D10.7 Final technical architecture update and interoperability guidelines update.

Lead Partner:	INFN
Version:	1
Status:	FINAL
Dissemination Level:	Public
Document Link:	https://documents.egi.eu/document/3708

Deliverable Abstract
This document describes how the architecture and the interoperability guidelines developed within EOSC-hub have been updated in accordance with the evolution of the project activities and the feedback collected among various stakeholders, including user communities. It also provides information about their future evolution after the end of the project.

COPYRIGHT NOTICE

DELIVERY SLIP

Date	Name	Partner/Activity	Date
From:	Alessandro Costantini Giacinto Donvito	INFN/WP10 INFN/WP10	16/04/2021
Moderated by:	Malgorzata Krakowian	EGI Foundation/WP1	
Reviewed by:	Enol Fernandez	EGI Foundation	16/04/2021
Approved by:	AMB		

DOCUMENT LOG

Issue	Date	Comment	Author
v.0.1	30/03/2021	ToC ready	Alessandro Costantini Giacinto Donvito Diego Scardaci
v.0.2	15/04/2021	Suggestion addressed	Alessandro Costantini Giacinto Donvito Diego Scardaci Luciano Gaido Tiziana Ferrari
v.0.3	15/04/2021	Document ready for review	Alessandro Costantini Giacinto Donvito Diego Scardaci
v.0.4	16/04/2021	Document reviewed	Enol Fernandez
v.0.5	16/04/2021	Suggestion from reviewer addressed by the authors	Alessandro Costantini Giacinto Donvito
v.1	16/04/2021	Final	Alessandro Costantini Giacinto Donvito Diego Scardaci

TERMINOLOGY

https://wiki.eosc-hub.eu/display/EOSC/EOSC-hub+Glossary

Contents

Executive summary

EOSC-hub worked since its start to foster the EOSC interoperability and facilitate the integration of EOSC services with the aim to create an integrated environment where researchers can easily find, combine and execute services for data access, sharing and analysis.

In this context, the EOSC-hub Technical Committee (TCOM) defined an overall framework to describe the EOSC services (in terms of architecture, protocols, APIs, etc.) and their interoperability capabilities. Such a framework is based on the concept of "building block", a technical function that can be offered by one or more services. A building block is defined through a technical specification that includes a high-level architecture, suggested EOSC standards and APIs, and interoperability guidelines. 18 different architecture and interoperability guidelines for EOSC federation/core services and horizontal/common services were developed in the context of this framework covering several technical areas (AAI, accounting and monitoring, Cloud IaaS, Cloud Orchestration, Data repository, etc.).

The "building block" approach has demonstrated to be successful in promoting a flexible and dynamic approach for integrating new services in the EOSC environment as proved by the several technical integrations achieved in the project. For example, the EOSC-hub thematic services achieved more than 40 technical integrations with 19 different services from e-infrastructures.

When the EOSC WGs were launched, the EOSC-hub TCOM widened its scope by participating in the activities of the EOSC Architecture WG with its own representative and brought its results to the wider EOSC environment. Collaboration with the EOSC Architecture WG enabled the creation of a virtual cycle that, from one side, allowed the EOSC Architecture WG to benefit from the main EOSC-hub outputs (interoperability guidelines, federation services to establish the core, work on AAI, PID, etc.) and, from the other side, EOSC-hub to align its points of view on the EOSC technical aspects to the wider vision provided by the Architecture WG. This collaboration brought two relevant results:

- The EOSC Interoperability Framework, developed jointly by the Architecture and FAIR WGs, adopted the building block concept initially conceived in EOSC-hub for technical interoperability;
- EOSC-hub mapped its services to the different components identified in the overall EOSC Architecture developed by the Architecture WG.

The architecture and interoperability guidelines produced by the project were also updated considering the suggestions from many relevant EOSC stakeholders collected through two campaigns to gather feedback launched in the EOSC Liaison Platform[1].

The work on the architecture and interoperability guidelines will be followed-up in the EOSC Future project that is in charge to harmonise and integrate all the past efforts in interoperability by different EOSC projects. EOSC-hub will provide a very relevant contribution to this effort hand overing the guidelines for services of the EOSC Core (AAI, accounting, helpdesk, monitoring, etc.) and for key technical areas of the EOSC Exchange (Cloud/HTC/HPC computing environments, data management systems, analytics, ML/AI, etc.). INFRAEOSC-07 projects will also contribute with

---

[1] https://www.eoscsecretariat.eu/eosc-liaison-platform

interoperability guidelines for services of their respective technical areas and will deliver real instances of services compliant with the EOSC interoperability guidelines. This makes evident how the EOSC-hub effort on interoperability of services has been effective in fostering open science and addressing real users' needs.

1 Introduction

The document summarises the activities carried out by the EOSC-hub project in the context of EOSC Architecture and interoperability framework.

The document is organised as follows:

- Section 2 highlights the EOSC-hub contribution to the development of the EOSC Architecture and of the interoperability framework;
- Section 3 presents the feedback received from the stakeholders on the architecture and interoperability guidelines developed by the project;
- Section 4 contains the description of the updated architecture and interoperability guidelines;
- Section 5 describes how the EOSC-hub work on the architecture and guidelines will be taken over and further enhanced by follow-up projects like EOSC Future, EGI-ACE, DICE, and other initiatives such as the EOSC Association Advisory groups.

2   EOSC-hub contribution to the EOSC architecture and interoperability framework

One of the main aims of EOSC is fostering and facilitating the collaboration between operators of research infrastructures and e-infrastructures. This can be achieved developing an integrated environment where researchers can easily find, combine and execute services for data access, sharing and analysis. In particular, this integrated environment should enable the easy composition of services to facilitate the creation of new added-value solutions for researchers allowing them to deal with the increasing complexity of the science problems.

For this reason, EOSC-hub worked since its start to foster the EOSC interoperability and facilitate the integration of EOSC services. The technical committee of the project (TCOM) designed a reference EOSC technical architecture with the aim of defining a framework to describe the EOSC services (in terms of architecture, protocols, APIs, etc) and their interoperability capabilities. Contextually, the TCOM developed the first interoperability guidelines for EOSC federation/core services and horizontal/common services. The overall framework and the interoperability guidelines developed by the TCOM were presented in the D10.4 'EOSC Hub Technical Architecture and standards roadmap v2'[2].

