

D5.1: Initial EOSC Service Architecture

Author(s)	L. Candela (CNR), D. Castelli (CNR), G. La Rocca (EGI), A. Lukkarinen (CSC), P. Manghi (CNR), P. Pagano (CNR), E. Papadopoulou (ATHENA RC)
Status	Final
Version	v1.0
Date	6/2/2018

Dissemination Level

<input checked="" type="checkbox"/>	PU: Public
<input type="checkbox"/>	PP: Restricted to other programme participants (including the Commission)
<input type="checkbox"/>	RE: Restricted to a group specified by the consortium (including the Commission)
<input type="checkbox"/>	CO: Confidential, only for members of the consortium (including the Commission)

Abstract:

EOSC is expected to be *“a federated, globally accessible, multidisciplinary environment where researchers, innovators, companies and citizens can publish, find, use and reuse each other’s data, tools, publications and other outputs for research, innovation and educational purposes”*.

This deliverable documents the initial architecture of the EOSC underlying IT system. In order to identify the constituents of the EOSC System, the deliverable identifies (a) the primary roles that the actors involved in EOSC system exploitation and development are expected to play, and (b) the major activities these actors are performing when exploiting, contributing and managing the EOSC system. Then, by analysing these elements, an initial list of service typologies is derived. This list is expected to be iteratively revised and enriched following the better understanding of the Open Science approach and the emergence of related policies, governance models and technologies able to support it.

The European Open Science Cloud for Research pilot project (EOSCpilot) is funded by the European Commission, DG Research & Innovation under contract no. 739563

Document identifier: EOSCpilot -WP5-D5.1	
Deliverable lead	CNR
Related work package	WP5
Author(s)	L. Candela (CNR), D. Castelli (CNR), G. La Rocca (EGI), A. Lukkarinen (CSC), P. Manghi (CNR), P. Pagano (CNR), E. Papadopoulou (ATHENA RC)
Contributor(s)	D. Bailo (INGV), J. Bot (SURFSara), M. Campanella (GARR), G. Coro (CNR), S. Holsinger (EGI), E. Huizer (GEANT), D. Lecarpentier (CSC), T. Ferrari (EGI), E. Fernández (EGI), R. C. Jimenez (ELIXIR), P. Kahlem (ELIXIR), S. Kuijpers (SURFSara), N. Manola (ATHENA RC), M. Scott (GEANT), A. Sevasti (GRNET), G. Sipos (EGI), M. van de Sanden (SURFSara), J. van Wezel (SURFSara), D. Vianello (EMBL_EBI), M. Viljoen (EGI), A. Whyte (DCC), F. Zoppi (CNR)
Due date	31/12/2017
Actual submission date	30/01/2018
Reviewed by	P. Oster (CSC), C. Calvin (CEA)
Approved by	J. van Wezel (SURFSARA), T. Ferrari (EGI.eu)
Start date of Project	01/01/2017
Duration	24 months

Versioning and contribution history

Version	Date	Authors	Notes
0.1	02/10/2017	L. Candela (CNR), D. Castelli (CNR)	Initial draft circulated (by re-structuring the document prepared for the EOSCpilot service provider workshop on 13-14 September).
0.8	20/12/2017	L. Candela (CNR), D. Castelli (CNR), G. La Rocca (EGI), A. Lukkarinen (CSC), P. Manghi (CNR), P. Pagano (CNR), E. Papadopoulou (ATHENA RC)	First complete version ready for internal review.

0.9	20/01/2018	L. Candela (CNR), D. Castelli (CNR), G. La Rocca (EGI), A. Lukkarinen (CSC), P. Manghi (CNR), P. Pagano (CNR), E. Papadopoulou (ATHENA RC)	First complete version ready for external review
1.0	6/2/2018	L. Candela (CNR), D. Castelli (CNR)	Added changes after Internal Review. Final version produced to respond to reviewer comments.

Copyright notice: This work is licensed under the Creative Commons CC-BY 4.0 license. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0>.

Disclaimer: The content of the document herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While the information contained in the document is believed to be accurate, the author(s) or any other participant in the EOSCpilot Consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the EOSCpilot Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the EOSCpilot Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

TABLE OF CONTENT

EXECUTIVE SUMMARY	6
1. INTRODUCTION	8
1.1. Deliverable in context	10
1.2. Deliverable organisation	10
2. DEFINING THE EOSC SYSTEM	12
2.1. EOSC System Users.....	14
2.2. Activities to be supported via the EOSC System	17
2.2.1. Activities performed by Researchers.....	17
2.2.2. Activities performed by Research Admins.....	18
2.2.3. Activities performed by Third-Party Service Providers.....	18
2.2.4. Activities performed by EOSC Suppliers	19
2.2.5. Activities performed by EOSC System Managers	19
3. EOSC SYSTEM SETTINGS AND DESIGN DECISIONS.....	21
3.1. System of Systems.....	21
3.2. EOSC Principles of Engagement and Governance.....	23
3.3. EOSC Service Added-Value Federation Models	24
3.4. Distributed Service Provisioning	26
3.5. FAIRness	26
4. EOSC SYSTEM COMPONENTS	27
4.1. EOSC Core Services.....	28
4.2. Services for Researchers	30
4.3. Services for Research Admins	31
4.4. Services for Third-party Service Providers	32
4.5. Services for EOSC Suppliers.....	33
4.6. Services for EOSC System Managers.....	34
5. CONCLUSIONS AND NEXT STEPS.....	35
REFERENCES.....	37
ANNEX A. GLOSSARY.....	38
ANNEX B. DETAILED DESCRIPTION OF ACTIVITIES TO BE SUPPORTED BY THE EOSC SYSTEM.....	40
ANNEX C. EOSC STAKEHOLDERS AND THEIR ROLES IN THE EOSC SYSTEM	51
ANNEX D. TYPICAL SCIENTIFIC SCENARIOS	55

LIST OF FIGURES

Figure 1. Open Science opens up the entire research enterprise (inner circle) by using a variety of means and digital tools (outer circle) [EC - DG R&I 2016, p. 36].	8
Figure 2. EOSC System and actor roles.....	12
Figure 3. EOSC Services stacking	14
Figure 4. EOSC System Roles	15
Figure 5. Pattern of an EOSC Service implementing the Invisible coordinator model.....	25
Figure 6. Pattern of an EOSC Service implementing the Matchmaker model	25
Figure 7. Pattern of an EOSC Service implementing the One-stop-shop model.....	25
Figure 8. EOSC Services	28
Figure 9. EOSC Core Services	28

Figure 10. EOSC Services for Researchers	30
Figure 11. EOSC Services for Research Admins	32
Figure 12. EOSC Services for Third-party Service Providers	33
Figure 13. EOSC Services for EOSC Suppliers.....	33
Figure 14. EOSC Services for EOSC System Managers.....	34
Figure 15. EPOS/VERCE Science Demonstrator	56
Figure 16. DP-HEP Science Demonstrator	57

LIST OF TABLES

Table 1. Activities performed by Researchers.....	40
Table 2. Activities performed by Research Admins.....	43
Table 3. Activities performed by Third-party Service Providers.....	45
Table 4. Activities performed by EOSC Suppliers	47
Table 5. Activities performed by EOSC System Managers	48
Table 6. Associating EOSC Stakeholders Primary Roles with EOSC System Actor Roles	51
Table 7. Associating EOSC Stakeholders Supplementary Roles with EOSC System Actor Roles	52
Table 8. Associating EOSC Stakeholders with EOSC System Actor Roles	52

EXECUTIVE SUMMARY

The European Open Science Cloud (EOSC) will be one of the most relevant instruments for meeting the Open Science challenge. It is envisaged to be the environment “where researchers, innovators, companies and citizens can publish, find and re-use each other's data and tools for research, innovation and educational purposes” [Ayrís et al. 2016].

*The aim of this deliverable is to lay the foundations for the architecture of the **EOSC underlying IT system** in terms of the services called to provide the envisaged functionality. This reference architecture is expected to iteratively evolve over time as an effect of better understanding the Open Science approach and of the emergence of appropriate policies, governance models and technologies able to support it.*

*The EOSC system architecture is designed to take into account not only the functional needs of the EOSC **end users**¹ called to facilitate and profit from an open approach to science (e.g. researchers, research administrators, innovators). It also takes into account the needs of many other actors, like those exploiting the EOSC services as the base for implementing new added value services, those dealing with the provision (e.g. **supplier**) and the management EOSC and its IT services (e.g. **managers**). The availability of appropriate services supporting these actors is key for the success of EOSC since their impact on EOSC aspects such as its maintenance costs, sustainability, expected evolution, business models, governance and so on can be large.*

*The design of this architecture is based on a number of initial assumptions. Among them: (i) functionalities are provisioned **as-a-Service**, i.e. they are made available by an online service operated by a provider that takes care of the technical and organisational approaches that it needs to deliver its functionality; (ii) EOSC is regulated by a set of **Principles of Engagement**. All the services included in the EOSC Service Catalogue (hereafter named “EOSC Services”) should satisfy these principles; (iii) the EOSC system is modelled as an open and evolving **System of Systems (SoS)** where the component systems providing services include existing and emerging Research Infrastructures (including e-Infrastructures) and other types of Service Providers, (iv) EOSC services provision is based on an open and evolving set of **EOSC Nodes** spread across several organisations and regions; (v) EOSC Services should promote and support **FAIRness**, e.g. the data managed by EOSC Services should implement the FAIR principles, some EOSC Services should be explicitly envisaged to enable users implementing FAIR principles.*

*The presented initial EOSC system architecture identifies **a set of service types** that the EOSC IT system should provide to implement an initial step towards the Open Science objective.*

*The EOSC Services are the instruments constituting the EOSC System offering, i.e. they are the elements that, at a given point in time, provide the functionalities that can be exploited by the EOSC users. EOSC Services are expected to (a) be listed in the **EOSC Catalogue**, (b) closely adhere to the **EOSC Principles of Engagement**; (c) be operated by an **EOSC Service Provider** (i.e. any service provider qualified as EOSC service provider), and (d) be delivered with well defined **Service Level Agreement(s)**. EOSC Service Providers are expected to leverage **EOSC Service Components** offered by **EOSC Suppliers** according to **Underpinning Agreement(s)** to realise their **added-value services**.*

*The EOSC services might be offered by or procured to existing and emerging components of the SoS (e.g. Research Infrastructures, Data Centers, Service Providers) or they might be partially implemented and operated by an “EOSC entity” (whatever form it will take). The decision among these alternative deployment models is **out of the scope of this deliverable**. This deliverable is simply meant to support the future implementation plan by highlighting what are the services **required to support the EOSC expected functionality**.*

¹ In D2.2 [Hienola et al. 2017] three stakeholder roles are envisaged: **Consumers** (i.e. stakeholders that will make use of services, data, or other resources from EOSC), **Providers** (i.e. stakeholders that provides services, data or other resources - e.g. scientific instruments, training - into EOSC), and **Decision-makers** (i.e. stakeholders that will be involved in the strategic direction, compliance and funding of EOSC). A mapping among these roles and those envisaged in this deliverable to characterise the actor roles involved in setting up and operating the EOSC IT system is given in Sec. 2 and Appendix C. The two sets of roles almost correspond yet some specificities lead to different naming.

