, Jisc is a lead partner of the UK Open Research Data Forum and is working with the UK Government to develop its Open Science and e-Infrastructure policies and strategies through membership of the Government's e-Infrastructure Advisory, as well as being a member of UKRI's e-Infrastructure Expert Group, working with UUK and UKRI to establish a roadmap for a more efficient and effective cross-disciplinary research infrastructure within the UK which also works internationally.

Jisc has been instrumental for many years in governance of pan-European e-Infrastructures, playing a strong leadership role in developing the governance of DANTE, which managed the GÉANT pan-European research network, and in the restructuring of DANTE, TERENA and GÉANT into a single organisation. Jisc lead a taskforce reviewing and revising the governance structure for EGI, lead the EUDAT2020 task in defining a governance structure for the EUDAT CDI. Jisc was also a member and rapporteur of the first High Level Expert Group on EOSC. Most recently Jisc has been leading the Governance Framework and Policy tasks within EOSCPilot, resulting in publication of the first draft of the Governance Framework in November 2017 which has informed both the EOSC Implementation Roadmap and the 2nd EOSC High Level Expert Group draft recommendations.

Key personnel

Matthew J. Dovey *[M], works at Jisc as Head of e-Infrastructure Strategy. He oversees work within Jisc Technologies to develop digital infrastructure services which support and enhance aspects of the research lifecycle – from discovery of information and data, to data analysis and manipulation, and collaboration and research impact and dissemination. He has developed governance structures for both EGI and EUDAT, as well as developing the EOSC Governance Framework in EOSCPilot. Matthew is currently the Chair of the Executive Board and Council for EGI.eu, member of the EUDAT CDI Council, member of the GÉANT General Assembly, and a member of the e-IRG advisory board. He is also a member of the UK Cabinet Office Open Standards Board and the Software Sustainability Institute Advisory Board. Previously, Matthew was Technical Manager at the Oxford University e-Science Centre, where he advised scientific research projects based on WebService and GridService architectures. Prior to this, he worked for the Oxford University Library Services, implementing numerous library and digital library technologies and projects on preservation of digital material, and conducted research on music information retrieval at Kings College London.*

Helen Blanchett *[F], provides advice, training and consultancy in the area of scholarly communications, with a particular focus on open access. Helen has over 20 years' experience working in training and staff development in the education sector. With a background in academic libraries, Helen has a keen interest in all aspects of information and digital literacy, and in supporting staff and students in their development. She worked for Jisc Netskills, a Jisc training service, for 13 years and delivered training programmes around e-learning, 'train the trainer', information and research skills and plagiarism awareness. She has led a multi-stakeholder Jisc-funded projects to develop a continuing professional development framework and assessment tool for business and community engagement and knowledge exchange.*

Dale Robertson *[F], has worked for Jisc since September 2017, and leads the Policy Engagement work of the EOSCpilot project, coordinating delivery of recommendations for policy action to encourage the EOSC as well as contributing to the Governance WP and liaising with the Rules of Participation team on the interplay of policies with the Rules of Participation. Prior to this she worked for GÉANT, heading their EU Policy Engagement function to position GÉANT relative to EU Digital Single Market, Research & Innovation and other relevant policies, and delivering future funding streams to secure GÉANT's*

future. Before that she led the GÉANT Marketing Communications team, developing it into a broad department which incorporated online presence and audio-visual content, events, literature, analytics and user surveys as well as the nascent user support function. Dale has a background in European Public Affairs, having previously worked inside the European institutions in Brussels and also for Sony Europe and the British Broadcasting Corporation, and in addition became an accredited SAP Human Resources and Payroll consultant, working in the UK, Ireland and France. She holds an MA in Mathematics from the University of Cambridge, a Postgraduate Diploma in International Studies from Johns Hopkins SAIS Bologna Center, and a Public Affairs Diploma from the Chartered Institute of Public Relations in the UK.

4.1.9 IRD/LEGOS

IRD will lead the validation of services for Marine Earth Observation in EOSC-synergy (WP4).

Description of the partner

French National Research Institute for Sustainable Development (IRD)

The French National Research Institute for Sustainable Development (IRD; *http://en.ird.fr/the-ird/presentation*), an internationally recognized multidisciplinary organization working primarily in partnership with Mediterranean and inter-tropical countries, is a French public establishment under the joint authority of the French Ministry of *Higher Education and Research* and the Ministry of *Foreign Affairs and International Development*. Via its network and presence in fifty or so countries, it takes an original approach to research, expertise, training and knowledge-sharing, to the benefit of countries and regions that make science and innovation key drivers in their development.

Laboratory of Geophysical Studies and Spatial Oceanography (LEGOS)

The LEGOS is a "mixed unit" comprising staff from four employing organizations: *Centre National d'Etudes Spatiales* (CNES), the *National Center for Scientific Research* (CNRS), the *French National Research Institute for Sustainable Development* (IRD) and the *University Paul Sabatier* (UPS). It is involved in environmental research centered on physical oceanography (large scale and coastal), marine biogeochemistry and geochemistry, satellite hydrology and glaciology, using modelling approaches and *in situ* oceanographic data and remote sensing as primary observation tools. It plays a leading role in major international programs such as SOLAS, Megha-TROPIQUES, SWOT and GEOTRACES. LEGOS has about 100 members: 40 researchers (CNRS, IRD, UPS, CNES), 20 technicians and engineers, 30 postgraduate students and postdoctoral and contract staff.

Key personnel

Fabien Durand *[M], PhD has a 20-year long experience in hydrodynamics of the tropical oceans. He has published 50 peer-reviewed publications (H-index of 21). He is an expert in the hydrodynamics of the rivers-estuaries-ocean continuum, with specific interest in the flooding dynamics during extreme events (cyclone surges). He has been a co-I in the SMOS satellite mission (ESA) and AltiKa altimetric satellite (CNES/ISRO).*

4.1.10 UPVLC

The University Polytechnique of Valencia will lead the Work Package of thematic services integration and validation (WP4). It will also participate in WP3 to bridge the activities of software maturity with the work developed in WP4. UPVLC will also promote the developments of WP6 at the University of Valencia.

Description of the partner

The Grid and High Performance Computing Group (GRyCAP) from the UPV, has a very strong background in Distributed, Cloud Computing and Data Management. GRyCAP has shared and extended its expertise through active participation in over 30 national and European R&D projects on HPC, Grid and Cloud technologies. These projects include HPCN-TTN Network, EUTIST-M and TT@MED – technology dissemination; EGEE-I, II, III, EGI- InSPIRE, EGI-ENGAGE and VENUS-C – e-infrastructures. The group took part in the International HealthGrid Association; has developed a roadmap on the use of grids in health (SHARE), and boasts a solid research line in the distributed storage of medical imaging data through the TRENCADIS platform (CVIMO project). Relevant GRyCAP expertise includes: the coordination of the cloud end-user community (27 applications, 20 coming from an Open Call) in the VENUS-C project, one of the first cloud projects to feature in the DAE ; the development of distributed and Grid-Computing platforms with Latin American countries: CyTED-GRiD, EELA and EELA-2, the EU-BrazilOpenBio Project,the EUBrazil Cloud Connect project and the EUBRA-BIGSEA project (last two coordinated by the UPV), and development of high-level middleware components for cloud infrastructures in the Spanish Codecloud and CLUVIEM project) and coordination of the Spanish Network for e-Science to implement the Spanish National Grid Initiative (NGI). The UPV also participated in the INDIGO-DATACLOUD H2020 project on cloud middleware services development.

Currently the UPV leads the ATMOSPHERE project and participates in EOSC-Hub and DEEP Hybrid-DataCloud H2020 projects.

Key personnel

Ignacio Blanquer *[M], associate professor (accredited full professor) of the Computer System Department at UPV since 1999 and he has been involved in Parallel Computation and Medical Image processing, participating in more than 60 national and European Research Projects, has authored and co-authored 40 articles in indexed journals and book chapters and in more than 80 papers in national and international journals and conference proceedings. He has served as coordinator of the application area in the Spanish Network for e-Science. He has been the project coordinator of EUBrazilCloudConnect (FP7) and EUBra-BIGSEA (H2020) and currently he is the coordinator of ATMOSPHERE (H2020) and co-principal investigator in the BigCLOE national research project.*

Germán Moltó *[M]****,*** *associate professor at the Department of Computer Systems and Computation (DSIC) at UPV. He has published over 70 papers in conference proceedings and more than 22 contributions to JCR-indexed journals in the areas of Cloud Computing, Grid Computing, High Performance Computing and Scientific Computing. He has been involved as a researcher in 20 R+D projects. Since 2009, he has been responsible for two national research projects in the area of Cloud Computing (CLUVIEM and BigCLOE).*

4.1.11 LNEC

LNEC will participate in the integration and validation of services related to Marine Earth Observation (WP4) where they will co-lead the work package with UPVLC. They will also provide input to the service integration platform in terms of service validation (WP3). At the national level they will also provide input for policy development from the point of view of engineering communities (WP5).