The EOSC technical architecture described in the D10.4 is based on the concept of "building block", a technical function that can be offered by one or more services. A building block is defined through a technical specification that includes a high-level architecture, suggested EOSC standards and APIs and interoperability guidelines. The "building block" approach has demonstrated to be successful in promoting a flexible and dynamic approach in integrating new services in the EOSC environment as proved by the several technical integrations achieved in the project[3]. For example, the EOSC-hub thematic services achieved more than 40 technical integrations with 19 different services from e-infrastructures (EGI, EUDAT and INDIGO). Technical integrations in EOSC-hub were also achieved thanks to the TCOM members that supported user communities willing to integrate their services in EOSC with their expertise, identifying the EOSC-hub services the communities could benefit from and suggesting how to perform the integrations according to the interoperability guidelines.

At the end of the project, 18 interoperability guidelines[4] have been released relying on the building block concepts. 5 of them are related to EOSC Core/federation services and allowed the project to enable the first beta EOSC Core. These guidelines for EOSC Core services also detail different integration scenarios that each EOSC provider can choose to integrate its services into EOSC. Two campaigns to gather feedback were launched through the EOSC Liaison Platform[5] to collect improvement suggestions from many relevant EOSC stakeholders (CESSDA, European Spallation Source ERIC, DESY, INRIA, GEANT Association, RENAM Association, etc.), more details are available

---

[2] https://www.eosc-hub.eu/deliverable/d103-technical-architecture-and-standards-roadmap

[3] EOSC-hub integration report: https://eosc-hub.eu/sites/default/files/EOSC-hub%20Integration%20Report.pdf

[4] https://www.eosc-hub.eu/technical-documentation#overlay-context=

[5] https://www.eoscsecretariat.eu/eosc-liaison-platform

in section 4. New versions of the guidelines were produced taking into account the collected feedback.

The EOSC technical architecture framework proposed by EOSC-hub was widely presented in conferences and webinars (e.g. EOSC Symposium 2019) and it was the base for interacting with EOSC architecture WG.

EOSC-hub also actively participated in the work of the EOSC Architecture WG with its own representative to propose the adoption of its interoperability framework in the wider EOSC environment. Results achieved in the project were shared with the WG and the methodology to develop interoperability guidelines based on the building block concept was recommended for inclusion in the EOSC Interoperability Framework (EOSC IF). For this aim, EOSC-hub also participated in the task force on the interoperability jointly launched by the EOSC FAIR and Architecture WGs that released the first version of the EOSC Interoperability Framework document. In addition to the direct involvement in this task force, the project also contributed to the feedback gathering campaign launched by the task force during the summer 2020 providing TCOM suggestions to improve the initial draft of the document.

As a result, the final version of the EOSC IF document[6] adopted the building block concept initially conceived in EOSC-hub for technical interoperability, this is described in the Section 5 of the document "TOWARDS THE EOSC IF: REFERENCE ARCHITECTURE": *The Reference Architecture contains framework definitions and uses abstract Building Blocks as a tool to group functionality that will be needed to meet the requirements for the EOSC Interoperability Framework…* Details on the feedback provided by the EOSC-hub TCOM to the first draft of the EOSC EIF document are available in the following section.

Furthermore, EOSC-hub contributed to the effort of the EOSC Architecture WG to design an overall EOSC Architecture bringing the large expertise of the 3 main initiatives behind the project (EGI, EUDAT and INDIGO-DataCloud) on operating and developing e-infrastructure services to serve research communities. The interaction with the EOSC Architecture WG was twofold, from one side, the EOSC-Hub representative regularly reported updates of the EOSC-Hub activities into the WG, and for the other side, they provided update information about the activities and results of the WG to the project management boards (AMB and TCOM). In this manner, a virtuous cycle was established that allowed:

- the EOSC Architecture WG to benefit from the main EOSC-hub outputs (interoperability guidelines, federation services to establish the core, work on AAI, PID, etc.);
- EOSC-hub to align its points of view on the EOSC technical aspects to the wider vision provided by the Architecture WG. Examples of this are the effort done by the project to map its services to the EOSC Architecture defined by the WG and the adoption of a new terminology to classify EOSC services to be coherent with the documents produced by the EOSC WGs (e.g. classify services as components of EOSC Core and EOSC Exchange instead of federation and common services).

---

[6] EOSC Interoperability Framework document released by the EOSC Architecture and FAIR WG:
https://op.europa.eu/nl/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1

EOSC-hub contributed to the EOSC Architecture Working Group View on the Minimum Viable EOSC (MVE) where the MVE is defined as a dynamic set of EOSC resources which will evolve over time, consisting of:

- The subset of EOSC resources necessary for forming the added-value and opportunities considered essential to be provided by the EOSC at a given moment in time, i.e., to allow essential services and research products (e.g., publications, datasets, software) to be discovered, composed, accessed and analysed via the EOSC, which could not be otherwise;
- The subset of EOSC Core components/services required to operate and deliver such resources.

The document shows broad agreement with the EOSC-hub proposals on the Federating Core and confirms that many of the functions of the MVE, and of the EOSC Core in particular, will be provided by EOSC-hub services.

A graphical representation of the MVE and its components is shown in Figure 1.
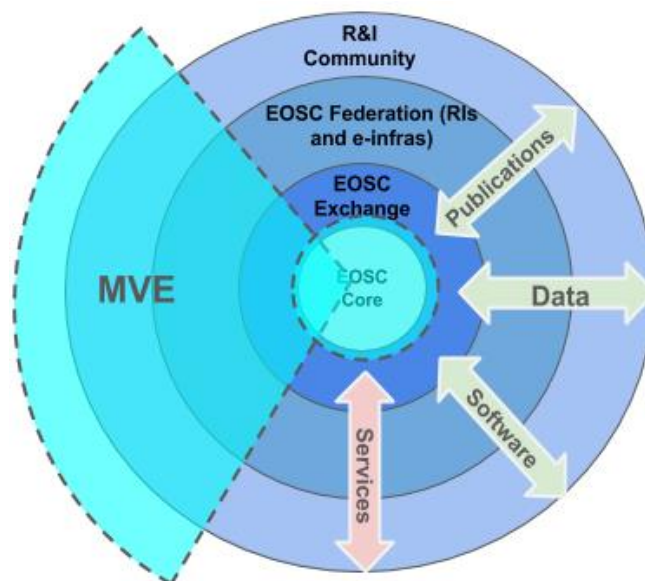
Figure 1: Architecture Working Group's High-level Diagram of the EOSC Depicting the Relationship Between EOSC Core, EOSC-Exchange, EOSC-Federation and the MVE

The EOSC Architecture designed by the EOSC Architecture WG is presented in more detail later in subsection 2.2 and the mapping of the EOSC-hub services to the MVE components is described in subsection 2.3.