The list of actors, services and the architecture assumptions presented in this document have been largely discussed with project partners through questionnaires, shared documents and face-to-face meetings. This exchange has included feedback from the shepherds of the EOSCpilot Science Demonstrators and from the stakeholders of the Research Infrastructures represented in the EOSCpilot project.

*Due to its complexity and novelty, the definition of the EOSC system architecture was necessary **to be carried out incrementally**. In this first release of the initial EOSC architecture, a lot of attention has been paid primarily to identify and agree on an initial understanding of the basic principles and scope of the EOSC system. It has also been chosen to focus on a system for Open Science that reflects what can be supported with mature technological solutions available today. This choice has been done primarily to be able to manage the inherent complexity of the architecture definition task and to enable an easier mapping of the produced model to the services of existing component systems (e.g. Research Infrastructures, e-Infrastructures, Data Centers, Repositories, Registries, etc.) can provide to EOSC.*

The architecture definition process will continue iteratively through enrichment, validation and revision steps during the second period of the project. A final richer and more consolidated version of the architecture will be presented in deliverable D5.4, to be released at M24.

*The discussions around the architecture held at EOSCpilot organised events and calls have highlighted relationships of mutual contribution with many other project activities in the same and across work packages (WP2, WP3, WP4, WP6, WP7). Due to the parallel timing of these activities, and the evolving nature of this deliverable, some alignment steps are still needed. As contribution to this alignment objective, a shared **Glossary** has been recently created. A first version of this Glossary, which contains the terms used in this deliverable, is added at the end of this document. The plan is to refine and complete it in the future, with the contribution of other WPs, in order to derive a core set of concepts agreed among the many initiatives participating in the project.*

1. INTRODUCTION

The European Open Science Cloud (EOSC) is will be one of the most relevant instruments for **meeting the Open Science challenge**². The first step in designing an appropriate EOSC system architecture is therefore to clearly understand this challenge. In the book “Open Innovation, Open Science, Open to the World” [EC - DG R&I 2016], Open Science is defined as follows:

*“Open Science represents a new approach to the scientific process **based on cooperative work and new ways of diffusing knowledge by using digital technologies and new collaborative tools**. The idea captures a systemic change to the way science and research have been carried out for the last fifty years: shifting from the standard practices of publishing research results in scientific publications **towards sharing and using all available knowledge at an earlier stage in the research process.**”*

*“Open Science **has an impact on the entire research cycle**, from the inception of research to its publication, and on how this cycle is organised. The outer circle in Figure 1 shows the new interconnected nature of Open Science, while the inner circle shows the entire scientific process, from the conceptualisation of research ideas to publishing. **Each step in the scientific process is linked to ongoing changes brought about by Open Science**, such as the emergence of alternative systems to establish scientific reputation, changes in the way the quality and impact of research are evaluated, the growing use of scientific blogs, open annotation and open access to data and publications.”*

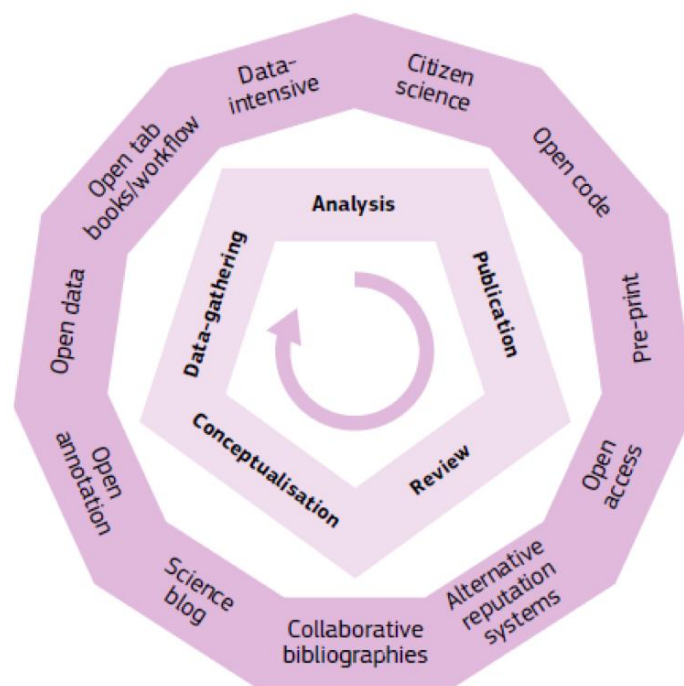


Figure 1. Open Science opens up the entire research enterprise (inner circle) by using a variety of means and digital tools (outer circle) [EC - DG R&I 2016, p. 36].

[Nielsen 2012] contains a further clarification of this challenge:

“Open science is the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process.”

² There will never exist a definition of Open Science that fits all. A discussion on this is available at <https://im2punt0.wordpress.com/2017/03/27/defining-open-science-definitions/> In the remainder of the text a set of characterisations of Open Science are exploited to clarify what is driving the rest.

This implies that:

- *Open Science is expected to deal with **scientific knowledge of all kinds**, i.e. the subject is open ended including journal articles, data, code, online software tools, questions, ideas, and speculations; anything which can be considered knowledge;*
- *Open Science is expected to support the **as early as is practical** release of this knowledge, i.e. it should be generally available as soon as it is produced in a version deemed sufficient for deriving further knowledge.*

EOsc is one of the five lines of policy actions to support the development of Open Science in Europe [EC - DG R&I 2016, p. 45]. This specific line aims to:

“Developing research infrastructures for Open Science, to improve data hosting, access and governance, with the development of a common framework for research data and creation of a European Open Science Cloud, a major initiative to build the necessary Open Science infrastructure in Europe.”

The EOsc first High Level Expert Group (HLEG) report [Ayrís et al. 2016] advises on what the EOsc system should/could offer by presenting major characteristics and expectations:

- *It should be a “**federated, globally accessible** environment where researchers, innovators, companies and citizens can **publish, find and re-use** each other's **data and tools** for research, innovation and educational purposes under well defined, secure and trusted conditions, supported by a sustainable and just and value-for money model”³.*
- *“It should enable **trusted access to services, systems and the re-use of shared scientific data** across disciplinary, social and geographical borders”.*
- *It should “build on **existing capacity and expertise** where possible”.*
- *It is approached as a “**federated environment** for scientific data sharing and re-use, **based on existing and emerging elements in the Member States, with lightweight international guidance and governance and a large degree of freedom regarding practical implementation**”.*
- *“It includes the required **human expertise, resources, standards, best practices** as well as the underpinning **technical infrastructures**”.*
- *It supports “systematic and professional data management and long-term stewardship of scientific data artefacts and services in Europe and globally”.*

The EOsc Declaration [EC 2017] issued by the EC in October 2017 further expands the EOsc requirements:

*The EOsc will be developed as a **data infrastructure commons serving the needs of scientists**. It should provide both common functions and localised services delegated to community level. Indeed, the **EOsc will federate existing resources across national data centres, European e-infrastructures and research infrastructures**; service provision will be based on local-to-central subsidiarity (e.g. national and disciplinary nodes connected to nodes of pan-European level); it will top-up mature capacity through the **acquisition of resources at pan-European level by EOsc operators**, to serve a wider number of researchers in Europe. Users should contribute to define the main common functionalities needed by their own community. A continuous dialogue to build trust and agreements among funders, users and service providers is necessary for sustainability.*

The aim of this deliverable is to lay the foundations for an initial architecture of the underlying **EOsc System** (i.e. the IT system supporting EOsc) that is expected to evolve over time with a better understanding of the Open Science approach and with the emergence of appropriate policies, governance

³ The one reported here is actually a slightly revised definition produced in the context of the Open Science Policy Platform Working group [Andreozzi et al. 2017].

models and technologies able to support it. This architecture is designed to take into account ***the entire spectrum of the aforementioned requirements*** that concern not only the functional needs of the different actors involved in the implementation of the open science approach, but also aspects like (re-)use, maintenance costs, sustainability, expected evolution, business models, governance and so on.

1.1. Deliverable in context

Given the novelty and complexity of this task, the EOSC architecture has been, and will continue to be, defined iteratively through validation and revision steps. Following this approach, a number of internal releases of this document were produced. The first public release has been discussed with representatives of EOSC Science Demonstrators and Research Infrastructures (RIs) in a face-to-face meeting organised on 14-15 September 2017 in Pisa. Input was also collected through dedicated questionnaires addressed to WP4 shepherds, with exchanges with representatives of research infrastructure and also through feedback collected at the “EOSCPilot Summit” held on 28-29 November 2017 in Brussels. This line of work will continue and will be presented in deliverable (D5.4) that will be released at M24.

1.2. Deliverable organisation

The remainder of the deliverable is organised as follows.

Section 2 introduces the actor roles directly served by the EOSC system (be they end-users, managers and/or suppliers) as well as their typical activities to be facilitated by the EOSC system.

Section 3 highlights key settings and design decisions characterising the EOSC system. In particular, it underlines that (a) the EOSC system is modelled as an open and evolving **System of Systems (SoS)** where the component systems are existing and emerging Infrastructures and Service Providers (public or private) and the EOSC system offers **added value services**, i.e. it performs functions and carries out purposes that do not reside in any component system; (b) the EOSC System is regulated by a set of **Principles of Engagement**, i.e. all the component systems, service providers and the services they contribute to EOSC should adhere to the regulations and practices established by these principles; (c) the EOSC functionalities are provisioned **as-a-Service**, i.e. they are made available by an online service operated by a provider that takes care of the technical and organisational approaches that it needs to deliver its functionality; (d) EOSC services provision is based on an open and evolving set of **EOSC Nodes** spread across several organisations and regions; (e) EOSC Services should promote and support **FAIRness**, e.g. the data managed by EOSC Service should implement the FAIR principles, some EOSC Services should be explicitly envisaged to enable users implementing FAIR principles.