Description of the partner

National Laboratory for Civil Engineering (Laboratório Nacional de Engenharia Civil - LNEC), established in 1946, is a Science and Technology (S&T) public research institution devoted to the various fields of civil engineering and related environmental areas. It has a staff of about 500 members and 125 research fellows. Researchers represent 30% of the personnel and have got a PhD or equivalent degree. Over the past seven decades, LNEC conducted 500+ studies in over 40 countries, within the framework of R&D studies and advanced technological consultancy. Between 2007 and 2016, LNEC participated in over 150 R&TD projects, about 50 being funded by European Programmes (e.g., H2020: BINGO, WADI, Co-ReSyf, USE-iT, RESCCUE, EOSC-hub, CEDR: PROPER).

LNEC's Hydraulics and Environment Department (DHA) carries out R&TD in all the components of the water cycle (maritime and coastal engineering; surface and groundwater water resources; sanitary engineering; information technologies applied to water and environmental domains). DHA has a vast

experience in the development, maintenance and operation of state-of-the-art modeling systems for the dynamics of water bodies, being the coordinator of the H2020 project BINGO on the mitigation of the impacts of climate change on the water cycle. DHA is also partner of the H2020 project Co-ReSyF, where an algorithm to retrieve bathymetry using Synthetic Aperture Radar (SAR), from Sentinel-1 and TerraSAR-X satellites was developed. DHA has also a strong background on real-time forecast systems and monitoring data networks, to investigate the hydrodynamics of water systems and support the emission of alerts to the authorities on emergency events. DHA has a body of researchers devoted to hydrodynamic research in coastal regions, rivers and cities, supported by information technology experts, which intertwines water and computer developments. LNEC is also one of the partners of the Portuguese National Distributed Computing Infrastructure, integrated in the Portuguese Science Foundation Digital Research Infrastructures Roadmap. In this scope, DHA lead the pilot cloud user validation initiative, exploring the use of CPUs and GPUs for computationally demanding tasks, such as forecast systems and remote sensing image processing. Presently, DHA is also partner of the H2020 EOSC-hub project, which resulted in the integration of the OPENCoastS service in the EOSC digital infrastructure.

Key personnel

Alberto Azevedo *[M], is a Research Officer of the Hydraulics and Environment Department of LNEC's Estuaries and Coastal Zones Division. He has a Ph.D. in Geophysical Sciences and Geoinformation, with specialization in physical Oceanography (University of Lisbon, 2010). He has 16 years of experience in remote sensing, High-Performance/Cloud Computing and numerical modelling (hydrodynamic and oil spills). Since 2004, he has collaborated in over 12 projects and published 16 scientific papers related with oil spill forecast systems, High Performance Computing (HPC) applied to numerical modeling and image processing algorithms, coastal zones protection and inundation studies. He is responsible for the real-time forecasts of the oil spills modeling systems and satellite remote sensing activities in the Estuaries and Coastal Zones Division of LNEC.*

Anabela Oliveira *[F], is a Senior Researcher of the Hydraulics and Environment Department and the head of the Information Technology in Water and Environment Group. She has a Ph.D. in Environmental Sciences and Engineering and a B.Sc. in Civil Engineering, with 25 years of experience in hydrodynamics and water quality estuarine modelling. In the last decade she has been devoted to the establishment of forecast platforms for hydrodynamics and water quality in coastal regions and on-line, real-time monitoring networks, combining conventional and innovative sensors. She has published over 60 papers in Science Citation Index Expanded indexed journals and has an h-index of 17. She has participated in several research projects, both nationally and internationally-funded. Examples include as coordinator: MADyCOS, G-cast, pac:man, C-WOS, BGEM, SI-GeA; and as team member: PREPARED FP7, INTERREG SPRES, MOLINES, RealQual, AQUAGISMON. She is currently the responsible at LNEC of H2020 project WADI.*

André Fortunato *[M], is a Senior Research Officer of the Hydraulics and Environment Department and former head of LNEC's Estuaries and Coastal Zones Division (2003-2012). He has a Habilitation degree in Maritime Hydraulics, a Ph.D. in Environmental Sciences and Engineering, and a B.Sc. in Civil*

Engineering. He has 25+ years of experience in the development and application of numerical models to simulate estuarine and coastal processes (hydrodynamics, wave propagation, sediment dynamics and water quality). He has participated in over 20 research projects involving hydrodynamic modeling (e.g., FP5 HarmoniQua, FP7 Prepared, H2020 BINGO), and is responsible for LNEC's NE Atlantic hydrodynamic forecast system. He has published 70+ papers in scientific journals and has an h-index of 18 in the Science Citation Index Expanded.

4.1.12 ACK-CYFRONET

CYFRONET will participate in expanding EOSC capacity in Poland with the resources of the institution (WP2). In WP3, it will provide support to the integration of services in the EOSC-hub Marketplace.

Description of the partner

The Academic Computer Centre CYFRONET (http://www.cyfronet.pl/en), University of Science and Technology AGH (referred to as AGHUST) is one of the leading academic computer research institutions in Poland. The research unit of the Academic Computer Centre CYFRONET AGH, together with the Department of Computer Science AGH, focuses on scalable distributed systems, cross-domain computations in loosely coupled environments, knowledge management and support for life sciences. The team took part in a series of successful scientific projects funded by EU, including CrossGrid, Pellucid, K-Wf Grid, Int.eu.grid, ViroLab and GREDIA. Other projects include gSLM and EDA EUSAS. Since April 2004 AGHUST participated in series of EGEE projects (I, II and III) and EGI-InSPIRE, being responsible for operations in Central Europe region including SLA enforcement and coordination of resource allocation activity in the project. Currently, AGHUST is the coordinator of the PL-Grid project, co-funded by the European Regional Development Fund as a part of the Innovative Economy Program (National Grid Initiative, Polish NGI - http://www.plgrid.pl/en), with the goal to provide the Polish scientific community with an IT platform based on Grid computer clusters and cloud environments, enabling e-science research in various fields. The PL-Grid infrastructure is both compatible and interoperable with existing European and worldwide Grid frameworks. Due to this fact CYFRONET has substantial expertise in coordinating federated Grid infrastructures. Furthermore, CYFRONET has substantial experience in Virtual Organization management and automated contract negotiation and enforcement in heteregenous IT infrastructures, based on research performed in several EU and national projects. The research in this area involved in particular development of knowledge-based automated contract negotation framework utilising domain specific knowledge for reaching a consensus among multiple agents as well as automatic establishment of the Virtual Organization in the IT infrastructure of

the participating organization and further enforcement of the contract and specific negotatied SLA's by utilizing underlying monitoring infrastructures.

Currently AGHUST closely cooperates with EGI.eu and participates in EGI- Engage project. AGHUST computing centre CYFRONET is the most powerful computer centre in Poland offering all kinds of computer and storage resources to the scientists.

CYFRONET's 2 most powerful machines include the Prometheus and Zeus clusters. Prometheus, ranked in the TOP 500 list (November 2017) at 77th position, is a most powerful supercomputer in Poland.

Dedicated computing environments, so-called domain grids, and specialised IT platforms facilitate conduction of increasingly complex research problems. Modelling of energy demands, drug and new material design, or simulation of complex metallurgical processes is just a fraction of research issues studied within nearly thirty domain grids operating in the PLGrid infrastructure.

The research portfolio carried out with the help of the Zeus and, recently, Prometheus is quite reach. It includes: prediction of 3D protein structures, study of semiconductor nanostructures and catalytically activity molecules as well as effective biosensors. Computations are used to study the behaviour of galaxies in a wide range of electromagnetic spectrum, for nuclear magnetic resonance modelling for the purposes of structural analysis of molecular systems, antidots in quantum world, for structural characteristics of human telomeres and complexity of the financial markets. Scientific computations do not include simulations only. Computing power is utilised by Polish scientists also within international projects like CTA, EPOS, LOFAR and Large Hadron Collider in CERN. With the help of dedicated software packages the supercomputers perform analyses of large and dispersed data sets as well as provide advanced visualisations.