2.1 EOSC Interoperability Framework: feedback from EOSC-hub TCOM experts

In the second half of 2020, members of the FAIR and Architecture Working Groups developed the initial draft of the EOSC Interoperability Framework. This draft was aimed at identifying general interoperability principles and organising them into the four layers: technical, semantic, organisational and legal. In such respect, the EOSC-hub TCOM experts provided their comments and recommendations to the document with the aim to improve it.

The comments and recommendations have been provided by answering to the following questions:

- Q1 Is this what you expected to see or are some things missing?
- Q2 Are the concepts clear or do some aspects need further clarification?
- Q3 Are the minimum requirements and recommendations appropriate?
- Q4 Is it clear who is responsible for what and how this should be followed?
- Q5 As a service provider, could you conform with / implement the framework?
- Q6 Is the model for FAIR Digital Objects sound?

All the comments and recommendations collected by the EOSC-hub TCOM experts have been discussed during the periodic TCOM meeting and have been collected and published in the EOSC Liaison Platform (EOSC Interoperability Framework Out for comment by TCOM[7]) presented to the editors.

A short summary of the comments and recommendations is given in Table 1.

Table 1: The most significant comments and recommendations on the first draft of the EOSC IF document provided by the EOSC-hub TCOM experts.

	Comment	*Recommendation*
Q1	The current Interoperability Framework (IF) seems to omit plans for the promotion of recommended standards and guidelines for their adoption in the implementation roadmap of the EOSC	We recommend the extension of the IF to integrate and define strategic actions for the increased adoption of the IF by relevant stakeholders
Q1	The current document does not outline how the IF will evolve in the future	Add an innovation roadmap, that considers existing guidelines and best practices emerging from EOSC projects such as the EOSC-hub interoperability guidelines[8], the OpenAIRE guidelines[9] and the FAIRSFAIR FAIR Data Policies and Practices[10], making sure

---

[7] https://www.eoscsecretariat.eu/eosc-liaison-platform/post/eosc-interoperability-framework-out-comment

[8] https://www.eosc-hub.eu/technical-documentation#overlay-context=

[9] https://guidelines.openaire.eu/en/latest/

[10] https://www.fairsfair.eu/fairsfair-open-consultation-fair-data-policies-and-practices

		that EOSC is leveraging as much as possible community good practices and recommendations
Q2	Section 4 discusses semantic interoperability in a quite comprehensive manner. However, this unfortunately highlights the relative lack of details of other aspects of interoperability (technical, organisational and legal) in the proposed model	Provide more details on how IF enables technical, organisational and legal interoperability
Q3	We think that the application interoperability is not fully sufficient to satisfy user needs within EOSC - it would need to be complemented by resource interoperability. In this way, research communities could profit from EOSC by accessing multiple resources from multiple providers (e.g. scaling up their setup using computing resources from more data centers). Currently, they have to use different interfaces to access different providers	EOSC should promote the adoption of standards to achieve interoperability also at resource level.
Q3	Data collections are - as described in the related section of the IF - usually poorly documented, in terms of the metadata that is made available for them. However, not all of the issues can be resolved by only improving metadata documentation	For the real interoperability also the data documentation with context and content description is needed (for example codebooks etc.). The context is crucial for the data products and research outputs
Q3	As presented in the document, organisational interoperability looks to be only/mainly for the EOSC organisation, not between EOSC and the community. There is mentioned "need of interoperability certification mechanisms for service providers" which is a good suggestion. IDS (International Data Spaces) already have this kind of mechanisms that are already tested with service providers	It would be desirable implementing a co-operating model with IDS
Q3	On developing the recommendations on legal interoperability for the next version of the document, data ownership has to be taken properly into account. Indeed all	Data ownership has to be taken properly into account

	data is not owned by organisations but also by individual researchers	
Q4	In our opinion, as a next step, the broad involvement and engagement of research communities and other EOSC initiatives is needed to find a consensus and deliver a widely accepted EOSC interoperability model	Setup an interest group with open participation to enhance the model with input from all the main EOSC stakeholders and deliver a widely agreed interoperability model
Q5	There are still too many undefined variables to state "[as a service provider, I will conform with the framework]". Prerequisites to implement the framework is that it will be widely accepted and the costs to implement it is reasonable	Find consensus on the proposed model and keep low the cost for service providers to be compliant
Q6	The document presents the FAIR Digital Object (FDO) (section 4) as the adopted design solution to implement the EOSC IF, without analysis and comparison with the state of the art	It would be useful presenting an analysis of the pros/cons of the FDO model in the context of its alternatives
Q6	It is not fully clear in the document what is the current status of the FDO model (e.g. availability of tools implementing the model, degree of adoption, long-term sustainability, etc.)	It would be desirable that current status and target status are described clearly and separately. Now it is not clear what we already have in the FDO context, what is missing from the target and what are the major steps to go from the current states to the target (to reach the state where the interoperability is functioning)

Some of the proposed comments and recommendations have been addressed by the editorial committee. In particular, among the different comments and recommendations, the editorial committee decided to adopt the building block concept provided by EOSC-hub for technical interoperability (see previous section) and integrate it in the latest published version of the EOSC Interoperability Framework.

The EOSC Interoperability framework[11] builds on a subset of frameworks aligning and supporting digital infrastructure solutions that act as enablers within EOSC and EOSC core capabilities. The different components within the frameworks also align and support development of other services and solutions. Each framework included within the EOSC Interoperability Framework is composed of a set of components targeting a specific topic or element within the framework. The EOSC

---

[11] https://op.europa.eu/nl/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1

Interoperability Framework and the individual frameworks included must be flexible in nature to meet the evolving needs of EOSC in a way that gives the highest benefit to the research community.