Section 4 delineates an initial **reference architecture** for EOSC by highlighting an initial list of services that the EOSC system should offer to satisfy its envisaged mission, i.e. to be an environment “*where researchers, innovators, companies and citizens can publish, find and re-use each other's data and tools for research, innovation and educational purposes under well defined, secure and trusted conditions, supported by a sustainable and value-for money model*”. The services are meant to support the requirements of the actors identified in Section 2 in a framework, discussed at the beginning of the session, that is characterized by a series of prerequisites on re-use and evolution and by different roles played by EOSC as value added provider.

Section 5 concludes the report by discussing next steps.

Some Annexes are also included.

Annex A reports a **Glossary** listing acronyms and terms used in the document and their definitions. It is meant to support the development of a common understanding and definition for some key terms characterising this and other related deliverables (e.g. D5.2). This Glossary has been discussed so far among the WP5 members. The plan is to progressively involve members of other WPs and transform it into a shared project resource.

Annex B details the activities the EOsc System should support per user role.

Annex C reports a **mapping between the EOsc Stakeholders**, as identified in WP2, and the **user roles** presented in this document to implicitly highlight the needs of those operating in the stakeholder contexts.

Annex D reports an early discussion on EOsc Demonstrators by taking into account the proposed architecture.

2. DEFINING THE EOSC SYSTEM

The objective of this document is to lay the foundations of the IT system architecture that is expected to facilitate the emergence of the Open Science approach to the research process. For simplicity in the rest of this document the terms ‘architecture’ and ‘EOSC’ are used to mean, respectively, ‘reference architecture’⁴ and ‘EOSC system’ when no confusion arises, i.e. we do not consider other elements like technologies, protocols, and products nor human expertise, best practices, funding models, etc.

The introductory part of this section is meant to provide readers with an overview of the *main concepts and relations* among them characterising the EOSC system⁵. Such concepts are in **italic bold** (first occurrence) in the text and informally defined to simplify the reading of this introductory part. A definition of these key terms is in the Glossary in Annex A.

The ***EOSC System*** (i.e. the IT system implementing EOSC) aims to serve the needs of several *actors playing diverse roles* (cf. Sec 2.1) by providing them with functionalities enacting the implementation of a series of *tasks and activities* (cf. Sec. 2.2). By design, EOSC is expected to leverage, harmonize and federate multiple existing systems and solutions operated by various providers (e.g. Research Infrastructures, e-Infrastructure, services providers) to support research activities.

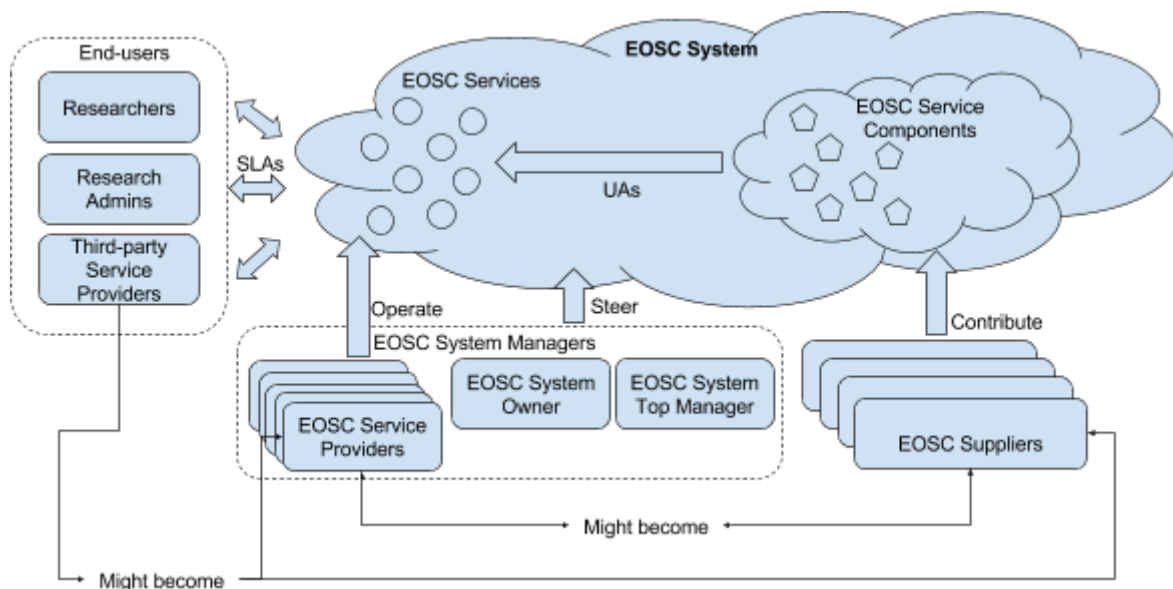


Figure 2. EOSC System and actor roles⁶

⁴ By *reference architecture* here we refer to the OAIS definitions (<http://docs.oasis-open.org/soa-rm/soa-ra/v1.0/soa-ra-cd-02.html>): “The abstract architectural elements in the domain independent of the technologies, protocols, and products that are used to implement the domain”.

⁵ Several aspects have been considered in identifying these concepts including IT Services Management methodologies (e.g. FitSM) as well as EOSC basic choices like building upon existing “systems” (actually to leverage services and facilities offered by Research Infrastructures and other Service Providers), and offering added-value services aiming at supporting Open Science. To reduce the complexity of the proposed model few concepts have been selected. Future versions of the EOSC Architecture will reconsider this decision and the resulting set of concepts at the light of (i) feedback, comments and request for changes, (ii) developments in understanding and modelling of aspects affecting the EOSC architecture (EOSC Governance), and (iii) emergence of best practices, standards and shared solutions (e.g. Open APIs <https://www.tmforum.org/open-apis/>).

⁶ The envisaged actor roles have a different naming than the stakeholder roles envisaged by D2.2 [Hienola et al. 2017] simply because the differences in scope between the two documents. In particular, D2.2 identified three primary roles for EOSC Stakeholders: **Consumers** (i.e. stakeholders that will make use of services, data, or other resources from EOSC), **Providers** (i.e. stakeholders that provides services, data or other resources - e.g. scientific instruments, training - into EOSC), and **Decision-makers** (i.e. stakeholders that will be involved in the strategic direction, compliance and funding of EOSC). Consumers correspond to the End-users, Providers include both EOSC Service Providers and EOSC Suppliers, while Decision-makers impact on the activities of EOSC System Manager and EOSC System Top Manager. A mapping among these roles and those envisaged in this deliverable to characterise the actor roles involved in setting up and operating the EOSC IT system is given in Appendix C.

Figure 2 highlights the primary actor roles involved in *EOOSC System* development and exploitation.

On the left side there are the EOOSC **End-users**⁷ (cf. Sec. 2.1), i.e. the primary beneficiaries and consumers of the *EOOSC Services* (cf. Sec. 4). These actors⁸ include **Researchers**⁹ (i.e. those performing the research activities), **Research Administrators** (i.e. those in charge of the administrating and taking decision regarding the research activities in organizations and agencies), and **Third-Party Service Providers** (i.e. those that develop and operate new services by exploiting the EOOSC provided capabilities).

The right and main part of the figure introduces (a) the two major constituents of the *EOOSC System*, i.e. *EOOSC Services* and *EOOSC Service Components*, and (b) the roles involved in the *EOOSC System* development and operation, i.e. the *EOOSC System Managers* (cf. Sec. 2.1) and the EOOSC **Suppliers** (cf. Sec. 2.1).

EOOSC Services represent the *EOOSC System* offering, i.e. they are the components that, at a given point in time, implement the functionalities EOOSC is responsible for. *EOOSC Services* are expected to (a) be the artefacts of the *EOOSC Catalogue*, (b) closely adhere to the *EOOSC Principles of Engagement (PoE)* (cf. Sec. 3.2), (c) be operated by an *EOOSC Service Provider* (i.e. any service provider qualified as EOOSC service provider according to the rules established by the EOOSC Managers for delivering a target service), (d) be delivered with well defined *Service Level Agreement(s)* (either negotiated per-use or pre-defined across users). *EOOSC Service Providers* are expected to leverage *EOOSC Service Components* offered by *EOOSC Suppliers* according to *Underpinning Agreement(s) (UAs)* to realise their **added-value services** (cf. Sec. 3.3).

It is worth stressing that an *EOOSC Service* can be in a one to one relationship with an *EOOSC Service Component*. In this case the EOOSC Service is homologous (from the functional perspective) to the EOOSC Service Component yet it adheres to the *PoE* and it must conceptually add some value to the component service (e.g. non-functional aspects like certification, performance explicitly declared). Moreover, the provider called to operate/deliver an *EOOSC Service* and an *EOOSC Service Component* can be the same entity / organisation yet playing different roles (i.e. *EOOSC Service Provider* and *EOOSC Supplier* respectively) each characterised by diverse settings, duties, and implications. Following the same reasoning, an actor playing the role of *Third-Party Service Provider* can become an *EOOSC Supplier* or an *EOOSC Service Provider* if the service he/she develops and operates becomes an *EOOSC Service Component* or an *EOOSC Service* on its own.

The *EOOSC System Managers* are responsible for the *EOOSC System* development and operation. These include the *EOOSC System Owner* (a collective role implementing the EOOSC System Steering Committee), the *EOOSC System Top Manager* (a collective role implementing the EOOSC System Executive Committee), and, most importantly, the set of *EOOSC Service Providers*. Each of the *EOOSC Service Providers* is responsible for the management of the *EOOSC Service(s)* he/she operates. Overall, the managers are called to implement the decisions established by the EOOSC Governance boards [Hienola et al. 2017].

Regarding the service provisioning and its topology, this document is not suggesting neither structure nor allocations (e.g. global, regional, national) apart from envisaging the notion of *EOOSC Node* (cf. Sec. 3.4). EOOSC Nodes are the settings where the pieces (be them EOOSC Services of EOOSC Service Components) of the EOOSC system reside. The logics governing nodes identification are various ranging from service specific constraints and peculiarities (e.g. some services are single-node by design) up to community-specific deployment decisions (e.g. to have a nation-based set of services each serving users based on the specific region).

In addition to the *EOOSC Services*, the class of *EOOSC Compatible Services* is envisaged [Hienola et al. 2017] to capture the plethora of existing services not being (yet) fully fledged *EOOSC Services*, nevertheless,

⁷ This term might be problematic because of the fact that a specific case of End-user is expected to be a Third-party Service Provider, i.e. an actor role making use of EOOSC Services to serve others (the end-users of the developed service). It was decided to retain the term End-user because in the rest of the document it is used only as a container, activities and services are envisaged for the specific roles.