Key personnel

Tomasz Szepieniec *[M], IT project manager and architect, lead of the FID team in Cyfronet. He was involved in series of EU and Polish projects both in academia and in industry acting as project manager, team leader or consultant in IT architecture. In the frame of EOSC-hub project he is responsible for marketplace service. In Indigo-DC he was responsible for SLA Manager and federated CMDB. Tomasz received M.Sc. in Computer Science in 2003, authored about 30 scientific papers and co-authored FitSM Standard.*

Dr Lukasz Dutka *[M], has significant expertise in Grid systems, large-scale systems, development of applications for business purposes, team and project management in commercial projects as well as EU IST projects. He obtained his M.Sc. in Computer Science from the Jagiellonian University, Poland and a Ph.D. in Computer Science from the University of Science and Technology, Cracow, Poland. He has actively participated in several EU IST projects including CrossGrid, EGI- Engage, Indigo-DataCloud. Since 2008, he is a Technical Director of PL-GRID project and he is responsible for full operation of the infrastructure including R&D tasks.*

Roksana Dobrzańska *[F], Member of Federated Infrastructure Development team in ACC Cyfronet AGH, Poland. She took part in series of EU and Polish projects, coordinating the delivery of numerous access tools for research. Responsible for family of solutions supporting IT Service Management integrated with large scale computing and infrastructures. She is an expert in and contributor to FitSM standard family. Person responsible for implementation and delivery of EGI Marketplace, leader of Marketplace activity in EOSC-hub. Few years' experience with designing and implementing operational tools for various infrastructures in scope of European and Polish research projects.*

4.1.13 FCT

The Portuguse foundation for Sceince and technology will participate in the alignment of policies towards FAIR data practices (WP5) and will collaborate in the identification and integration of Portuguese data repositories working with INCD and LIP.

Description of the partner

FCT is a public institution of the Ministry of Education and Science. On 1 October 2013, FCCN became one of FCT's organic units. Originally, in January 1987, FCCN started its activity as a private not-for-profit organisation. In that period, with the support of universities and various national R&D institutions, it has played a central role in the expansion of the Internet in Portugal.

The main activity of FCT is to fund science and technology. The FCCN unit's central role concerns the planning, managing and operating of the Science, Technology and Social Community Network (Rede Ciência, Tecnologia e Sociedade – RCTS), the Portuguese NREN.

RCTS is a high-performance network servicing organisations with the most demanding communication requirements, and is thus an experimentation platform for advanced communications software and services. It uses Internet protocols to provide a communication and cooperation platform between public and private institutions acting in the fields of education, science, technology and culture.

More information about FCT and FCCN is available at: www.fct.pt and www.fccn.pt/en.

Key personnel

João Mendes Moreira *[M], has a degree in computer science. He has 20 years of experience managing infrastructures and advanced services aimed at the national Research and Education community. Presently he is the Head of Scientific information of FCCN, a unit of the Foundation for Science and Technology of Portugal. This area comprises the Online Library of Knowledge (b-on), the Portuguese Open Access Initiative - RCAAP (Portuguese Open Access Scientific Repository) and the Portuguese Current Research Information Ecosystem (PTCRIS). He is member of several committees and working*

groups: : EOSC Board, Advisory Board member of European projects OpenAIRE and eInfracentral, DSpace Steering Committe, SCOAP3 Repository Steering and Support Working Group.

4.1.14 EGI.eu

The EGI.eu foundation will take care of communication of outreach for EOSC-synergy, building on then extensive network of contacts of the organization (WP1). It will also have a policy role (WP5) towards the coordination of activities with other international initiatives and projects.

Description of the partner

The EGI Foundation (also known as Stichting EGI and abbreviated as EGI.eu) is a not-for-profit foundation established under the Dutch law to coordinate the EGI Federation (abbreviated as EGI), an international collaboration that federates the digital capabilities, resources and expertise of national and international research communities in Europe and worldwide.

The main goal is to empower researchers from all disciplines to collaborate and to carry out data- and compute-intensive science and innovation.

The EGI Foundation has participants and associated participants drawn from representatives of national e-infrastructure consortiums (NGIs), EIROs, ERICs, and other legal entities. These entities provide the physical resources and shared services that enable EGI to deliver, improve and innovate services for communities.

The EGI Foundation coordinates areas such as overseeing infrastructure operations, user community support, contact with technology providers, strategy and policy development, flagship events and dissemination of news and achievements. The EGI Foundation holds certifications in both ISO/IEC 9000 “Quality Management” and ISO/IEC 20000 “IT Service Management”. As part of its mandate, the EGI Foundation actively represents the EGI federation at European level with policy makers and funding agencies, it provides expert advice to shape policies and funding programs and also support the implementation of the policy priorities. In particular, in the area of open science, EGI is a key stakeholder represented in and contributing to the EC Open Science Policy Platform.

Through its services for High Throughput Computing, Cloud, Federated Operations and Community-driven innovation and support (https://www.egi.eu/services/), EGI actively supports the European Open Science Cloud initiative and leads the EOSC-hub project.

Publications

Answering the Cost Assessment Scaling Challenge: Modelling the Annual Cost of European Computing Services for Research, Matti Heikkurinen, Sandra Cohen, Fotis Karagiannis, Kashif Iqbal, Sergio Andreozzi, Michele Michelotto, Journal of Grid Computing, May 2014, DOI: 10.1007/s10723-014-9302-

y Matti Heikkurinen, Sandra Cohen, Fotis Karagiannis, Kashif Iqbal, Sergio Andreozzi, Michele Michelotto. 2014.

EGI: an Open e-Infrastructure Ecosystem for the Digital European Research Area, Proceedings of eChallenges e-2012 Conference, Lisbon, Portugal, ISBN: 978-1-905824-35-91 Sergio Andreozzi, Sy Holsinger, Damir Marinovic, Steven Newhouse. 2012.

Open Science Commons, Sergio Andreozzi, Owen Appleton, Sara Coelho, Tiziana Ferrari, Sy Holsinger, Yannick Legré, Jan 2015, http://go.egi.eu/oscwp Sergio Andreozzi, Owen Appleton, Sara Coelho, Tiziana Ferrari, Sy Holsinger, Yannick Legré. 2015.

"Report on the governance and financial schemes for the European Open Science Cloud https://ec.europa.eu/research/openscience/pdf/ospp_euro_open_science_cloud_report-.pdf" S. Andreozzi & Al. 2017.

The user support programme and the training infrastructure of the EGI Federated Cloud, in the proceeding of High Performance Computing & Simulation (HPCS) 2015, Amsterdam, the Netherlands. DOI: 10.1109/HPCSim.2015.7237016. Fernandez, E., Scardaci, D., Sipos, G., Chen, Y., Wallom, D. Fernandez, E., Scardaci, D., Sipos, G., Chen, Y., Wallom, D.. 2015

Key personnel

Sergio Andreozzi *[M], is strategy and policy manager of the EGI Foundation, the coordinating body of EGI (e-infrastructure supporting data- and compute-intensive research and innovation). In his role, Sergio contributes to strategic planning and execution, governance, business models and evolution of services. Sergio holds the role of Service Portfolio Manager and Chair of the Services and Solutions Board (SSB) for EGI. Other responsibilities comprise contributing to developing project proposals to implement the EGI strategy or innovative ideas, leading activities in projects, authoring external communication messages as well as organising and participating in meetings, forums and conferences. Sergio is also member of the EC Open Science Policy Platform. He has been involved in grid and distributed computing since 2002 contributing to technical activities, standardisation within OGF, analysis and development of business models, interoperability, policy analysis, sustainability, IT service management among others. He holds a PhD in Computer Science (University of Bologna) and a MSc in Computer Science Engineering (University of Pisa).*

Roberta Piscitelli *[F], is Strategy and Policy Officer at EGI.eu with a focus on strategic planning and execution, business model development, and federated service management. Other responsibilities comprise contributing to developing project proposals to implement the EGI strategy or innovative ideas, leading activities in projects, authoring external communication messages as well as organising and participating in meetings, forums and conferences. She has been involved in R&D projects in Big Data Analytics, Deep Learning covering both research and engagement with SMEs in the Big Data sector. She holds a PhD in Computer Science from University of Amsterdam.*