2.2   Introduction to the EOSC Architecture

Alongside the work of the Sustainability working group, which looks at EOSC Core from the point of view of capabilities and frameworks[12], EOSC Architecture Working Group has taken a more technical approach, considering the functional components necessary to provide a Core of EOSC. These were prepared as part of the Working Group's internal work and have subsequently been shared with other Working Groups and community members as a way to express the likely contents of EOSC Core, and therefore be able to explain plans for a Minimum Viable EOSC as an initial subset of EOSC Core. As a result of this effort, the EOSC Architecture Working Group designed an overall EOSC Architecture diagram[13] that provides a function overview of EOSC. As shown in Figure 2.
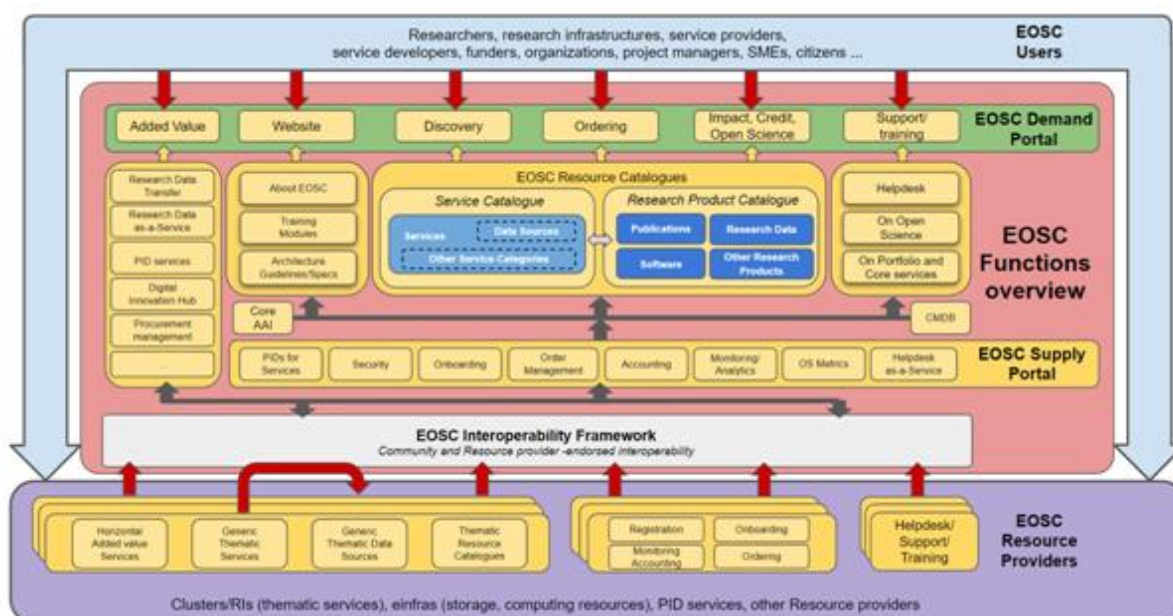
Figure 2 - EOSC architecture. The Red box comprises EOSC Core.

In particular, this diagram highlights the internal architecture of the EOSC Core (the big red box) and its interactions with external components including services in the EOSC Exchange (the purple box at the bottom).

As from the diagram, the heart of EOSC is a set of resource catalogues which include services and research products available in EOSC. On the demand side, those services and products are exposed to EOSC users (researchers, etc.) in a demand-facing portal that includes several functions like discovery, ordering, support, training, etc. EOSC services can be also accessed through other

---

[12] Solutions for a sustainable EOSC: https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1/language-en/format-PDF/source-175468053

[13] EOSC architecture working group view on the minimum viable EOSC: https://op.europa.eu/nl/publication-detail/-/publication/91fc0324-6b50-11eb-aeb5-01aa75ed71a1

channels beyond the EOSC Portal as highlighted in the diagram by the 2 blue arrows on the left and on the right.

The supply side takes care of the interactions between resource providers and EOSC. This is done by using a supply portal which is based on the guidance, formats and standards of the EOSC Interoperability Framework. Through this, providers contribute generic and thematic services via an onboarding mechanism. Such services can be integrated with several different elements like helpdesk, monitoring, accounting, AAI. The EOSC Interoperability Framework also supports the integration and composability of resources across providers of horizontal and/or thematic services.

EOSC-hub contributed to the establishment of the initial beta EOSC Core with several services that include the EOSC Portal AAI, the helpdesk, the monitoring system, the CMDB and many others (see D5.6 Final report on the integration of federation and collaboration services[14] for more details). This beta EOSC Core is currently operated according to the EOSC SMS developed by the project and has been hand overed to the follow-up project EOSC Future. EOSC-hub is also contributing to the EOSC Architecture with several EOSC Exchange services (both horizontal and thematic). More details about the mapping of the EOSC-hub services to the overall EOSC architecture are available in the next section.

2.3   EOSC architecture: mapping the EOSC-hub services

As stated in the previous sections, EOSC-hub is already providing a number of services which can fit into the EOSC Core and EOSC Exchange categories (and this is also true of some other EC-funded projects). The consolidation of the EOSC-hub services in the EOSC Core is expected to happen in the next couple of years in the context of the EOSC Association and the next EOSC-related projects. EOSC-hub services in the EOSC Exchange are already used by thousands of researchers as demonstrated by the EOSC-hub virtual access metrics[15].

Mapping the EOSC-hub services into the EOSC Core and EOSC Exchange categories gives a useful view of the contribution of the EOSC-hub activities on establishing the EOSC. Information on how and where these services are supported after the end of the project is also important for paving the way to what may happen in the frame of the new Horizon Europe programme (see Section 5).

In such respect, Table 2 (see EOSC-hub D2.5 Final Governance and Sustainability implementation roadmap[16] for more details) shows this mapping and shows which Horizon 2020 projects will sustain them until 2023.