⁸ Human actors are primarily considered in envisaging the EOOSC actors' roles. However, the envisaged roles can be played by machines, e.g. sensors storing and publishing the collected data by EOOSC Services.

⁹ This term has to be read with the widest scope possible, it includes professional researchers as well as citizens.

considered of value for *EOOSC End-users*. From the perspective of this document, they are not expected to enter into the picture in a prominent way simply because the EOOSC System is not expected to offer any management or added value other than facilitating their discoverability and somehow attesting their “compatibility” with EOOSC. In fact, both *EOOSC Services* and *EOOSC Compatible Services* are expected to be discoverable by the **Service Catalogue**, an **EOOSC Core Service** (c.f. Sec. 4.1). Details on how this service is expected to work and how it is related with the EOOSC Portfolio and its management are given in deliverable EOSCPilot D5.2 “EOOSC Service Portfolio” [Andreozzi et al. 2018].

In order to serve the needs of the variety of actors mentioned above, the EOOSC system must include services (EOOSC Services and EOOSC Compatible Services) operating at **different levels of abstraction**. Figure 3 shows¹⁰ possible classes of services based on a generalization of the “as-a-Service” model as it appears in a number of articles in the literature, e.g. [Duan et al. 2015], and their conceivable exploitation scenarios both at EOOSC Service Use and at EOOSC Service Development (the circles represent where the service offered and consumed is). The way of reading this picture suggests that EOOSC Services, EOOSC Compatible Services and well as EOOSC Service Components can be at any level of abstraction from Infrastructure-facilities up to software, data, applications, etc. When analysing the exploitation scenarios from the perspective of End-users the cases range from that where EOOSC provides these users with typical IaaS facilities and they have to build and maintain the rest to cases where EOOSC is providing them with the entire solution. Similar cases emerge when using the EOOSC Service Development perspective: scenarios range from cases where an EOOSC Service Provider is counting on IaaS facilities operated by EOOSC Suppliers and has to build and develop on its own what is needed to operate the EOOSC Service to cases where the EOOSC Supplier is providing the entire facility to the EOOSC Service Provider.

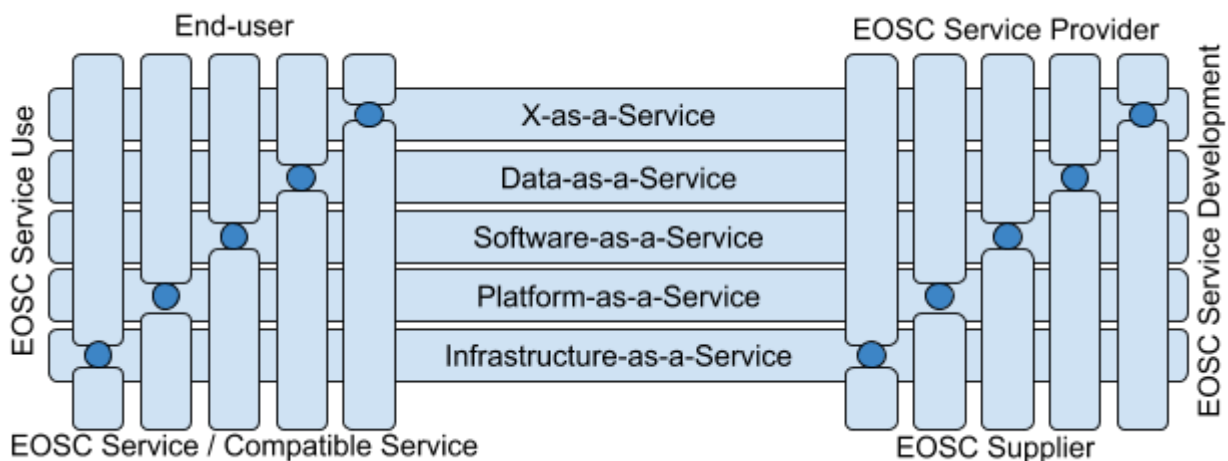


Figure 3. EOOSC Services stacking

The rest of the section is dedicated to illustrate the actor roles introduced above (cf. Sec. 2.1) and the set of activities that the EOOSC System should facilitate for each of them (cf. Sec. 2.2).

2.1. EOOSC System Users

Several **actors** are expected to exploit the functionality of the EOOSC system in order to meet different objectives. These objectives range from the exploitation of EOOSC services, to their provision and to the management of the EOOSC itself and of its services.

Each actor may play many **roles** as well as a role can be played by many actors (see Figure 4). We have named the most general of these roles **EOOSC System User**. For the sake of this document, the three major sub-roles of a User have been identified together with a number of further specialised sub-roles. The latter

¹⁰ A picture similar to this is in D5.2 too [Andreozzi et al. 2018] where Users is used instead of End-users and E-Infrastructure is used instead of EOOSC Service / Compatible Service.

are not meant to be an exhaustive list of all sub-roles but only a means to highlight the most frequently mentioned ones.

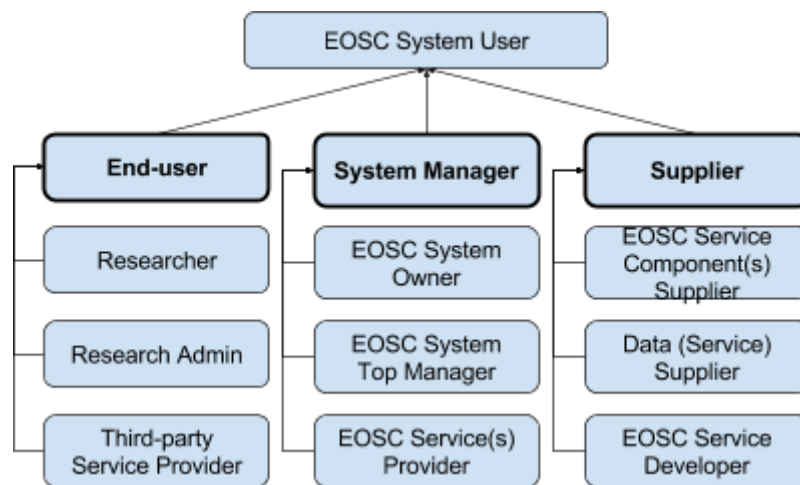


Figure 4. EOsc System Roles

End-user^{11,12} is the role of the actors that exploit the facilities offered by the EOsc system as a mean to support their working activities. Three major sub-roles of End-user are envisaged:

- **Researcher.** Actors exploiting EOsc facilities to accomplish science-related tasks. These actors range from professional researchers performing scientific activities by following an “open science” approach up to citizens willing to have access to scientific results. When exploiting the EOsc Service, these actors are actually both “consumers” and “producers” of EOsc, e.g. via EOsc Services they have access to certain data and they can publish other data. They can carry out the entire scientific workflow by themselves (i.e. from the idea conception, to data collection and collation, till the analysis and publication of results) or they can reuse existing artefacts and collaborate with others to achieve a common goal. In order to support these actors in meeting their objectives the EOsc system is required to offer services for: (i) favouring as much as possible the sharing of scientific artefacts, e.g. by facilitating publishing of any research artefact, the implementation of FAIR data management and the implementation of alternative research productivity metrics, (ii) simplifying the access and (re-)use of scientific results and functionalities that others have produced and shared. Several typologies of Researchers can be identified, including *Data scientists* and *Citizen scientists*.
- **Research Admin.** Actors who manage activities in research funding agencies or research performing organisations in order to meet research lifecycle goals, to enhance collaboration, to evaluate research impact from set objectives, to shape future research goals and to support informed policy making. They use EOsc services for collecting and monitoring information on research activities and their results, and combine them to assess research productivity for all actors involved, short and long term impact and trends. Among Research Admins several specialised profiles can be envisaged including the *Research Admin Liaison*, i.e. the person responsible for the organisation’s communication activities with respect to research and research services.
- **Third-party Service Provider.** Actors playing this role exploit the EOsc Services to develop and operate their own services. Such actors are expected to leverage the wealth of functionality offered

¹¹ As already clarified, the name of this role might be problematic because of the fact that one of the specializations is a Third-party Service Provider, i.e. someone making use of the EOsc Services to develop a new Service used by others. For the time being the term is retained because it is used only as a collector for roles somehow making use of EOsc Services.

¹² Such role can be played by human users as well as machine users, i.e. devices capable to connect with EOsc Services and exchange data with them.

by EOSC to simplify and make faster the implementation of innovative services dedicated to their designated users.

System Manager. This is a comprehensive umbrella role for all the actors that manage and operate EOSC and its services. Actors playing this role are called to “tune” the EOSC system thus to implement the decisions of the EOSC Governance [Hienola et al. 2017]. Actors with this role need facilities to efficiently and effectively perform their specific activities so to globally reduce the cost of the EOSC system operation and maintenance while preserving, at the same time, a high quality of service. Relevant sub-roles are:

- **EOSC System Owner.** This is the role played by the actors responsible for the development and maintenance of the EOSC system as a whole. This is expected to be a “collective role” assigned to a committee acting as the “steering committee” of the EOSC system. The committee is the primary responsible for making the EOSC System compliant with the decisions of the EOSC Governance by liaising with the EOSC System Top Manager(s);
- **EOSC System Top Manager.** This is the role played by the actors that are responsible for the continuous planning, implementation, and revision of the overall EOSC system. This is expected to be a “collective role” assigned to a committee acting as the “executive committee” of the EOSC System. The committee is responsible for putting in place the decisions of the EOSC System Owner by liaising with the EOSC Service Provider(s);
- **EOSC-Service Provider.** Actors playing this role are responsible for a specific *EOSC Service*. They are responsible for everything pertaining the development, operation and quality of the specific service including the establishment of the needed *underpinning agreements*. This documented agreement between the EOSC Service Provider and an EOSC Service Supplier specifies the underpinning (supporting) service to be provided by the EOSC Service Supplier, together with the related service targets (e.g. quality of service). Several actors with specialised profiles are expected to contribute to the provisioning of a service including, for example, *service operators*, i.e. system administrators of the IT system realising the service; (ii) *data curator / steward*, i.e. actors responsible for the data / content collated and made available by the service; (iii) *technology developer*, i.e. actors responsible for the development and maintenance of the technology supporting the service.