Sara Coelho *[F], works at the EGI Foundation implementing and leading the communications strategy for the organisation and the EGI-Engage project. She has been a part of the EGI Foundation team since 2010 and has been involved in all aspects of the organisation's communications activities including copy writing, website maintenance, production/design of promotional materials, community building and event organisation. Alongside a PhD in Geology, Sara has an MSc in Science Communication and her previous experience includes working as a reporter for AAAS Science and the UK's Natural Environment Research Council.*

Iulia Popescu *[F], works as Communications Officer at the EGI Foundation. She joined the communications team in 2015 and has an international background in Communications and Marketing. She obtained her master's degree in Communication at the University of Amsterdam and previously worked for large companies such as Hewlett-Packard and TomTom as well as two SMEs in Barcelona. Iulia supports the EGI Communications activity in all aspects, from website maintenance, writing for all communications channels and social media management.*

Yannick Legré *[M], is the director of EGI.eu since February 2014. Formerly he was a senior research engineer at the French National Scientific Research Centre - Grid and Cloud Institute (CNRS-IdGC). He holds a Master of Science in Information Technology (MScIT) and a degree in Law (LL.L). Over the last 15 years, Yannick has been involved in more than 50 projects in the areas of e-Infrastructures, healthcare and biomedical research, as well as biodiversity and environmental research. Yannick has successfully coordinated several EC, NATO and nationally funded projects. He has also been a co-founder and the president of the international HealthGrid association, and the director for International Relations of a French SME.*

4.1.15 INCD

INCD will integrate the resources of the Portuguese National Distributed Computing Infrastructure in EOSC (WP2) and promote the EOSC-synergy training program in the Universities of the North of Portugal.

Description of the partner INCD

The Portuguese National Distributed Computing Infrastructure (INCD) is a digital research infrastructure that provides scientific computing and data services to the Portuguese academic and research communities. INCD was established in 2015 resulting from the evolution of the Portuguese National Grid Initiative towards an autonomous body with a broader scope of technical activities and competences in the context of the Portuguese Science Foundation (FCT) roadmap of research infrastructures.

The INCD infrastructure delivers cloud computing, high throughput computing and high performance computing services to a wide range of users in many scientific domains. On-top of these core services INCD provides a set of added value services tailored to needs of specific research domains and user

communities. The INCD infrastructure is currently being enlarged both in terms of the service portfolio and to encompass new resource centers in different geographic locations. The INCD infrastructure is member of the Iberian distributed computing infrastructure (IBERGRID) and of the European grid Infrastructure (EGI), in addition INCD plans to establish in 2019 an EUDAT data storage node. INCD is currently participating in EOSC related activities indirectly via its member organizations LIP and LNEC. Within this project proposal INCD aims to bridge this gap and engage directly in EOSC. Through partnership with its users and member organizations, INCD participates in the integration and delivery of national thematic services.

Key personnel INCD

João Machado *[M], is a computer engineer and researcher at INCD. He has a Ph.D. in Marine and Environmental Sciences obtained at Instituto de Ciências Abel Salazar of Porto University, and a bachelors's degree in Computer Engineering from Instituto Superior de Engenharia do Porto (ISEP). He worked in genomics research projects at Instituto de Ciências Abel Salazar and CIIMAR in the area of bioinformatics. He has experience in software development with agile methods, scientific data analysis, and computing infrastructures. He currently works in the deployment and development of the INCD infrastructure services*

Catarina Ortigao *[F] is member of the Portuguese National Distributed Computing Infrastructure (INCD). She holds a PhD in Physics by the Lisbon Technical University with research in the development and experimental study of positron emission tomography and an MsC also in Physics. She has been involved in research and development projects in the areas of Medical Physics, Virtual Reality and Space Science in academia and industry. She is the INCD liaison officer coordinating the relations and communications with user communities. She is currently managing the INCD participation in national and international projects.*

João Pina *[M], is a computing researcher at LIP. He has a PhD in Physics with research work in the ATLAS detector at the CERN Large Hadron Collider (LHC). He joined the LIP computing group as post-doc researcher, to work on grid computing technologies acting as contact point for the Portuguese LHC community. In 2013 joined the European Grid Infrastructure (EGI) Software Provisioning team, as deputy of the Staged Rollout and later in the same year as leader. Under this role has become a member of the EGI UMD Release Team, responsible for the coordination of the EGI software stack and to liaise between the several international development teams. In 2014 was nominated Regional Contact Point for the EGI Spanish federation and National Infrastructure Liaison for Portugal. In 2015 became member of the FCT Advisory Group for the Future Emerging Technologies (FET) and Marie Curie H2020 programs. He participated in the EGI-Inspire and EGI-ENGAGE projects as well as in the WLCG and IBERGRID infrastructures. He is currently collaborating with the Portuguese National Distributed Computing Infrastructure (INCD) and leading the Configuration Management, Change Management, Release and Deployment Management in the EOSC-hub project.*

4.1.16 INDRA

INDRA will participate in the integration and expansion of the G-Core service in WP4.

Description of the partner

Indra is a global company with the vocation, experience, and capabilities to attract third-parties. It is a publicly traded company. All the shares are officially listed on the Spanish stock markets since 1999, on the selective IBEX 35 index, which includes the 35 most representative. It also forms part of major international indexes such as the Dow Jones STOXX Broad Market Index, which includes the main European listed companies, and the STOXX Europe Technology and IT Services indexes, which list the leading European technological shares. Similarly, since September 18, 2006 Indra's shares have been listed on the Dow Jones Sustainability World Index (DJSWI), which selects companies that make the greatest effort to adapt their activities to meet sustainability criteria from among the largest companies in the world. The company has a diversified market structure. 39% of the sales are originated in Spain, whereas the remaining 61% comes from international. Indra business model is based on innovation and sustainability with more than 192 agreements with universities and research centres and with 241 alliances with technology partners. Indra also has 55 Centers of Excellence and 22 Software Labs worldwide, operating branches in 46 countries.

The space unit contributes significantly to defence market incomes with strong presence in the most important projects developed in Europe in the aerospace field. This unit involves four major areas: Earth Observation Solutions, Satellite Control Systems, Satellite Communications Systems & Equipment and Satellite Navigation. In particular, the Earth Observation group is the leading Spanish team in Ground Segment for Space Systems and Earth Observation Applications, with more than 25 years of specific experience in the Space sector and a highly specialized staff. A multidisciplinary group of 60 engineers specialized in Satellite Ground Segment Systems Engineering, Remote Sensing, GIS and advanced Earth Observation product development, data processing, value adding and consultancy, forms this unit.

It is precisely this combination of Software and System Engineering with Earth Observation and geospatial data value adding what enables the company to provide advanced solutions to its customers. The next bullets identify the relevant activities developed by this unit:

- Design, development, integration and operation of complete Earth Observation ground segments. Covering PDGS, full ground segment (including both PDGS and FOS) and Service segment (imagery sales and high level exploitation system for the satellite operator).
- Design and development of high level EO data/geoinformation exploitation centres, including automatic value adding processing of satellite imagery and Enterprise GIS for defence and civilian environments. Strong experience in Spatial Data Infrastructures, interoperability and OGC web services
- EO data processing, geo-information supply and consultancy. This includes a leading role in Copernicus Services in Security (EUSC), Land (EEA) and Emergency (JRC). Working in GMES/Copernicus since 2003.

The reference customers of this unit are the Spanish Ministry of Defence, the Spanish Ministry of Environment, the European Space Agency, the Spanish National Geographic Institute (IGN), the National

centre of Geographical Information, the European Environmental Agency, the Municipality of Madrid, Spanish Civil Protection, the Satellite Centre of EU, the French DGA, the Belgium MOD, the European Commission, Hispasat, Hisdesat, EUTELSAT, EUMETSAT, AENA, INTA, Thales Alenia Space and Airbus.