---

[14] https://documents.egi.eu/document/3636
[15] EOSC-hub D13.4 Periodical assessment of the services: https://documents.egi.eu/document/3629
[16] https://documents.egi.eu/document/3634

Table 2: services in scope of the EOSC-Hub project are expected to become part of the EOSC Core and EOSC-Exchange[17] services

EOSC-hub service	EOSC-hub Work Package[18]	Included in EOSC Core capability – Iron/FAIR Lady (yes/no)	Included in EOSC-Exchange (yes/no)	Continuing in an EC-funded project[19] (F=EOSC Future, A=EGI-ACE, D=DICE, O=Other)
Hub Portfolio Services				
EOSC Portal (Including web content, Marketplace, Provider portal, portal metrics)	WP5	Yes		F
EOSC Order Handling System (SOMBO)	WP5	Yes		F
AAI (EGI Check-in, EUDAT B2ACCESS, INDIGO-IAM, eduTEAMS, RCauth, PERUN)	WP5	Yes		F, A (EGI Check-in, INDIGO-IAM, PERUN, RCauth), D (B2ACCESS), O (INDIGO-IAM)
Helpdesk (GGUS, EUDAT RT)	WP5	Yes		F, A(GGUS), D(EUDAT RT)
Monitoring (ARGO)	WP5	Yes		F, A, D

---

[17] the attribution to this category has been derived by the definition of EOSC-Exchange in the Fair Lady Report: "EOSC-Exchange is a digital marketplace that builds on the EOSC Core to offer a progressively growing set of services exploiting FAIR data and encouraging its reuse by publicly funded researchers. It is expected that services, such as those that store, preserve or transfer research data as well as those that compute against it, will be made available via EOSC-Exchange".

[18] WP4: Federated Service Management; WP5: Federation and Collaboration Services; WP6: Common Services; WP7: Thematic Services; WP9: Joint Digital Innovation Hub; WP10: Technical coordination

[19] the services continuing in a follow-on project are not totally maintained by these projects, they are typically also sustained by other actors/funding

Accounting (APEL, portal, repository, message broker)	WP5	Yes		F, A
Configuration Management System (GOCDB, DPMT, SVMON)	WP4, WP5	Yes		F, A (GOCDB), D (for DPMT)
EOSC Portal Metrics Dashboard	WP5	Yes		F
Common Services[20]				
EGI DataHub	WP6	No	Yes	A
B2FIND	WP6	No	Yes	D
B2STAGE	WP6	No	Yes	
B2DROP	WP6	No	Yes	D
EGI Cloud compute	WP6	No	Yes	A
EGI cloud container compute	WP6	No	Yes	A
EGI Workload management	WP6	No	Yes	A
EGI Online Storage	WP6	No	Yes	A
EGI High Throughput Compute (CREAM)	WP6	No	No (this service has been phased out)	[21]

---

[20] See D6.3: https://documents.egi.eu/document/3558

[21] CREAM has been phasing out during the last part of the EOSC-hub project, due to a strategy change of its main developer

TOSCA for Heat	WP6	No	Yes	[22]
Infrastructure Manager	WP6	No	Yes	A
PaaS Orchestration System	WP6	No	Yes	A, O
Future Gateway	WP6	No	Yes	[23]
B2HANDLE	WP6	No	Yes	D
B2SAFE	WP6	No	Yes	D
B2SHARE	WP6	No	Yes	D
B2NOTE	WP6	No	Yes	
eTDR	WP6	No	Yes	
TSD	WP6	No	Yes	D
ePOUTA	WP6	No	Yes	O (ELIXIR related project and health area)
Thematic Services				
CLARIN	WP7	No	Yes	F, O
DODAS	WP7	No	Yes	A, O
ECAS	WP7	No	Yes	A
GEOSS	WP7	No	Yes	A, O
OPENCOASTS	WP7	No	Yes	A, O
WeNMR	WP7	No	Yes	A

---

[22] TOSCA for Heat will continue to be supported by a community
[23] Future Gateway will be supported in future by one of its partners

EO Pillar	WP7	No	Yes	A, O
DARIAH	WP7	No	Yes	F, O
LIFEWATCH	WP7	No	Yes	A
Compliance Framework and Other Services				
Security Vulnerability Group, Information Security Management process in EOSC SMS	WP4	Yes		F, A
EOSC Security Coordination	WP4	Yes		F, A
EOSC Digital Innovation Hub	WP9	No		F
SMS	WP4	Yes		F
Interoperability guidelines (project documentation not formatted as a service)	WP10	Yes		F, A

3 Interoperability guidelines - feedback from EOSC stakeholders

As already discussed in the previous section, the EOSC technical specification and interoperability guidelines have been developed to the aim of identifying standards, well-known interfaces and best practices to facilitate the service integration in EOSC.

To improve the review process of these documents, in late 2020 an open consultation process has been started among the EOSC-hub partners and communities involved in the project. A set of six questions have been provided for each EOSC-hub technical area. The questions are aimed at identifying criticalities and stimulating comments and recommendations:

- *Does the specification clearly explain the main features offered by this service?*
- *Is the high-level architecture well described with highlighting all the main functions and interfaces?*
- *Do the interoperability guidelines provide clear and complete instructions to make services compliant with this specification interoperable?*
- *Do the interoperability guidelines provide clear and complete instructions to allow services to exploit this service?*
- *Are the "Examples of solutions implementing this specification" clear and well documented, and do they refer to concrete available services?*
- *Any additional comment?*

Comments and recommendations have been collected in a document (*Technical Specifications Survey results*[24]) that has been presented and discussed during the last TCOM meeting in March. The TCOM experts integrated such comments and recommendations by providing an updated version of the EOSC technical specification and interoperability guidelines[25].

A short summary of the comments is shown in the following subsections.

3.1 Federation/Core services

The EOSC-hub Federation/Core services are providing (and enabling) federation functions, needed to operate in EOSC, and support the processes and procedures of the EOSC Service Management System (SMS). This category includes services dealing with AAI, monitoring, accounting, operations management, order management, security and others.

In this section, a short summary of the comments (most significant ones) required for the Federation/Core services technical areas is given and shown in Table 3, together with the partner/community that provided them.

---

[24] https://docs.google.com/document/d/1vZ1ci9hsfZmAUV-4NsmO5vf_qZNSQVYw/edit
[25] https://www.eosc-hub.eu/technical-documentation

Table 3: summary of the comments (most significant ones) received for the Federation/Core services technical areas.