Supplier. Actors playing this role contribute their services to realise and operate the EOSC system. In order to reduce the time spent in performing this task they demand services for easily sharing their facilities through EOSC, making them accessible and checking the compatibility of the service policies with the EOSC ones. The sub-roles of this role include, among the others:

- **EOSC-Service-Component(s) Supplier.** Actors playing this role provide their own services to EOSC. By federating / integrating these services, EOSC Service Providers implement an *EOSC Service*. The nature and typology of the provided services is open ended; service components include functionality at all the levels of a technological stack, e.g. networks, virtual machines, software. Supplier stand thus for a large class of actors that may vary a lot according to the type of resource offered and the step of the provisioning step. These include: (i) *service operator*, i.e. system administrators of the IT system realising the service; (ii) *data curator / steward*, i.e. actors responsible for the data related objects collated and made available by the service, like datasets, workflows, publications and software tools; (iii) *technology developer*, i.e. actors responsible for the development and maintenance of the technology supporting the service. Moreover, a special case of supplier is the *one offering EOSC compatible services*. Actors playing this role are responsible for registering their services into the EOSC Service Catalogue to make them (at least) discoverable via the EOSC System. These actors are the suppliers of **EOSC Compatible Services**, i.e. services (i) existing independently on EOSC, (ii) worth being advertised to the widest “open science” community, (iii) on which EOSC is not yet planning to offer any added value facility apart facilitating the discovery and certifying their compatibility with the EOSC Principles of Engagement.
- **Data-(Service) Supplier.** This role is a specialisation of the **EOSC-Service-Component(s) Supplier**.

Actors with this role manage a Data Service that is mainly called to provide access to data (i.e. the data these services offer are the real “artefacts” that the users of the service are interested in). Data(-Service) can be fully compliant with the EOSC Principles of Engagement or they can be *Compatible*. Suppliers of these services register them into the EOSC Service Catalogue to make them (at least) findable via the EOSC System (cf. Sec. 4.1). This does not necessarily mean that the corresponding data are also made accessible through an EOSC Service. It depends on the Data Supplier access policies and on the underpinning that it has established with EOSC.

- ***EOSC-Service Developer***. This is the role played by the IT actors contributing the technology (the “entire” stack from the machine(s) and networking infrastructure(s) up to software) needed to operate an *EOSC Service*, namely the technology to combine selected EOSC Service Components and complement them to realize the service specific “added-value” (cf. Sec. 3.3).

2.2. Activities to be supported via the EOSC System

A necessary step to identify the EOSC classes of services is to clarify the needs of the different EOSC actors when playing the roles presented in the previous section. This section exemplifies functional requirements by presenting a number typical tasks performed by the EOSC users when playing the envisaged roles (cf. Sec. 2.1). Some of the envisaged actions are expected to be further clarified and expanded by other documents and activities (e.g. EOSC Service Portfolio Management is described in D5.2 [Andreozzi et al. 2018], Service Management will be described in D5.3).

2.2.1. Activities performed by Researchers

The usual activities researchers perform are related with all the phases of a typical research lifecycle. Broadly speaking, the major tasks they are called to do are related to (i) finding, accessing, linking and (re)using existing research outcomes to create new ones, (ii) publishing (make available) the artefacts they have produced, (iii) collaborating with others.

All these activities deal with primary entities: “research artifacts” and “individuals conducting science”. They act according to the semantics of these entities and the requirements related to their exploitation and management.

Typical research activities are sketched below. These reflect the current research practices and are expected to evolve in parallel with the progress towards the open science vision. A detailed description of the identified activities is in Sec. B.1. Envisaged activities include:

- ***Finding research artefacts***, i.e. the class of actions researcher perform to ***discover*** research outcomes of potential interest.
- ***Accessing research artefacts***, i.e. the class of actions researchers perform to ***obtain*** an artefact of potential interest.
- ***(Re-)using research artefacts/outcomes***, i.e. the class of actions Researchers perform in order to ***create*** new research outcomes by exploiting existing ones¹³.
- ***Depositing research artefacts***, i.e. the class of actions researchers perform in order to ***store and preserve*** the artefacts they own.
- ***Publishing research artefacts***, i.e. the class of actions researchers perform in order to ***communicate and make available*** their research outcomes to peers and to other stakeholders that may be interested in exploiting them.
- ***Collaborate to address a common challenge***, i.e. the class of actions researchers perform in order to engage others to ***jointly work*** on a research endeavour.

¹³ Artefacts expected to be reuse should be intended with an open-ended scope, they range from papers and data up to software, processes, and services operating at any level of the service stacking (cf. Figure 3).

- **Open and track requests for enhancements¹⁴**, i.e. the class of actions researchers perform in order to provide EOSC Managers with requests for changes wrt the currently offered services.

2.2.2. Activities performed by Research Admins

Research admins are involved in informing their organization leadership in advising on policies or IPR. In order to address their mission, they perform activities that aggregate research results for their organization in order to produce metrics and indicators allowing to assess the past (impact) and shape the future (trends).

In particular, the following activities are envisaged (detailed descriptions are in Sec. B.2 where for every action there is also the indication of the responsibilities assigned to each role involved):

- **Monitoring research outcomes**, i.e. the class of actions research administrators perform in order to be informed on the list of artefacts resulting from a research activity they promote and the impact these artefacts have produced (e.g. citation metrics, acknowledgments).
- **Assessing research impact**, i.e. the class of actions on research analytics that research admins perform in order to assess the impact of selected research artefacts.
- **Operating Communication hubs¹⁵**, i.e. the class of actions Research Admins perform in order to inform and communicate with Third-Party affiliates, as well as to encourage/promote the development of community-driven networks of researchers and general public/citizens.
- **Open and track requests for enhancements¹⁶**, i.e. the class of actions Research Admins perform in order to provide EOSC with requests for changes and/or additions with respect to the currently offered services.

2.2.3. Activities performed by Third-Party Service Providers

The activities Third-Party Service Providers deal with are mainly related with the development and operation of services that specifically address the needs of one or more designated communities by building upon one or more EOSC services. Such services are not natively conceived to be EOSC Services on its own yet, but they may contribute to the EOSC Catalogue.

In particular, the following activities are envisaged (detailed descriptions are in B.3 where for every action there is also the indication of the responsibilities assigned to each role involved):

- **Finding EOSC Services**, i.e. the class of actions Third-Party Service Providers perform to get acquainted with the EOSC System services that they might usefully exploit in developing their own.
- **Accessing EOSC Services**, i.e. the class of actions Third-Party Service Providers perform to get access to the EOSC Services they are interested in.
- **Using EOSC Services**, i.e. the class of actions Third-Party Service Providers perform to actually exploit the facilities offered by one or more EOSC Services.
- **Managing the SLA**, i.e. the class of actions Third-Party Service Providers perform to establish the Service Level Agreement regulating the provisioning of the target EOSC Service(s).
- **Providing feedback on EOSC Services¹⁷**, i.e. the class of actions Third-Party Service Providers perform to provide any EOSC Service Provider with feedback resulting from the concrete exploitation of the service in their own settings.
- **Open and track requests for enhancements¹⁸**, i.e. the class of actions Third-Party Service Providers perform in order to provide EOSC Managers with feedback and requests for changes wrt the currently offered services.

¹⁴ This is among the actions falling under the “Service Management” umbrella. It will be further developed in D5.3.

¹⁵ These are expected to be services and tools developed by the organisation to communicate with its designated community.

¹⁶ This is among the actions falling under the “Service Management” umbrella. It will be further developed in D5.3.

¹⁷ This is among the actions falling under the “Service Management” umbrella. It will be further developed in D5.3.

2.2.4. Activities performed by EOsc Suppliers

The activities that the **EOsc Suppliers** are expected to perform depend on the asset they make available to EOsc. According to the asset they may play different roles, e.g. **EOsc Service Component(s) Suppliers**, **Data (Service) Suppliers**, or **EOsc Service Developers**.

In particular, the following activities are envisaged (detailed descriptions are in Sec. B.4 where for every action there is also the indication of the responsibilities assigned to each role involved):

- **Managing the EOsc Service Component Provisioning**, i.e. the class of actions EOsc-Service-Component Suppliers perform to provide EOsc Service Providers with the target EOsc Service Component.
- **Publishing a future/possible EOsc Service Component**, i.e. the class of actions potential forthcoming EOsc Service Component Suppliers perform in order to advertise the service they own to the EOsc Managers.
- **Managing the Data (Service) Provisioning**, i.e. the class of actions Data (Service) Suppliers perform to provide EOsc Service Providers with their data¹⁹.
- **Managing the Underpinning Agreement**, i.e. the class of actions EOsc Service Component Suppliers perform to establish the agreement regulating the provisioning of their own services in the context of the provisioning of an EOsc Service.
- **Developing an EOsc Service**, i.e. the class of actions EOsc Service Developers perform to develop the technical solution underlying an EOsc Service.

2.2.5. Activities performed by EOsc System Managers

The activities **EOsc System Managers** are expected to perform via EOsc depends on their responsibility in managing the EOsc System, i.e. whether they are playing the role of **EOsc System Owner**, **EOsc System Top Manager**, or **EOsc Service Provider**. In the context of IT Service Management, a number of processes have been envisaged ranging from Service Portfolio Management and Service Level Management to Release & Deployment Management. For the sake of this deliverable we selected some of them and complemented the list of proposed activities with some EOsc specific ones.²⁰

In particular, the following activities are envisaged (detailed descriptions are in Sec. B.5 where for every action there is also the indication of the responsibilities assigned to each role involved):

- **Governing the EOsc system**, i.e. the class of actions EOsc System Owner(s) perform to develop, control and regulate the EOsc system as a whole.
- **Operating the EOsc system**, i.e. the class of actions EOsc System Top Manager(s) perform to make the EOsc system working as planned.
- **Managing the EOsc Portfolio**²¹, i.e. the class of actions EOsc Managers perform in order to define and maintain the EOsc Portfolio.
- **Managing the EOsc Service**²², i.e. the class of actions each EOsc Service Provider is called to perform in order to make a specific EOsc Service available for use and behaving like expected (e.g. being compliant with the established SLA).

¹⁸ This is among the actions falling under the “Service Management” umbrella. It will be further developed in D5.3.

¹⁹ Data is used here to mean any research artefact, including datasets, publications, software, workflows, maps, etc.

²⁰ The resulting list is expected to be revised in future versions of this deliverable once the outcomes resulting from T5.3 on Service Management and WP2 on EOsc Governance will be better defined and developed.

²¹ This activity is further described in D5.2 [Andreozzi et al. 2018].

²² This is a very broad class including many of the IT Service Management processes, e.g. Service Level Management, Service Reporting Management, Service Availability & Continuity Management, Capacity Management, Information Security Management, Customer Relationship Management, Supplier Relationship Management, Incident & Service Request Management, Problem Management, Configuration Management, Change Management, Release & Deployment Management. It will be detailed in D5.3.