The relevant projects developed by this unit are describe below

- Indra has a relevant role in the Galileo program. It provides the control stations since 10 years.
- Indra has contributed and leaded the following Ground segment; Helios, Pleiades, SMOS (Soil Moisture and Ocean Salinity mission), PAZ, Ingenio and S3T (Space Surveillance and Tracking system of ESA)
- Indra is participated in the following R&D programs in the field of Earth Observation services in Europe; G-SEXTANT Service Provision of Geospatial Intelligence in EU External Actions Support. GNEXT pre-operational services of Copernicus to support the foreign operations of EU. G-MOSAIC GMES services for management of operations, situation awareness and intelligence for regional crises. DECUMANUS Earth Observation component of the smart city solution of Indra. PROMETEO, National R&D project promoted by Spanish Centre for the Development of Industrial Technology to provide services and products for surveillance and monitor wildfires.
- Indra has contributed to develop the Spatial Data Infrastructure (SDI) of Spain according INSPIRE directive.
- Indra has developed during more than 10 years a technical staff in the National Geographic Institute and the National Center of Geographic Information to perform different web sites and contributes to national Spatial Data Infrastructure (SDI).
- Indra has developed a Direct Access Terminal (DAT) to provide direct access to PAZ and Terrasar-X data directly for public institutions or private company in accordance with Hisdesat.
- Indra is developing the Commercial services of Hisdesat, Governmental operator of Spanish Earth Observation missions.
- Indra is currently hosting one of the PAC (Processing and Archive Center) of Sentinel program in its own Data Center.
- Indra has a specialized staff to operate Ground Segment and Control Stations.
- Indra has a specific unit for Digital data production: reference mapping with references as MGCP or Copernicus Security Border Monitoring; land mapping, including active participation in Copernicus Land Monitoring Service; and digital imagery, including satellite imagery distribution. This unit is currently developing Earth Observation engines to generate added value products and services based mainly on Sentinel data. Also it develops an AIRINT service to manage and generate products from UAV data The main strengths supplied by Indra are described as follow:
- Suitable staff and resources to satisfy the customer´s requirements and expectations, with the proper skills and deep technical knowledge, specially with extensive background on the subjects of interest of space solutions. More than 30 years developing space solutions for public sector in Europe and private companies abroad the world. The most relevant references are described below.
- Acknowledged experience in managing large projects, including multi-million Euro programs for defence and large institutional customers. The company has a large base of skilled project managers and most of them have achieved the Project Manager Professional certification from the Project Management Institute. Furthermore, the company is a fully ISO certified company with strong methodologies and tools to support the management of large programs
- Ensures mature solutions, products, engineering methods and practices along the years. Indra has mastered all technologies related to Earth Observation products needed to provide complete state-

of-the art solutions. Indra periodically undergoes audits by independent bodies for the certification of its management and production systems in accordance with leading international standards, amongst these: UNE-EN-ISO 9001, PECAL 2110, PECAL 2210, PECAL 2310, UNE-EN 9100 aerospace series, UNE-EN 9110, CMMi, TMMi, UNE-ISO/IEC 27001,UNE-ISO/IEC 20000-1: Excellence of IT services management

- Ensure the sustainability of products or services to develop in R&D projects due to strong presence in different countries and markets. The commutation unit of Indra has a relevant position in the social media market. The company international communication policy is permanently supported by the communication teams at its national offices and by local communication agencies, including Burson-Marsteller, Kreab and Gavin Anderson, Cohn & Wolfe, Llorente y Cuenca, Máquina, Apoyo Comunicación, and Inrete. It has professionals working specifically on Communication tasks in Italy, Germany, the United Kingdom, Portugal and the majority of Latin American countries.

Key personnel

Juan Sánchez-Ferrero de Pablo *[M], Indra staff with more than 17 years of experience in EO projects with a strong background in Ground Segment Systems participating in its design, development and technical management with large industrial teams in the European context for defence and civilian markets. Juan (MSc in Physics at Complutense University of Madrid) is responsible of the System's Engineering area at Indra EO unit. Juan also has participated as teacher in Master courses for the university and private institutions. He is actually technical responsible of the User Service project for Hisdesat to set the commercial services for coming satellites PAZ and SEOSAT. He has participated in several missions participating in SEOSAR/PAZ satellite Ground Segment system as deputy technical manager in Ground Segment engineering, participating also in the SEOSAT/INGENIO consolidation study and in the SMOS DPGS project. He also has been technical responsible for Helios-II supporting tools, SMOG (Impact of Small Missions on EO Ground Segment Systems), UGEI (Universal Geometry Engine) projects for the ESA and also in the SIGESTREDI (System of Spanish Department of defence for the management, exploitation and distribution of the geospatial data) for the Spanish Ministery of defence (MoD)*

José Manuel Sánchez Granero *[M], Indra staff with 12 years of experience with a strong background in EO system engineering, leading the development group of Earth Observation unit. He is the product manager of GCore suite, specific framework for processing EO data. José Manuel is a computer engineer at Complutense University of Madrid specialized in systems engineering and computation. He has leaded the processing component of SMOS (The Soil Moisture and Ocean Salinity mission) and PAZ Ground Segments, currently he is leading the processing framework of Ingenio, Processing Archive Centre (PAC) of Sentinel data and Sentinel preoperational service for reprocessing campaigns. He has very good skills for technical leadership and coordination, very strong analytical and problem-solving, great experience in innovation management and technological surveillance. He also support the technical activities required for business development and commercial activities thanks to technical background and experience in systematic processing system and new technologies.*

Manuel Pavesio Blanco *[M], Indra staff with more than 17 years of experience in EO projects with a strong background in SDI projects, working in the National Center of Geographic Information to develop the SDI national programs. He has also complemented a sound experience in ground segment operations participating in its development and technical management of Earth Observation system operations with large industrial teams in the European context for defence and civilian markets. Manuel is a Geomatic and Surveyor Engineer, EMBA in new digital markets by the Technical University of Madrid (UPM), Founding member of National Professional Association of geodetic and geospatial Engineers and Legal Expert in Surveying and geospatial information. At present time he is the project manager of commercial services of the "Service Segment" of Paz under contract from Hisdesat. He has a very strong team spirit with initiative and proactivity and high degree of commitment to ongoing work with a very good overall interpersonal awareness.*

José González Ruiz *[M], Indra staff, head of remote earth observation unit, with more than 30 years of experience encompassing the definition, design, verification and management. Currently he leads the management and technical activities in Indra related to ground space elements. Besides, representing Indra in different working groups for the definition of the ground space policy. José is MSc in Physics at Science University of Valladolid, specialized in electronic field. He has a wide knowledge in Space solutions and an extensive international experience as advisor, lecturer, CEO and CTO in the space field such as Galileo program, Copernicus and Spanish Earth Observation program.The nature of his experience, expertise and wide range of skills would be a valuable assess in general and in particular to several domains with cover from designing large complex systems such as ground control, mission and processing segments in the field of remote sensing and space surveillance areas.Currently, he is director at Indra, managing a large group of experts. He can offer you initiative, flexibility and open-mindedness attitude to generate new ideas and my great capability to discuss, analyse and reach consensus and agreement inside a working group.*

4.2 Third parties involved in the project (including use of third party resources)

In order to simplify the project administration at the Spanish level, the following partners participate under the form of JRU[2] lead by CSIC. In what follows we provide the description of each of those partners, together with the activities and budget foreseen in EOSC-synergy (included as CSIC budget in the global Form A).

[2] See https://digital.csic.es/handle/10261/171799

4.2.1 BSC

Barcelona Supercomputing Center is leading the integration and validation of services for the environmental community in WP4 (Task 4.5), and participates in the integration and validation of services for Biomedicine (Task 4.3).

WP4	PMs	Personnel Cost	Travel Budget	Direct Cost	OH (25%)	Total
BSC	24	108000	8000	116000	29000	145000

Description of the partner

The Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS), established in 2005, serves as the National Supercomputing Facility in Spain. The Center hosts MareNostrum, one of the most powerful supercomputers in Europe with more than 150,000 cores and more than 11PFlop of computing power and MinoTauro a cluster of GPU's composed of 2 partitions, one with 90 nodes, each with 2 Intel Nehalem processor (six cores each) and two NVIDIA M2090 processors, and 24 GB memory, and a second partition, with 39 nodes of 2 Intel Hashwell processors (8 cores each) and 2 K80 NVIDIA GPU cards, with 128GB of main memory. The mission of the BSC-CNS is to research, develop and manage information technologies with the larger goal of facilitating scientific innovation. It strives to attain these goals by focusing on technology and innovation in the areas of Computational Sciences, Life Sciences and Earth Sciences. The Computer Science Department of the BSC-CNS focuses on adapting both currently available and cutting edge hardware and software technologies to supercomputing infrastructures. The Earth Sciences Department carries out research in earth systems modeling, focusing on atmospheric physics and chemistry. The Life Sciences Department integrates a broad range of interconnected research areas in computational biology, from genomics to computational biochemistry. BSC is an active and well recognized participant of different initiatives for the integration and consolidation of High Performance Computing and data management in Europe and Spain. The BSC manages all the information generated by the HPC simulations executed in their computational resources. In order to fulfill the data requirements of scientists, BSC has more than 20 PB of available disk space and a tape library with up to 6PB. At the international level, BSC is currently working in more than 42 European projects. At infrastructure level, the main e-infrastructures projects where BSC is working are EUDAT2020, PRACE (where BSC is one of the tier-0 partners) and HBP, in all of these projects BSC is acting at the same time as a Data center and HPC center.