Partner/community	Tech. area	Comments/Recommendations
EGI Foundation GEANT Association CESSDA European Spallation Source ERIC DESY SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF RENAM Association	Helpdesk	The scope of its intended use and therefore the relevance of the functions would benefit from context and clarification. Include and define functionalities for incident management . Include the methodology / framework for Service management to be used. We would also like the service to establish some sort of prioritization procedure, to try to mitigate the flood of tickets and deal with the most important and/or urgent first. There are gaps in the description: Only one User is specified. It is not clear who this user is, whether it's a User of the Helpdesk as an EOSC-Hub internal user, or a Researcher User. Registered Service Provider services are implied by references to Thematic Services and 'Services of the EOSC Portfolio with integrated ticketing system', but most of the diagram would imply it is for internal services only. Missing the use case of a thematic service using the EOSC helpdesk with the ""direct usage"" integration. Improve the document where it is lacking in context and scope statements. It should also be made clear that the service provider can choose to use its own Helpdesk system either independently or can interoperate. It says that it uses xGUS. There is not enough description of it, but we suppose that we could go there and read more. The document is focussed on GGUS but it is unclear if this system is generic enough to support all aspects of EOSC and if this would be the only system that is supported. The section is called "Examples …" but in fact only one example is given - is this intentional?

		Review the glossary and then use some practical examples / use cases.
		The generation of relevant and useful feedback is dependent upon a clear understanding of the scope of the intended helpdesk system and is therefore limited because: It is not clear whether the 'Users' are Researcher Users or Service Provider Users.
EGI Foundation GEANT Association CESSDA European Spallation Source ERIC DESY SECD, Bineo, DDQ, INRIA, 52°North GmbH CREAF RENAM Association	Accounting	Note that the persistent use of the word 'Accounting' implies that this process offers financial transactions.
		It would benefit from a description/definition of the nature of usage that is actually logged.
		This would benefit from clarification: is it intended that I as a resource provider would be able to compare the usage of my services with the services of another resource provider?
		Explain how this accounting information is shared.
		It is too oriented towards the mechanism and UI, talking about views rather than capabilities.
		Resource centres that are providing compute or storage to the EOSC infrastructure have to implement a collector (a stand-alone script or program, or a built-in function of their resource system) that gathers accounting metrics formatted into a standardised record format (see next section for details). This creates an additional burden for the service providers. Has an assessment been made of how onerous this would or whether there is appetite to do so?
		Guidance and requirements appear specific to grid computing systems.
		It is not clear how the allocation of cost between all research projects and institutions a single researcher is affiliated with is mapped.
		Many components seem to be using different standards which made the interoperability guidelines complicated to follow.
		Define a more practical / use case-oriented approach.
		In general this must be seen as a very high-level document outlining some strategic consideration on how to move forward. Detailed design is needed and something to test

		against soon will clearly be needed. Overall you wrote a clear and interesting document.
EGI Foundation GEANT Association CESSDA European Spallation Source ERIC DESY SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF RENAM Association	Monitoring	If EOSC Monitoring/ARGO is to be offered to resource providers as a service, it would be useful for the benefits of ARGO over any other open source, free to use monitoring tool to be described, as the benefits described are fairly generic, and for any potential cost or licensing obligations to be noted. What sort of alerts/messages does Argo send? It would be interesting to see how this system works, as monitoring and alerting are very important to us. It provides a good high-level description of a generic monitoring service but does not contextualise it in the EOSC landscape (in the way that the 'Accounting' diagram does), however, the next chapter goes on to describe specific use cases that provide more context. Explains features at a high level but would benefit from scope and positioning statements. Why is SAML2 and not OpenIDConnect used for authentication? Why is Lavoisir chosen? What is ARGO and where and why is it used? There is some description of the components (e.g. Monitoring Engine, Sources of Truth) but otherwise it is too vague to judge whether or not it is a good specification. Based on only one service network: ARGO. It is unclear why what appears to be quite low-level events are being passed through to the EOSC Portal when each infrastructure is responsible for the correct operation of its own resources and the services they support. If the EOSC portal needed *any* status information, it would seem more efficient and effective to have this as simple and high level.
GÉANT Association SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF	Security	Security here is not a service so this is not appropriate. Standards and approach examples are listed but no concrete requirements that "the federation service" shall fulfil.

RENAM Association		There are best practices, but this is not a specification relating to a system with such interoperability.
		The document "EOSC Security'' seems to function as a compendium on some practices related to cyber security and how to ensure processes to deal with exploits. Not sure how an EOSC compliant service must provide more than a valid email address to fulfil the Security requirement.
GÉANT Association SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF RENAM Association	Software Quality	This is very vague and just talks about possible topics. Which is fine but not a specification. Based on a few examples, such as GitHub and Jenkins, these are best practices not a specification. It explains what should be necessary but does not describe standards a lot. Any software developer will recognize the necessary steps. It does not correspond to the diagram. It says that some federation services in EOSC are conformant to this specification, but the intro says, "no specific EOSC-hub services support for the assessment of the software quality". If that is true, how can be sure you are implementing the specification? The adopted standards (e.g. IEEE) should be put in context, rather than just links. Documents need to be harmonized among the different sections.
GÉANT Association SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF UKRI-STFC RENAM Association	AAI	If "this federation service" refers to the AAI component, then it does NOT clearly explain the main features as it looks concrete requirements how to leverage such a service. Very high-level diagrams not showing interfaces. The reader has to be a SAML + OpenID Connect + Federation Security expert. You need to do an effort to connect to less/non security expert people. The interoperability is via so called Proxies without defining the concrete interface and their functional / protocol requirements. For the listed services, more documentation is needed.

3.2   Common services

The EOSC-hub common services are services providing a middleware layer to exploit storage, compute and data resources and can be re-used by other services. This category includes services dealing with High-throughput computing, cloud computing, storage, data management and other specialised services. As an added value, those services enable scientific applications (thematic services) to efficiently use the resources provided by main European infrastructures.

In this section, a short summary of the comments (most significant ones) required for the Common service technical areas is given in Table 4, together with the partner/community that provided them.

Table 4: summary of the comments (most significant ones) received for the Common service technical areas.

Partner/Community	Tech. area	Comments/Recommendations
Sofia University "St.Kliment Ohridski" Vrije Universiteit Amsterdam RENAM Association SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF	Cloud Compute	The specifications are missing an overall overview, both of the cloud compute area as a whole and for cloud compute relative to the other proposals from EOSC-hub. As a consequence, as someone not already intimately familiar with the subject area or EOSC, I found it hard to understand the roles of all the different components being described. We believe that a context in cloud compute is needed to understand what is being offered. Although it was hard to understand due to the lack of overview, I do feel the specification covers all the necessary main functions. Specifically on the subject of VM management, the proposal seems to not require any compliance with a standard, only *preferring* to follow *an* open standard. This looks like an interface requirement description, but not detailed enough interoperability guidelines. An API is mentioned but there is no endpoint or list of operations that API supports. The only reference is the EGI service, other links are just pointing to containerisation services. We do not understand why 3 documents. Are these services independent?