- **Managing Catalogue(s)**²³, i.e. the class of actions EOsc Managers are called to perform in order to define and maintain a Catalogue service. Among them are worth mentioning the **EOsc Service Catalogue** (cf. Sec. 4.1) and the **EOsc Data Catalogue** (cf. Sec. 4.1).
- **Managing the requests for changes**, i.e. the class of actions EOsc Managers are called to perform to respond to the solicitations received from the End Users and stakeholders using the EOsc System and its Services [Hienola et al. 2017]).

²³ A Catalogue is a specific typology of Service particularly relevant in EOsc. Several catalogues are expected to be developed and operated by EOsc including the EOsc Service Catalogue listing EOsc Services, a Service Catalogue listing both EOsc Services and EOsc Compatible Services, a Data Catalogue listing datasets available in EOsc. Every catalogue should be conceived to serve the needs of a designated community (this might have implications on catalogue content and on catalogue views).

3. EOSC SYSTEM SETTINGS AND DESIGN DECISIONS

The design of the EOSC architecture is not only driven by the functional needs of the different actors listed in the previous section. As outlined in the introduction to this deliverable, there are many other factors that have to be taken into account in understanding “how” these functionalities can be provided. These concern primarily aspects like the use of services provided by existing initiatives (e.g. Research Infrastructures²⁴, service providers, data centers), maintenance costs, sustainability, expected evolution, business models, governance and so on.

The entire EOSC architecture presented in this deliverable is based on the assumption that facilities are provisioned **as-a-Service**, i.e. an approach in which a functionality is made available by an online service operated by a provider (***EOSC Service Provider***). The technical and organisational approaches needed for the service to deliver the functionality are completely up to the *service provider*. As a consequence of this choice, a key concept underlying this approach is the ***EOSC Service***, i.e. the ways to provide value to EOSC users through bringing about the results that they want to achieve.

Documents produced by high level Experts (cf. [Ayrís et al. 2016], [EC 2017]) have highlighted a general consensus on the necessity and opportunity of building EOSC by leveraging and integrating existing resources at European and Member State levels.

This consideration suggests to model EOSC as an ***open and evolving System of Systems*** (cf. [Maier 1996]) where component systems, e.g. Research Infrastructures and other existing and emerging Service Providers, supply services in accordance with the ***EOSC Principles of Engagement*** for Service Providers.

The provisioning of each EOSC service is regulated by well defined agreements (namely ***Service Level Agreements***²⁵ and ***Terms of Use*** for EOSC Services, ***Underpinning Agreements*** for supplied EOSC Service Components). These conditions must always be compliant with the EOSC established principles of engagement and may include other context or service specific constraints.

The added-value introduced by an EOSC Service on top of the exploited / federated Service Component(s) of the SoS is diversified and depends on the added-value mode (cf. Sec. 3.3) implemented by the Service Provider. It can range from a simple support for service compliance with the established PoE to a complex service adding relevant functional value to the federated services.

Below all the aspects mentioned above are discussed in more detail highlighting in particular their implications on specific architecture model choices.

3.1. System of Systems

In order to clarify the notion of System-of-Systems (SoS) below two characterizations ([Maier 1996] and [Boardman and Sausser 2016]) that are commonly exploited to define the distinguishing features of Systems of Systems are reported. These indicate properties of the component systems with respect to the overall system, e.g. the EOSC system in our case.

Maier’s characterisation envisages the following features:

²⁴ Research infrastructures are facilities, resources and services that are used by the research communities to conduct research and foster innovation in their fields. Where relevant, they may be used beyond research, e.g. for education or public services. They include: major scientific equipment (or sets of instruments); knowledge-based resources such as collections, archives or scientific data; **e-infrastructures, such as data and computing systems and communication networks**; and any other infrastructure of a unique nature essential to achieve excellence in research and innovation. Such infrastructures may be 'single-sited', 'virtual' or 'distributed'. (EC’s definition according to: http://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-infrastructures_en.pdf)

²⁵ In some cases, the Service Level Agreement is independent of the specific user of the service, i.e. the SLA is self-defined by the EOSC Service Provider to explicitly characterize the quality, availability, responsibility, and any other aspect qualifying the provisioning of the service. The presence of SLAs is key to enact an informed use of the service.

- **Operational Independence:** if the system-of-systems is disassembled into its component systems, the component systems must be able to usefully operate independently;
- **Managerial Independence:** the component systems are separately acquired and integrated but maintain a continuing operational existence independent of the system-of-systems;
- **Evolutionary Development:** the system-of-systems does not appear fully formed. Its development and existence is evolutionary with functions and purposes added, removed, and modified with experience;
- **Emergent Behavior:** the system performs functions and carries out purposes that do not reside in any component system. These behaviors are emergent properties of the entire system-of-systems and cannot be localized to any component system. The principal purposes of the systems-of-systems are fulfilled by these behaviors;
- **Geographic Distribution:** the geographic extent of the component systems is large. Large is a nebulous and relative concept as communication capabilities increase, but at a minimum it means that the components can readily exchange only information and not substantial quantities of mass or energy.

In addition to that, Maier envisaged three possible management models:

- **Directed:** “Directed systems are those in which the integrated system-of-systems is built and managed to fulfill specific purposes. **It is centrally managed** during long term operation to continue to fulfill those purposes, and any new ones the system owners may wish to address. The component systems maintain an ability to operate independently, but their normal operational mode is subordinated to the central managed purpose”;
- **Collaborative:** “Collaborative systems are distinct from directed systems in that the **central management organization does not have coercive power to run the system**. The component systems must, more or less, voluntarily collaborate to fulfill the agreed upon central purposes. The Internet is a collaborative system. The Internet Engineering Task Force works out standards, but has no power to enforce them. Agreements among the central players on service provision and rejection provide what enforcement mechanism there is to maintain standards. The Internet began as a directed system, controlled by the US Advanced Research Projects Agency, to share computer resources. Over time it has evolved from central control through unplanned collaborative mechanisms.”;
- **Virtual:** “Virtual systems lack a central management authority. Indeed, they lack a centrally agreed upon purpose for the system-of-systems. Large scale behavior emerges, and may be desirable, but the supersystem must rely upon relatively invisible mechanisms to maintain it”.

At this stage of the EOsc system definition and of its governance we envisage that **it falls in the collaborative management model**.

In essence, by adopting the SoS model just presented and applying the two characterizations it emerges that:

- the EOsc system is characterised by **constituent systems that are autonomous, independent and distributed**;
- the EOsc system **is exposing an emergent and evolving behaviour**, i.e. it delivers higher functionality than that delivered by the constituents and new/revised functionality may be developed during the EOsc lifetime;
- the **existence of EOsc induces** a sort of **interdependency among the constituent systems** when delivering the EOsc planned services (i.e. the emerging behavior, the services in the EOsc catalogue).

Interdependency and interdependency among the systems is the added value of a SoS. It does not only allow to exploit existing systems without requiring their refactorization. Rather it allows to integrate, promote, and deliver tailored solutions to the different disciplinary sectors in a coordinated fashion while preserving the specificities and emphasizing the level of interoperability among them.

An external “system”, operating independently to meet its own objectives, has to meet a number of conditions in order to become an EOSC supplier for part, or all, of its services including those envisaged in the EOSC Principles of Engagement (cf. Sec. 3.2). In particular, it has to sign, for each of the supplied services, an “Underpinning Agreement” with EOSC. The conditions reported in this agreement have to be consistent with the EOSC Principles of Engagement (cf. Sec. 3.2) that define the properties to be satisfied by all the EOSC components.

3.2. EOSC Principles of Engagement and Governance

The ***EOSC Principles of Engagement*** (PoE) define the umbrella for the ***policies, processes*** and ***procedures*** (be them organizational and technical) governing systems integration, service provisioning and service usage that have to be satisfied by every actor interfacing with the EOSC system when playing an envisaged role (cf. Sec. 2.1).

These principles are currently under definition as part of Task 2.3.1 “Investigation and analysis of organisational Rules of Engagement for EOSC” and Task 5.3. “EOSC Federated Service Management Framework”. In the current planning, they primarily concern EOSC Service Component Suppliers²⁶ and End-users.

Regarding the EOSC Service ***Component Supplier***, these principles may concern the suppliers as such, e.g. they can be requested to be certified actors, or concern ***the services that they offer***, e.g. these can be requested to guarantee a given service level, or be ***service-specific***, e.g. any data repository must expose its content metadata using a format compatible with EDM I guidelines²⁷.

The entire EOSC system operates assuming that the conditions established by these principles are satisfied. These conditions may have a “scope”, i.e. they may concern actors with certain roles or certain services. The selection of these rules has a strong implication on the services that are required to support the EOSC managers in charge to verify and monitor them. As a matter of fact, in order to guarantee that the conditions implicit with principles are indeed met, the EOSC system has to include appropriate ***registration, validation and monitoring of services***. For example, if it will be decided that only data repository services meeting certain FAIR criteria can be part of EOSC, the managers will have to be provided with services that enable to validate and certify these criteria before being registered in a registry of certified data repositories.

The extent that validation and monitoring services will be able to guarantee that the established principles are indeed met will largely depend on the form that these principles will take and on the advances that the research in the area of machine processable descriptions of these conditions will have.

Any EOSC Service Component Supplier signs an ***underpinning agreement*** with EOSC (actually, implicitly with every EOSC Service Provider). This agreement, that may specify more specialized conditions with respect to the more general EOSC principles of engagement, regulates in more detail “how” the service will be provided. It is envisaged that the EOSC management makes available a set of pre-defined agreements (sort of “agreement templates”) enabling constituent systems to choose among them according to a cost-benefit function.

²⁶ For the purpose of this document the need to distinguish among two concepts that we have named “supplier” and “provider” has emerged. The same need has not emerged in Task 5.3 that is using a single term “provider” to mean both.

²⁷ EDM I (EOSC Dataset Minimum Information) is a simple metadata guideline defining the minimum metadata properties that a data service should expose in their data models and programmatic interface. The EDM I guideline is being developed by EOSCpilot Task 6.2 EOSC Research and Data interoperability. For more information see <https://tinyurl.com/eoscpilot-d63>

The EOSC principles of engagement for *EOSC End-users* (cf. Sec. 2.1), instead, regulate the exploitation of the EOSC provided services. The generic form of these principles is given in terms of (i) the consumer, (ii) the services to be exploited and (iii) the possible usage of the services. Terms of use are usually part of a Service Level Agreement. Also in this case, there are different levels of conditions, from **general ones that apply to all the users**, e.g. any registered user can use EOSC services public interfaces, to **more specific ones acting on specific classes of services or on specific actions**, e.g. only users authorised by a specific VRE manager can access the VRE or only service suppliers or authorised data curators can modify the content of the service Catalogue.