Key personnel

Dr. Nadia Tonello *[F], is head of BSC's Data Management group since 2018. She holds a PhD in Astrophysics from the Technical University of Munich (D) and a Big Data specialization. As a Postdoc she worked as scientific liaison in the Astrophysics and Cosmology department at Port d'Informació*

Científica (HTC data center, Barcelona) for two international projects: the PAU Survey and the ESA mission called Euclid. She was responsible of the management of both observed and simulated data archives, the parallelization and optimization of pipelines code, and the operations of data analysis and validation. She was the deputy leader of the Spanish Scientific Data Center of the Euclid Science Ground Segment and part of the image simulations unit (OU-SIM), coordinating the code integration activities and the validation of the products. At BSC, she is currently involved in projects of Open Data and Open Science such as RDA, EUDAT CDI, EOSCpilot and EOSC-hub.

Francesco Benincasa *[M], is software designer and developer. He holds a Master of Science in Software Engineering with a thesis on Optimization Algorithms. He is GNU/Linux and Python language expert and author of various software applications, web and stand-alone, for web communities, for data managing, automatic shopping, network data upload/download, data processing, formatting and visualizing. He started his experience in supercomputing and data manipulation (SCS SuperComputing Solutions, a CINECA spin-off - Bologna - Italy) oriented to biomedical simulations, participating to several FP6 and FP7 European projects. Since 2010 he is at the Barcelona Supercomputing Center (BSC) working on air quality data management, processing and visualization. Currently is co-leader of the Data and Diagnostics Team inside the Earth Sciences Department. He is in charge of data processing and management and web development and maintenance of the WMO SDS-WAS NA-ME-E Regional Center (SDS) and the Barcelona Dust Forecast Center (BDFC), both projects operated by a consortium of BSC and AEMET (Spanish Meteorological Agency) under the umbrella of the World Meteorological Organization (WMO). He is also chairman of the "weather, climate and air quality Interest Group" inside the Research Data Alliance (RDA).*

Prof. Alfonso Valencia *[M], PhD, joined the Barcelona Supercomputing Center (BSC) recently as ICREA Research Professor. Before, he was Vice Director of Basic Research and Director of the Structural Biology and Biocomputing Programme at the Spanish National Cancer Research Centre (CNIO). He is also the Director of the INB. His research is based on the use of computation for the analysis of large collection of genomic information with particular emphasis in the study of protein families and protein interaction networks. His recent work focuses in the area of cancer (epi)genomics, tumour evolution, precision medicine and the study of co-morbidity. He has published more than 420 peer-reviewed papers (Scopus: H-index of 80, >36,000 citations; Google Scholar: H-index of 102, 45,000 citations).*

Dr. Salvador Capella *[M], PhD, joined BSC as the team leader of the INB. His current focus is on the development of long-term infrastructures to facilitate the scientific benchmarking and technical monitoring of bioinformatics tools, web services and workflows in the context of ELIXIR. The team is also responsible for the technical coordination of ELIXIR-ES covering a broad range of bioinformatics areas. Author of over 25 peer-reviewed scientific publications in the field of bioinformatics (Scopus: H-index of 16, >3,000 citations; Google Scholar: H-index of 18, >4,000 citations).*

4.2.2 CESGA

CESGA	PMs	Personnel Cost	Travel Budget	Direct Cost	OH (25%)	Total
WP2	18	82800€	6000€	88800€	22200€	111000€
WP3	10	46000€	4000€	50000€	12500€	62500€
TOTALs	**28**	**128800€**	**10000€**	**138800€**	**34700€**	**173500€**

Fundación Pública Gallega Centro Tecnológico de Supercomputación de Galicia (CESGA) was founded in 1993 to promote and provide supercomputing facilities to regional and national researchers. Nowadays, CESGA employs more than 35 people (6 of them are PhDs). CESGA provides more than 250 TFlops of computing capacity, which is updated continuously to solve the demands of its users. For more than 23 years, CESGA has provided computing and communications to researchers, mainly from the three Galician Universities, regional research centres, and the Spanish National Research Council (CSIC). In fact, CESGA is certified by ISO 9001:2008 for the provision of services related to computing and communications. Also, since 2004, it started an action to promote and enhance the usage of HPC in the SMEs, mainly for the manufacturing sectors. Additionally, to this work with companies and researchers, CESGA runs the simulations for the Regional Weather Forecast Agency (Meteogalicia) and the IBI (Iberian-Biscay-Ireland) zone of the European MyOcean project for Puertos del Estado of Spain. These services, which run on the same supercomputers as the other batch services, have created a large knowledge in the institution related to operational HPC or HPC in real-time. Also, since 2003, CESGA is providing big-data infrastructure on demand based on Hadoop, where several experiences have been executed as the analysis of astronomical data for GAIA project (see https://www.cesga.es/gl/biblioteca/downloadAsset/id/750). Other services that are provided to research institutions and companies are technical-software optimization, parallelization, and development of graphical interfaces for technical computing.

CESGA has been participating in European research projects for more than 15 years, such as FORTISSIMO, BonFIRE, EGEE-1,2,3, Crossgrid, int.eu.grid, BEinGRID, or SmartLM, where they have developed activities mainly in development, integration, and testing of the software as well as deployment and management of computing infrastructure. Also, it has been participating actively in EGI-Inspire and EMI. In EGI.eu, it has the accounting responsibility for all the world distributed computing infrastructure.

Relevant Projects:

FORTISSMO and FORTISSIMO2 EU FP7 project, from 1st July 2014 EU H2020 project, from 1st November 2015: FORTISSMO is a FP7 European project devoted to the provision of HPC resources to the European Manufacturing SMEs. In this project, CESGA has developed the accounting package and has run one of the finished experiments in 2014. Currently it is running another set of experiments (which have accessed to the project through the open calls). It has hosted the portal to support the open calls of the project. These experiments are related to Modeling and Simulation, point-cloud analysis, map

conversions, etc. All of them are based on the usage of mathematical algorithms and in some of them have helped research and ISVs to parallelise the code. FORTISSIMO 2 is the second stage of FORTISSIMO. It is still focused in the application of HPC technologies to Manufacturing SMEs, but having more complexity (i.e., using coupled models or multi-physics) or are related to Big Data.

EGI-INSPIRE. European Grid Initiative. FP7 e-Infrastructures project. From 1st May 2010 to 31st December 2014. This project deployed and managed the largest Grid infrastructure of the world, which has made the first experience to move to Cloud environment in the FedCloud experiment. CESGA participated in this experiment and it has the responsibility of the accounting and metrics portal for all the world infrastructure.

PRACE-4IP. CESGA also participates in the PRACE initiative as third party, where has plans to integrate its newest supercomputer, FinisTerrae2, on the PRACE infrastructure as a Tier1 system.

Significant infrastructure and technical equipment

CESGA data centre has over 340 square meters for IT equipment. CESGA will provide to the project access to its IT infrastructures, more specifically, the HPC and BigData infrastructures. Current computing infrastructure is composed of:

Finisterrae. A HPC supercomputer with 306 nodes, 24 cores Haswell 2680v3 and 128GB of RAM. It is attached to a LUSTRE filesystem with more than 768TB of disk. The low latency network is Infiniband FDR@56Gbps. It includes additional nodes with special characteristics as GPU (8 NVIDIA K80), Xeon Phi (4) and one node wih 2TB RAM and 8 Intel Haswell 8867v3, 128 cores.

SVG Cluster is based on x86 servers dedicated to provide high level HTC, Grid and Cloud services to scientific and industrial customers. It is based on more than 1200 CPU-cores AMD Opteron and Intel Sandy-bridge and Haswell.

CESGA has a Cloud computing service for users which delivers a virtual infrastructure, configurable to the requirements of the final user: operating system, number of processors, memory, disk, and number of nodes are configured to user's needs in a dynamic form. The management of the system uses OpenNebula software.

Visualization. In 2015, new equipment for remote visualisation has been deployed. It has 4 nodes, each one in NVIDIA Grid GPU.