Vrije Universiteit Amsterdam RENAM Association SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF	HTC/HPC	It is unclear how to use the services. If we have understood well, you are offering a Cloud Infrastructure environment services such as OpenStack in which each client could select the list of resources that they need for their projects or their processes. We get the idea, but it is not clearly defined. It would have been better to link it to include a description of how this relates to the cloud compute proposals, and to adopt a common model and terminology. It only provides a list of standards used with no explanation. It serves more as an overview and will require much deeper exploration. You are defining in some cases the interoperability like a list of functionalities. It is confusing.
Open Knowledge Maps Vrije Universiteit Amsterdam RENAM Association CODATA CNR DANS SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF	Metadata Management and Data Discovery	Need to be clearer in the definition of the service. Using a common model & terminology would have made it so much easier to understand, rather than making 3 disparate proposals. Common functionalities are not described, what are they? and for what purpose. Documents need to be harmonized in their different parts. More detailed comments have been collected in the document *Technical Specifications Survey results*[1]
Vrije Universiteit Amsterdam RENAM Association SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF	PaaS	We are not sure that we understand what is a "virtualized computing infrastructures with complex topologies". Is this solution fit for purpose or just a pile of cloud technologies bundled by a web service? Some parts of the document need to be reviewed to better express their meaning. More examples need to be provided. This looks like a huge stack of technology looking for a solution, not the other way around.

| Vrije Universiteit Amsterdam

RENAM Association

CNR, SECD, Bineo, DDQ, INRIA, 52°North GmbH, CREAF | Workflow | I do not understand why these specifications are grouped together. Workflow management is a thoroughly ambiguous term, and assuming it covers the Jira/Helpdesk-aspects of the Marketplace I feel it is also the wrong term (probably would use service delivery).

Some parts of the document need to be harmonised.

This seems to be a wholly technical specification of a generic modern-day cloud-based service latched onto a good overview of an AI-platform. Separately the parts certainly make sense, but I am not sure the interoperability requirements of this are at all AI-platform-specific.

In general this document does not take enough into accounts the needs and viewpoints of Service Providers.

The name does not have a direct relation with the internal goals of the service: Artificial Intelligence, Deep Learning, and Machine Learning.

SQA requirements for Service Providers are indicated, with no mention on Code Accessibility. I agree on this, since it opens the possibility for Service Providers to offer non-open-sourced solutions, but I'd suggest adding the condition that in this case, a restricted access to the code repository must be given to EOSC, to allow code inspections. |

4 EOSC-hub Technical architecture and interoperability guidelines update

The EOSC technical specification and interoperability guidelines have been developed within the activities of the project and their evolution have been guided by the EOSC-hub Technology committee (TCOM). These guidelines are aimed at identifying standards, well-known interfaces and best practices to facilitate the service integration in EOSC. Moreover, they cover both common/horizontal services (as part of the EOSC Exchange) and federation/core services (as part of the EOSC Core) and can be considered as building blocks of the technical aspects of EOSC Interoperability Framework.

EOSC-hub released 18 interoperability guidelines. They have been organized in different technical areas. Six of them are related to the EOSC Core/Federation services provide (and enable) federation functions needed to operate in EOSC. They are Accounting, Helpdesk, Monitoring, Security, Software Quality Assurance and AAI (see Table 5 for more details).

The rest of the interoperability guidelines are related to Common services providing a middleware layer to exploit storage, compute and data resources and can be re-used by other services. They are Cloud Compute, HTC/HPC Specification, Metadata Management and Data Discovery, PaaS Solutions and Workflow management and user interfaces and Data analytics (see Table 6 for more details).

The implementation of the EOSC interoperability guidelines is extensively described in the deliverables referring to the activities of WP5 and WP6 (ADD REF). The guidelines reflect the technical integration expertise that already belong to the e-infrastructures (EGI, EUDAT and INDIGO) and to the TCOM members that supported user communities willing to integrate their services in EOSC.

Table 5: EOSC interoperability guidelines for Core/Federated services with the related description and reference.

Federation services		
Interop. guideline title/Area	**Description**	**Reference**
Accounting	Collects, stores, aggregates, and displays usage information of HTC compute, storage space, cloud VM and data set resources. This usage data is collected from the Resource Centres of the EOSC infrastructure.	http://go.egi.eu/eosc-hub-accounting
Helpdesk	Is the entry point and ticketing system/request tracker for issues concerning the available EOSC services. Main features offered to the user are:	http://go.egi.eu/eosc-hub-helpdesk

	- Creation of a ticket for any of the EOSC Services (Hub and EOSC Portfolios) - Displaying all the tickets created by the owner - Finding a previously created ticket - Notifying the user of answers and changes to the tickets - Access which is integrated with the EOSC Portal AAI system	
Monitoring	Is the key service needed to gain insights into an infrastructure. It needs to be continuous and on-demand to quickly detect, correlate, and analyse data for a fast reaction to anomalous behaviour.	http://go.egi.eu/eosc-hub-monitoring
Security	Established by having policy and practices for infrastructure service providers, for end users and their communities, and between infrastructures (including any national CSIRT).	http://go.egi.eu/eosc-hub-security
Software Quality Assurance	Is the process responsible for the overall supervision of both software development lifecycle ensuring that the required quality level is achieved together with service quality.	http://go.egi.eu/eosc-hub-sw-quality
AAI services	Access to resources in EOSC plays a crucial role. It is therefore essential to have an Authentication and Authorisation Infrastructure (AAI) as well as AAI building blocks in place that can address the variety of requirements and the distributed nature of EOSC.	http://go.egi.eu/eosc-hub-aai

Table 6: EOSC interoperability guidelines for Common services with the related description and reference.