The definition of the Principles of Engagement will be an iterative process as it will be for the definition of the architecture. Clearly, the more specific the conditions imposed by these principles will be the easier it will be to support interoperability. However, it is clear that in the current situation not all the potential component “systems” of the SoS, and hence the potential EOSC Service Suppliers, are willing/ready to accept strict conditions. The plan is thus to start with light conditions and add more specific ones if/when the different components will be ready to accept/implement them.

As the Principles of Engagement are not yet consolidated at the time of the writing of this deliverable this initial version of the architecture includes only “placeholder” service types for validation and monitoring of underpinning and service level agreements. The nature of these services will be better specified when D5.3 “EOSC Federated Service Management Framework” (M18) and D2.5 “EOSC Rules of Engagement” (M19) will be released.

3.3. EOSC Service Added-Value Federation Models

The activities illustrated in Sec. 2.2 have to be interpreted in the context of an environment dedicated to support Open Science. This means, for example, that they may be performed collaboratively by multiple actors or they may concern resources that span across multiple domains service suppliers. For example, researchers may require to be able to combine and run different state-of-the-art models offered by different EOSC Service Suppliers or to search data across repositories provided by EOSC Service Suppliers of different domains. These functionalities are hardly made natively available by the Suppliers of the EOSC SoS. Actually, the majority of existing services supporting researchers’ data-centered activities respond to the needs of communities operating in the same domain. EOSC Service providers are thus required, as part of their federation role, to provide these cross-domain functionalities as “added value” building on what the existing different EOSC Service Suppliers can offer.

The **cross-domain functionalities** mentioned above are just an example of the supplementary value that the EOSC system can supply when observed as **federator of existing systems**. The extent and nature of added value may vary a lot, from the simple “certification of compliance” with respect to EOSC principles of engagement introduced in Sec. 3.2 or the “easy discoverability” of the service up to more complex “functional enhancements” as those exemplified to serve cross-domain and cross-sector applications.

Within the broad vision of the EOSC, there are many different service types and many modes of delivery. Added value can therefore be provided in many different ways. Being inspired by the FedSM federation business models [Appleton 2012] we highlight below some of possible EOSC added-value roles that may be appropriate for some services:

- **Invisible coordinator** (cf. Fig. 5): The *EOSC Service Provider* makes available a user functionality that is natively provided by a *Service Supplier* by an *EOSC Service Component*. The added-value perceived by the user with respect to the original functionality is mainly the transparency with respect to the supplier. Other “behind the scene” functionalities may be added, e.g. the EOSC Service may add value for facilitating the integration of the service in the system-of-systems or for simplifying its management (e.g. monitoring and checking the satisfaction of the principles of engagement or supporting its discoverability).



Figure 5. Pattern of an EOOSC Service implementing the Invisible coordinator model

- **Matchmaker** (cf. Fig. 6): the EOOSC Service Provider receives from the user the request for a specific functionality. It matches the request with respect to the multiple EOOSC Service Suppliers that can offer it and dispatch it to the “best option” with respect to established criteria. Matchmaking and dispatching criteria do not only take into account the requested functionality but also other non-functional criteria, like load balancing, proximity, costs, etc.

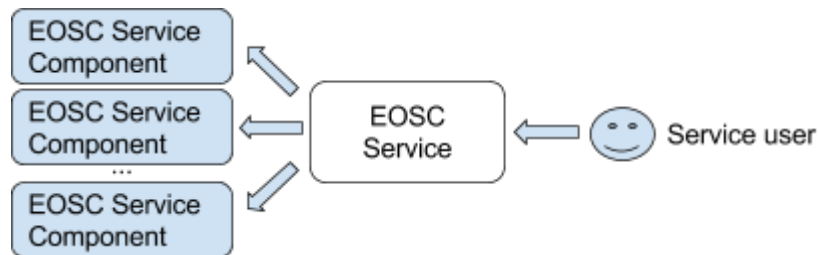


Figure 6. Pattern of an EOOSC Service implementing the Matchmaker model

- **One-stop-shop** (cf. Fig. 7): The EOOSC Service Provider acts as single entry-point for a functionality that it distributes across different Service Suppliers. The distribution is transparent to the user as all the details related to the heterogeneity that may exist behind it (e.g. protocols, data types, SLAs). All these aspects are managed and resolved by the business logic of the specific EOOSC Service as part of the added-value.

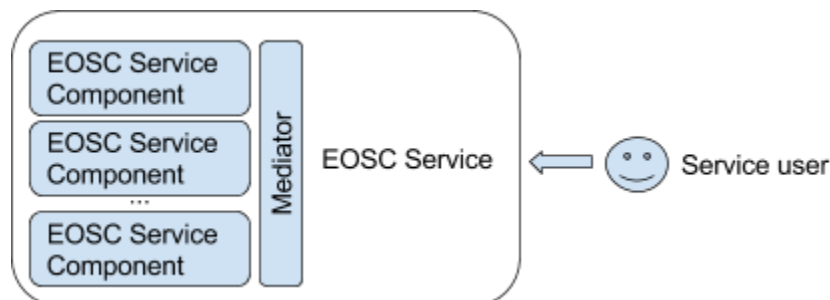


Figure 7. Pattern of an EOOSC Service implementing the One-stop-shop model

It is important to highlight that the added value is “*service specific*” meaning that EOOSC is not forced to adopt a unique federation model, but can make different choices for different typologies of EOOSC provided services.

The choice of the model to be implemented for each service has relevant implications for EOOSC as such and for the underlying IT system. As a matter of fact, the selected models contribute to define the business role of EOOSC as federator of resources (the added-value is the capability that the EOOSC system offer). Hence, the choice of the federation models to be supported is a governance decision that has to be taken by EOOSC governance/executive boards.

Regarding the IT system, every choice implies the demand for specific “mediator/orchestrator” services implementing the required added-value. Detailed indications on the nature of these services cannot be provided at the moment since no decision on the federation model for each identified typology of services has been taken yet. These services, generically named “**Mediators**”, are included in Sec. 2.2.1-2.2.4 as place-holders for the variety of them that will be introduced when more concrete versions of the EOOSC Architecture will be produced.

3.4. Distributed Service Provisioning

According to the EOsc Declaration [EC 2017] “... the EOsc will federate existing resources across national data centres, European e-infrastructures and research infrastructures; service provision will be based on local-to-central subsidiarity (e.g. national and disciplinary nodes connected to nodes of pan-European level); ...”. To reconcile this expectation with the model characterising the EOsc System (cf. Sec. 2) the notion of **EOsc Node** should be introduced and clarified.

EOsc Nodes are the “organisational pieces” of the EOsc System called to contribute to the provisioning of one or more EOsc Services. Because of this, EOsc Nodes are:

- the “places” / “settings” where the EOsc Services and/or the EOsc Service Component reside;
- operated by an EOsc Service Provider and/or EOsc-Service-Component Supplier/Data(-Service) Supplier.

In the majority of cases, EOsc Nodes are expected to correspond to existing “systems” (EOsc is a System of Systems). However, and depending on per-EOsc Service design and deployment decisions, new nodes can be developed.

3.5. FAIRness

The FAIR principles [Wilkinson et al. 2016] have been proposed to make all scholarly output *Findable, Accessible, Interoperable, and Reusable*. In the reality such principles focus on “data” and they describe “expectations” / “requirements” more than “implementations” options. These principles are commonly referred in Open Science discussions.

Since EOsc is called to provide its users with instruments for meeting the Open Science challenge, it must provide solutions for meeting the FAIR principles. This has several implications including:

- EOsc Services should be developed to guarantee that the data²⁸ managed by them adhere to the FAIR principles;
- Specific EOsc Services should be conceived, developed and provisioned to support EOsc Users implementing FAIR principles (e.g. the *Repository* service in Sec. 4.2);
- EOsc Services should be FAIR on its own, i.e. they should be developed thus to be *Findable, Accessible, Interoperable and Reusable*.

²⁸ Data is used here with the widest meaning possible, in essence it is including any typology of data going to be managed via EOsc Services.

4. EOSC SYSTEM COMPONENTS

As highlighted at the beginning of this document the EOSC system is intended to support and facilitate all the actors involved in implementing a FAIR data management and more in general an open approach to science.

This section presents an initial set of service classes that the EOSC System is expected to provide to serve the needs of the variety of users (cf. Sec. 2.1) involved in implementing and supporting the Open Science approach.

It is important to note that the list of services reported at this stage of the EOSCpilot architecture definition has to be interpreted as an *initial, in-progress one*. Other services are expected to be included in successive versions of the architecture to reflect the progresses that will emerge on the many fronts involved in the EOSC development, ranging from the variety of existing services that the various communities will decide to make available to EOSC, to the technological advances and to the change in the research practices that the Open Science approach will stimulate.

The current list of services has been discussed so far by exploiting the collaboration with the shepherds of the EOSCpilot Science Demonstrators, with the stakeholders of the Research Infrastructures represented in the EOSCpilot project and with partners of other related work packages (e.g. WP2, WP6, WP7).

The work done so far has been aimed only to identify the set of necessary typologies of services. As the provision model is “as-a-Service” (cf. Sec. 3), for each identified service (or class of services) there might be two diverse typologies: the *factory* and the *instance*. In essence, the factory provides functionality for the creation and monitoring of the service instance (tailored to serve the client/user needs under a well defined SLA), while the instance is intended for the real consumption of the facility.

Note that at this stage the set of service types and the internal architecture of these services is not specified yet. It will largely depend on the added-value federation model chosen.

Finally, it is important to take into account that at this stage service types are identified but it is not yet specified who (and how many) will be the underlying *Service Provider* that will operate and deliver these services. The EOSC services might be offered by or procured to existing and emerging components of the SoS (e.g. Research Infrastructures, Data Centers, Service Providers) or they might be partially implemented and operated by an “EOSC entity” (whatever form it will take). It is also likely that the same, or similar, EOSC Service is offered by multiple providers (e.g. peer and federated services created for safety reasons). The choice between these options is *beyond the scope of this first release of the EOSC Architecture deliverable*. The deliverable is simply meant to support the future deployment plan by highlighting what are the services *required to support the EOSC system expected functionality*.