Big Data cluster is composed of 36 nodes with Intel Xeon E5-2620 v3 @ 2.40GHz and a very large disk capacity (24TB per node). It runs Centos 7.1 con Docker, Hadoop Distribution, GlusterFS as filesystem, ElasticSearch, Logstash and Kibana, Custom Cluster Deployment as Mesos and Marathon and Hadoop YARN integration with Mesos. It is devoted to Big Data reserach.

Key personnel

Dr. Carlos Fernández Sánchez *[M], graduated in Physics in 1995 at the University of Santiago de Compostela and in Computer Systems Engineering at the Universidad Nacional de Educación a*

Distancia in 2002. He holds a Ph.D in Physics from Universidad de Santiago de Compostela since 1999. Since 2002, he is the coordinator of the Systems Department at CESGA. He has participated in more than 30 European, national and regional projects related to HPC, Grid, Cloud and IT as: EGI-Inspire, EMI, or BonFIRE. He coordinated the working group of Service Oriented Infrastructures of INES (NESSI Spanish mirror). He will coordinate the provision of infrastructure and services.

***Dr. Ignacio López** [M], obtained his B.S. in Physics from the Universidad Complutense de Madrid and his Ph.D. in Physics from the University of Santiago de Compostela. He worked as a senior system engineer for Fujitsu Spain for more than 5 years. In 1998 he moved to his actual position at CESGA, where is in charge of all the technical activities of the center. He has participated in several regional, national and European R&D projects related to research infrastructures, Grid or Cloud computing and network technologies as Crossgrid, TORGA.net, EGEE, EGEE2, EGEE3, Int.eu.grid, EGI-Inspire, EMI or Retelab, and SME oriented like Cloudpyme, Cloudpyme and FORTISSMO. He will have the responsibility of coordinating the provision of computing infrastructure and communications.*

4.2.3 BIFI

BIFI	PMs	Personnel Cost	Travel Budget	Direct Cost	OH (25%)	Total
WP2	12	48000€	6000€	54000€	13500€	67500€
WP6	15	60000€	9000€	69000€	17250€	86250€
TOTALs	**27**	**108000€**	**15000€**	**123000€**	**30750€**	**153750€**

BIFI is a research institute that promotes interdisciplinarity to face the big challenges of innovation and development. BIFI has a strong relationship with digital science and international projects at several levels including data analysis, scientific applications development and porting, user support and resource provision. BIFI is a reference in e-infrastructures (Grid, Cloud and HPC), participating in several EU-FP7 projects and hosting one of the nodes of the Supercomputing Spanish Network. Researchers from BIFI also constructed the awarded Dedicated Supercomputer called JANUS. BIFI is also a reference in Citizen Science projects and initiatives. It is leading the Spanish Ibercivis Foundation and is coordinating the Socientize project which is creating the White Book on Citizen Science for Europe.

BIFI is partner of many European projects related to industrial innovation like CloudSME, Fortissimo or CloudFlow from the Factures of the Future (FoF) Public-Private Partnership (PPP) and has also participated in SCI-BUS, Global Excursion, SCC-Computing, EGI, EGEE II, EGEE III int.eu.grid, EDGeS, EDGI and DEGISCO. BIFI also was leader in two regional Grid initiatives, Aragrid and PireGrid.

Key personnel

Prof. Dr. Alfonso Tarancón Lafita *[M], is the Director of BIFI. He obtained the Physics Science degree in the University of Zaragoza in 1982 with Extraordinary Award and did his doctoral thesis in the study of Quantum Theory regularised on the Lattice. He joined the APE Group (Universita de Roma I La Sapienza and INFN) under the direction of Nicola Cabibbo y Giorgio Parisi. He is responsible of one of the nodes of the Supercomputation Spanish Network (RES) and he is Director of the Aragrid and Piregrid projects among others. Along his career, he has published more than 100 scientific articles in international journals, has directed 9 PhD, many technological transfer projects and has been leading research in more than 20 R&D projects.*

Dr David Iñiguez *[M], holds a PhD in Physics from the University of Zaragoza, and an ARAID position at BIFI. His main topics of research include the development of parallel and dedicated computers and several branches of complex systems physics (e.g. field theories, spin glasses, fracture models). Nowadays, he is also focused in the application of this kind of physics models and multicore computing to business problems.*

4.2.4 CIEMAT

CIEMAT	PMs	Personnel Cost	Travel Budget	Direct Cost	OH (25%)	Total
WP2	6	24360€	7000€	31360€	7840€	39200€
WP4	24	97440€	5000€	102440€	25610€	128050€
TOTALs	**30**	**121800€**	**12000€**	**133800€**	**33450€**	**167250€**

CIEMAT is a Spanish Public Research Institution (www.ciemat.es). Since its formation in 1951, it has developed and led R&D projects in the fields of Energy, Environment and Technology, placing the institution at the forefront of science and technology. As a technological research centre, CIEMAT fosters links between academia and industry. Its main activities include: to promote the introduction and improve the competitiveness of renewable energies on the energy market; to improve the efficiency and environmental quality of fossil fuels; to optimise waste management and safety of nuclear fission energy; to demonstrate the role of nuclear fusion as a future energy alternative; to assess the environmental impact of energy; to promote the development of environmentally-friendly technologies that respect the environment; to transfer horizontal technologies to industry; to improve scientific returns derived from CIEMAT activities; to strengthen industrial participation in international projects with high technological content; and, to foster technology transfer, training and scientific outreach.

CIEMAT is the Spanish major R&D centre on energy. Related to HPC/HTC, it counts on supercomputing facilities since the sixties of the last century, being pioneer in Spain. It has participated in 58 projects co-

funded by H2020 and has a strong expertise in FP6, FP7, and H2020 projects on distributed computing infrastructures (EGI series, EELA series, CHAIN series. etc.) where it has held managerial responsibilities.

In addition to the horizontal WPs related to Management and Dissemination and Outreach, CIEMAT will participate in those related to the use cases integration with the LAGO experiment and to the computing infrastructure provisioning with almost 1,000 cores and ~3.5 TB

Key personnel

Dr. Rafael Mayo García *[M], is Senior Researcher at CIEMAT and Harvard University Fellow. He earned his PhD in Physics from the Universidad Complutense de Madrid (2004). From 2006 he has also been Adjunct Faculty and Honorary Fellow at the same University in the Physics of Materials Department. He has been involved in many experiments in the US, Bulgaria, Sweden and Ireland (funded, among others, by the European Commission with a Marie Curie fellowship). He has also obtained a postdoctoral fellowship in the Spanish Juan de la Cierva Programme. He is author of 130 scientific articles. He has been project coordinator of 4 Spanish and 1 international R&D IT initiatives and has been involved in several European and National projects working on HPC scientific developments and even on managerial activities as Work Package Manager and/or member of Executive Boards. He sums up to participation in 50 projects. He also has served to several institutions as evaluator for their competitive Calls, European Commission included, and has supervised 2 PhD theses.*

Guillermo Díaz *[M], received his M.Sc. degree in Telecommunication Engineering (2001) from the Universidad Politécnica de Madrid (UPM). Guillermo is also MBA by ITAE Business School of Extremadura (2011), and has a post-graduate degree in Home & Building Automation Systems (2006). He worked during 10 years in multiple ICT-based projects for the Spanish industry, with successful deployments in Japan, USA and Mexico. After that period, he joined CIEMAT in 2009 as technical coordinator of the Advanced Computing Technologies Centre of Extremadura (CETA), taking over his current position as Head of CETA in 2011. Since then, he has gained experience with best practices in the management of computing e-infrastructures, and in the use of European Regional Development Funds (ERDF). During the last ten years he has participated in multiple FP7 and H2020 European projects, within activities concerning management of e-infrastructures and provisioning of advanced computational services to research communities. Likewise, in the field of technology transfer activities, he has participated in projects of the Spanish INNPACTO call and R&D&I regional programmes, as well as various contracts and collaboration agreements with private companies in multiple areas of the ICT, energy and agricultural industries. He now belongs to the group of experts in ICT, evaluating and assessing the regional government in the implementation of the RIS3 (Research and Innovation Smart Specialisation Strategy) for Extremadura, committed to the European Commission.*

4.2.5 RedIRIS

RedIRIS, with more than 500 affiliated institutions, notably including Spanish universities and Public Research Organisations, is the national network of advanced communications for the academic and scientific community. From an organisational perspective, RedIRIS is a scientific infrastructure, under the Ministry of Economy, Industry and Competitiveness (MEIC), which sets its strategy and finances its operations. Since 2004, its technical and operational management has been entrusted to the public entity Red.es, under the Ministry of Energy, Tourism and the Digital Agenda. RedIRIS commenced its activity in 1988 and has pioneered the introduction of the Internet and various telematic services in Spain.