Common services			
Tech. Area	**Interop. guideline title**	**Description**	**Reference**
Cloud Compute	Cloud Compute - IaaS VM Management	Covers those services that provide on-demand API-based access to computing resources as Virtual Machines that can run user-defined arbitrary software (including operating systems and applications).	http://go.egi.eu/cloud-iaas-vm
Cloud Compute	Cloud Compute - IaaS Container Management	Provides on-demand API-based management of container-based applications supporting the (Automated) Orchestration of container-based applications which manage the deployment of a complete lifecycle of the containers that compose an application into a set of computing resources.	http://go.egi.eu/cloud-iaas-container
Cloud Compute	Cloud Compute - IaaS Orchestration	Covers tools and services that automate the deployment of applications by executing a set of tasks that interact with the cloud services to start Virtual Machines, create storage devices and objects, configure networking and any other kind of related services to install and run the application.	http://go.egi.eu/cloud-iaas-orchestration
HTC/HPC Specification	HTC/HPC Compute - Multitenant job submission	Exposes a macro-feature for Multitenant job submission, which relates to the capability of submitting HPC/HTC jobs with predefined constraints (both at resources and software) without a	http://go.egi.eu/htc-multitenant-job

		previously deployed virtual infrastructure.	
HTC/HPC Specification	HTC/HPC Compute - Multitenant Containerised job submission	This macro-feature is complementary to the Multitenant Job Submission and it should be considered as an extension.	http://go.egi.eu/htc-multitenant-container
HTC/HPC Specification	HTC/HPC Compute - HPC/HTC clusters on demand	Service will explicitly deploy a single-tenant cluster backend to be used by the user community managed by the user who deployed it.	http://go.egi.eu/htc-cluster-ondemand
Metadata Management and Data Discovery	Metadata Cataloguing and Indexing	comprise the entire metadata ingestion workflow, i.e. 1. Metadata harvesting from community repositories, 2. Metadata mapping on common schemas including curation and validation and, 3. Uploading and indexing of metadata records in the metadata catalogue, to enable Data Discovery and Access, see related macro feature.	http://go.egi.eu/metadata-catalogue
Metadata Management and Data Discovery	Data Discovery and Access	Comprises the ability for end-users to search for data resources and access the referenced data. This functionality requires and is based on the existence of an indexed metadata catalogue (see macro feature Metadata Cataloguing and Indexing).	http://go.egi.eu/metadata-access
Metadata Management	Annotation Service	Enables end-users to extend descriptions of datasets or parts of datasets with user-defined content, without modifying the underlying	http://go.egi.eu/metadata-annotation

and Data Discovery		dataset, e.g. adding comments, free text keywords or semantic tags (keywords from ontologies).	
PaaS Solutions	PaaS Solutions	Allows the users to deploy virtualised computing infrastructures with complex topologies (such as clusters of virtual machines or applications packaged as Docker containers) using standardized interfaces based on REST APIs and adopting the TOSCA (Topology and Orchestration Specification for Cloud Applications) templating language for the description of Cloud-based applications.	http://go.egi.eu/paas-solutions
Workflow management and user interfaces and Data analytics	Marketplace	Is a dedicated platform where services are presented to the users and made available to get access to.	http://go.egi.eu/marketplace-spec
Workflow management and user interfaces and Data analytics	Machine Learning/Deep Learning data analytics services	Provides tools and mechanisms to build, encapsulate and execute Artificial Intelligence, Machine Learning and Deep Learning applications across different platforms, covering the whole development life cycle.	http://go.egi.eu/machinelearning_services

5   Conclusions and future works

The Interoperability and Integration guidelines developed during the EOSC-hub project and the contributions from the project to the EOSC architecture and interoperability framework had the aim to define the high-level architecture for basic EOSC technical functions and promote EOSC standards and APIs. Facilitating the access to services, lowering the barriers to integrating and composing services and promoting the usage of services between adjacent communities, the guidelines have been used by a large number of related projects that continue beyond the EOSC-hub project itself to make their services EOSC-compatible.

The EOSC-hub project will be succeeded by projects funded by the calls for proposals INFRAEOSC-03 and -07 which were part of the final Work Programme of Horizon 2020. In very simple terms, the project funded by the INFRAEOSC-03 call is expected to implement the EOSC Core until mid- to-late 2023 by integrating the services it should contain, whereas the projects funded by the INFRAEOSC-07 call will provide services to the EOSC-Exchange.

he EOSC Future project, funded by the INFRAEOSC-03 call, is expected to setup the new EOSC Interoperability Framework as a library of EOSC architecture and interoperability guidelines, covering the whole research data lifecycle (discover and reuse, processing and analysis, curation and preservation, access and sharing), to promote the branding and adoption of standards and common best practices in EOSC. Such a library of architecture and interoperability guidelines will be the main instruments to foster secure resource integration, interoperability, and composability. EOSC Future will not start this effort from scratch but will rely on interoperability guidelines developed in the context of several EOSC projects (EOSC-hub, OpenAIRE Advance, AARC, EOSC Enhance, FAIRsFAIR, etc) and will rely on the experience in IT infrastructures for research of the 5 scientific clusters that are part of the project and will act also as an important source of requirements. The project will take care to harmonise all these different contributions in a homogeneous interoperability framework. EOSC-hub will provide a very relevant contribution to this effort hand overing the guidelines developed by the project for services of the EOSC Core (AAI, accounting, helpdesk, monitoring, etc.) and for key technical areas of the EOSC Exchange (Cloud/HTC/HPC computing environments, data management systems, analytics, ML/AI, etc) described in the previous sections of this document.

The INFRAEOSC-07 projects will collaborate with EOSC Future in the establishment of the interoperability framework contributing with guidelines related to the technical areas covered by their activities. In particular:

- EGI-ACE[26]: guidelines on data analysis services and compute infrastrastructure;
- DICE[27]: guidelines on data management services

---

[26] https://www.egi.eu/projects/egi-ace/

[27] https://www.dice-eosc.eu/

- OpenAIRE Nexus[28]: guidelines on research scholarship;

- C-SCALE[29] (Copernicus - eoSC AnaLytics Engine) and RELIANCE[30] : guidelines for EO data management and analysis.

Furthermore, the INFRAEOSC-07 projects will deliver live instances of services, compliant with the EOSC interoperability guidelines, and will make them available to user communities. This makes evident how the EOSC-hub effort on interoperability of services has been effective in fostering open science and addressing real users' needs.

It is worth noting that the EOSC Association is setting up a number of different advisory groups including one that will deal with interoperability between EOSC services, and this will contribute to the further evolution of the work done in

---

[28] https://www.openaire.eu/openaire-nexus-project
[29] https://c-scale.eu/
[30] https://www.reliance-project.eu/