Since its creation, it has been a fundamental support element for carrying out collaborative projects, above all in the e-Science sphere. The network's pivotal role has been acknowledged with inclusion in the listing of extraordinary scientific and technical installations, Unique Scientific and Technical Infrastructures (ICTS), which groups together the most important scientific infrastructure in Spain. In order to provide users with the best possible connectivity and to keep abreast of the latest telematic services, RedIRIS cooperates with other academic and research networks, both regional and international. Of particular note is the panEuropean academic network GÉANT, which is jointly managed by RedIRIS and through which it connects with other academic and research networks both in Spain and in other European countries, as well as with research networks on other continents, such as Internet2 (USA), RedCLARA (Latin America) or EUMEDCONNECT (North Africa).

Key personnel:

Antonio Fuentes Bermejo [M], *Head of Systems and Security at RedIRIS, qualified as a Computing Engineer (University of Murcia) in 1999, and he is now working towards his PhD at the Computer Architecture and Technology Department of the University of Murcia. Antonio has previously worked as a system engineer to Airtel/Vodafone (a Spanish telecommunication company) and at the University of Murcia. He has been the coordinator of the infrastructural services of the Academic Grid Initiative, IRISGrid and he was working in a number of grid-related and Fiware European Projects. His research works are oriented to cloud services and file transfer protocols.*

Dr. Alberto Perez [M] *works since 2.002, for Red.es, a Spanish governmental agency in charge of promoting digital transformation in Spain and which, inter alia, manages RedIRIS (the Spanish National Research and Education Network, or "NREN"), and Dominios.es (the Registry for the top-level domain name ".es"). In 2017, Dr. Pérez was appointed Director of RedIRIS, the department of Red.es which provides very high-speed broadband (multiple links of 10 Gbps) to 500 Spanish education and research centers (with 5 M end-users). RedIRIS also provides other related ICT services, which facilitate remote collaboration among those centers, at national and international level.*

RedIRIS, with a headcount of 20 staff members, and a yearly budget of 8 M€, manages an optic fibre backbone of 14.000 Km, called RedIRIS-NOVA, which will be in operation for more than 20 years, and, which had a cost of 105 M€ (approx. 50% covered by ERDF). RedIRIS is now running a competitive dialogue of 23 M€ to replace its optical equipment. Dr. Pérez also collaborates with the ".es" Registry (also managed by Red.es) in issues related to international relations and policy development. Before joining Red.es, Dr Pérez worked for the Spanish telecom regulator (1.999-2.002, and was a lecturer at the University of Alcalá (1994-1999).

***Esther Robles** [F] born in Valladolid (Spain) in 1973, is Deputy Director of Communications Networks of RedIRIS since December 2017. She joint RedIRIS in 1998 as Network Engineer and was appointed as Head of RedIRIS network area in 2000, leading the RedIRIS network evolution, both in technology and network services, and managing the network operation team responsible for planning and operating the network infrastructure and services. However, her major achievement in terms of network expansion and evolution was to lead the strategical project to establish an own dark fiber network (14.000 km of optical fiber, national coverage including subsea cables to Canary and Balearic islands and Melilla.) with the latest state-of-the-art WDM optical technology equipment.*

In July 1997 received her M.Sc degree in Computer Science by the Valladolid University.

From 1997 to 1998, completed a postgrade research stay at the Technische Universität Dresden working in ATM technology and advanced reservation mechanisms

5. Ethics and Security

The EOSC-Synergy project does not raise any ethical concerns as listed in the table part of the proposal submission administrative forms. The project does not involve activities or results raising security issues, or any 'EU-classified information' as background or results.

6. Letters of Support

See attached PDFs.

Dr. Isabel Campos
CSIC - Spanish National Research Council

Letter of Support
EOSC-synergy, H2020 INFRAEOSC-05 B

Dear Isabel

EOSC-hub is one of the flagship EC projects contributing to the European Open Science Cloud implementation. It mobilises providers from the EGI Federation, EUDAT CDI, INDIGO- DataCloud and major research e-infrastructures offering services, software and data for advanced data-driven research and innovation. Access to these services are provided by the integration and management system of the European Open Science Cloud, acting as a single entry point for all stakeholders. The EOSC-hub service catalogue is open to new contributing communities and business organisations.

On behalf of EOSC-hub, I am pleased to offer our wholehearted support for the EOSC-synergy project proposal.

EOSC-hub is willing to promote the integration and exploitation of the EOSC-synergy services and resources for adoption by a large international group of users in EOSC, through the EOSC Portal Marketplace, EOSC-hub will offer you a mechanism to promote access and re-use of your key exploitable to EOSC stakeholders.

We are interested in building a joint service and resource portfolio that includes EOSC-synergy national services and certified data repositories that are quality verified and assessed by your international research collaborations of pan-European relevance. In addition, we are interested in collaborating to the definition of a joint service catalogue roadmap that takes into account users and providers' needs from the countries in the EOSC-synergy scope.

EOSC-hub is interested in setting up and implementing a joint events and communications programme aiming at harmonizing and strengthening the projects' innovation management plans. EOSC-hub will promote EOSC-synergy services and research data for both research and commercial exploitation via the EOSC Digital Innovation Hub, the EOSC stakeholder forum and other relevant engagement channels. We regard the EOSC-synergy human capacity development programme as an important EOSC pillar: EOSC-hub is committed to provide services and resources for training.

The EOSC-synergy policy recommendations will further contribute to the definition of EOSC suitable procurement and service provisioning and models that are necessary components of a sustainable EOSC ecosystem of services and resources.

Finally the two projects will deliver coordinated technical support to user communities. The projects will be responsible in their respective service provisioning areas for technical support in the thematic and competency areas in scope in the respective descriptions of work

I strongly believe the collaboration between the two projects will be of great benefit to the advancement of the implementation of the European Open Science Cloud initiative.

Amsterdam, 15 November 2018
Yours sincerely,

Dr. Tiziana Ferrari
EOSC-hub Project Coordinator

The Latin American Giant Observatory (LAGO)

To: EOSC-SINERGY Project Coordinator

Dear Prof. Blanquer,

On behalf of the Latin American Giant Observatory (LAGO, http://lagoproject.org/) initiative, I hereby confirm our strong involvement in the EOSC-SINERGY proposal to be submitted to the INFRAEOSC-05-2018-2019 Call throughout our Spanish representative Centro de Investigaciones Energéticas y Medioambeintales (CIEMAT).

LAGO is an extended cosmic ray observatory composed of a network of water-Cherenkov detectors (WCD) spanning over different sites located at significantly different altitudes (from sea level up to more than 5000 m a.s.l.) and latitudes across Latin America, covering a wide range of geomagnetic rigidity cut-offs and atmospheric absorption/reaction levels. It is an observatory designed, built, and operated by the LAGO Collaboration, a non-centralized collaborative union of more than 30 institutions from ten countries.

LAGO WCDs are simple and robust, and incorporates several integrated devices to allow time synchronization, autonomous operation, on board data analysis, as well as remote control and automated data transfer. This detection network is designed to make detailed measurements of the temporal evolution of the radiation flux coming from outer space at ground level. LAGO is mainly oriented to perform basic research in three areas: high energy phenomena and the extreme universe, space weather and atmospheric radiation at ground level. It is also producing interesting results in determining radiation fluxes received by airplane cabin crew or volcanology forecasting from muology determination.

LAGO use case perfectly matches the requirements needed to be exploited on an infrastructure as the EOSC one is as complies with the EOSC data management standards and simulations can be seamlessly run by mature and well tested applications on distributed virtualized platforms (CORSIKA; GEANT4, etc.). In this way, the codes used to perform the simulations can be integrated into the EOSC portal and the LAGO measured data can be harvested from the EOSC services (prior similar experience can be derived from the EC Infrastructure co-funded project CHIAN-REDS where LAGO was presented as success story providing accessibility, reproducibility, and trustworthiness).

Data coming from both LAGO measurements and simulations will be directly provided by the LAGO consortium, i.e. data is host elsewhere by the community and researchers will perform the simulations using the EOSC services.

Kind regards,

Dr. Iván P. Sidelnik
LAGO Principal Investigator