Service Categories

Figure 8 illustrates the classes of services that are presented in more details in the following subsections. Service classes are organized per role. In addition, there are a number of Core Services classes that are needed across all roles. The presentation below starts with the latter.

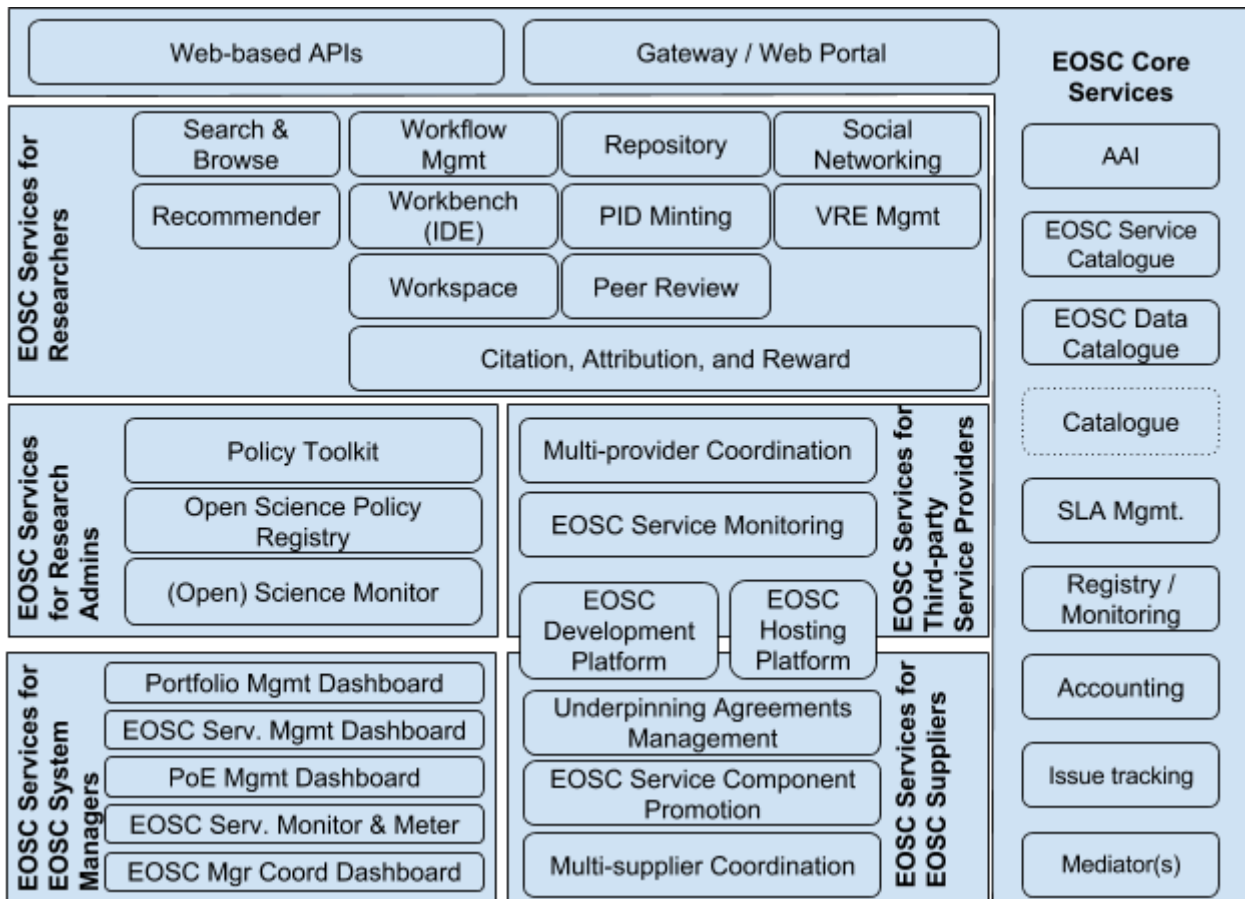


Figure 8. EOsc Services

4.1. EOsc Core Services

There are some functionalities that respond to needs that are common to any user, independently from her role. The functionality can be exactly the same for each role, e.g. Authentication, or it can correspond to different instantiations, each addressing the specific view of the resources manipulated by each actor role, e.g. Service Catalogue. In structuring the EOsc architecture we collectively name the services providing this functionality “EOsc Core Services”. A basic set of them is depicted in Fig. 9.

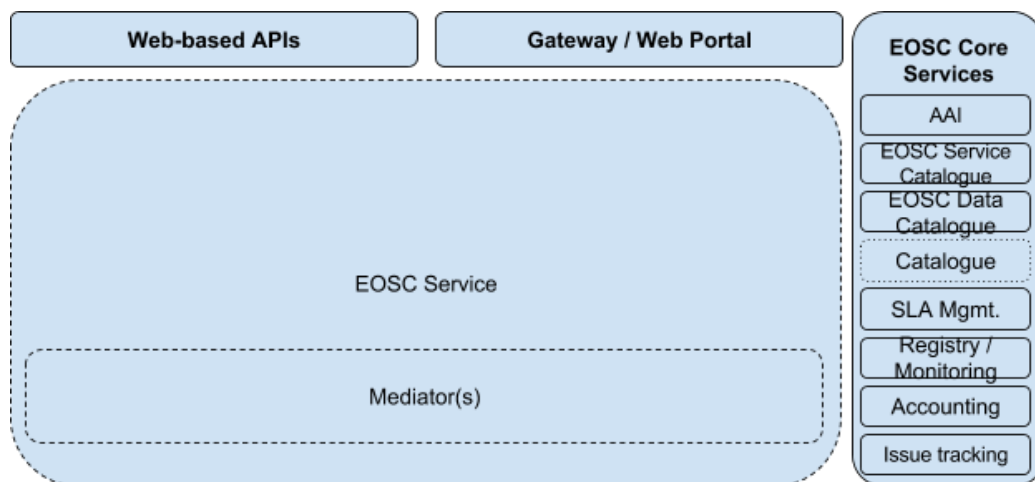


Figure 9. EOsc Core Services

AAI (Authentication and Authorization Infrastructure) is the approach EOsc System put in place for authorization and access control. At a conceptual level, this is a centralised system, however, it must be

implemented in a distributed way that guarantees that the existing solutions for authentication and authorization remain available and become interoperable.

Examples of existing services possibly falling in this class and/or exploitable to realize services in this class include eduGAIN, ELIXIR AAI, OAuth-based services, B2Access, [INDIGO IAM](#) and EGI Check-in.

EOSC Service Catalogue is a service offering a comprehensive map of the EOSC Services.²⁹ It supports seamlessly publishing and description of the EOSC Services so that they become easily discoverable and, to the most extent that it is possible, also FAIR (cf. Sec. 3.5).

Conceptually each actor role needs its own, tailored, service catalogue. This is because each actor when playing a specific role is actually interested in the subset of the EOSC services (both EOSC Services and EOSC Compatible Services) and in descriptions that meet its needs. These different service catalogues can be implemented through different catalogue services or through a single service offering different personalized views on the content and on rich, multi-faceted and multi-purpose metadata³⁰. Whatever solutions are chosen, every entry into the catalogue is required to have a unique and persistent identifier in order to support FAIR principles. Moreover, service catalogues can be developed by the one-stop-shop federator role (cf. Sec. 3.3), i.e. they are expected to federate existing service catalogues and provide the user with an homogeneous view of the existing services across the federated catalogues.

Examples of existing services exploitable to realize this typology of catalogues are the ones that are under development in the context of the eInfraCentral Project (einfracentral.eu), EGI (<https://www.egi.eu>), GEANT (<https://www.geant.org>), and EUDAT (<https://sp.eudat.eu>).

EOSC Data Catalogue, is a Service supporting findability of the datasets made available by the EOSC Services. Given the scope and rapid growth in the number of datasets available through the EOSC, such a catalogue will need to be automatically maintained. This will require the widespread adoption of standards enabling automatic cataloguing.

SLA Management is a Service enabling, where appropriate, to define and record the per-Service agreements characterising their provisioning. For some services there is only one agreement, for other services there may be an agreement per user. It may be the case that some well-established externally provided services may be made available through the EOSC with no specific agreement on service level. In this case default ones, in line with EOSC Principles of Engagement, will have to be adopted.

Registry is a Service maintaining records on EOSC Services for monitoring purposes.

Accounting is a Service recording the consumption of capabilities by users.

Issue tracking is a Service enabling every EOSC System User to create “tickets” documenting occurred issues and their resolutions. By exploiting this service, the various roles can easily interact according to their responsibilities.

Examples of existing service belonging to this class are the ticketing systems like Redmine and online support systems.

In addition to these “core” facilities, there are also three classes of “cross-cutting” services the EOSC System is offering for presentation / access and interoperation.

The EOSC can be access either through programmatic API’s or directly through a GUI:

Web-based APIs are the means EOSC Services are offered for programmatic exploitation, and

Gateway / Web portal is a web-based GUI entry point. The level of integration of the various Services in this GUI may vary a lot.

²⁹ Other Service Catalogues are likely to be offered by EOSC in addition to the EOSC Service Catalogue. The service catalogue has not to be confused with the Service Portfolio.

³⁰ Notice that the service catalogue is expected to include also data repository services. This means that this catalogue can also cover the role of data repository registry.

Mediators are set of services developed to support intermediation among other services, i.e. the latter no longer communicate directly with each other, but instead communicate through the mediator. Mediators in this context are used to make transparent the heterogeneity of certain service interfaces towards a common one, e.g. hiding the heterogeneity of metadata in different formats by exposing them through a pivot format. Even if they are developed to satisfy needs in specific contexts, it may be useful to share them to avoid replicating interoperability efforts, e.g. when they intermediate among two different standards. This class is almost open-ended.

4.2. Services for Researchers

Figure 10 illustrates a set of services supporting the Researchers' activities described in Section 2.2.1³¹. Such services are derived starting from the identified activities and taking into account major aspects that characterize the European Open Science Cloud as an environment for facilitating cross-domain research, i.e. enabling access to services for the *creation, deposition, processing, validation, interlinking, enrichment, of research artifacts to enable reuse* (which is translated in the request of managing these artifacts according to the FAIR principles) and availability of services for *collaboration across domains, organizations, and along the entire scientific workflows*. The requirements that derive from these two aspects are not always disjoint, e.g. recommending latest research results that meet an individual's interests is a way to facilitate sharing but also a way to facilitate collaboration.

As already highlighted these services operate understanding as primary entities "research artifacts" and "individuals making science". They offer functionality according to the semantics of these entities and the requirement related to their exploitation and management.

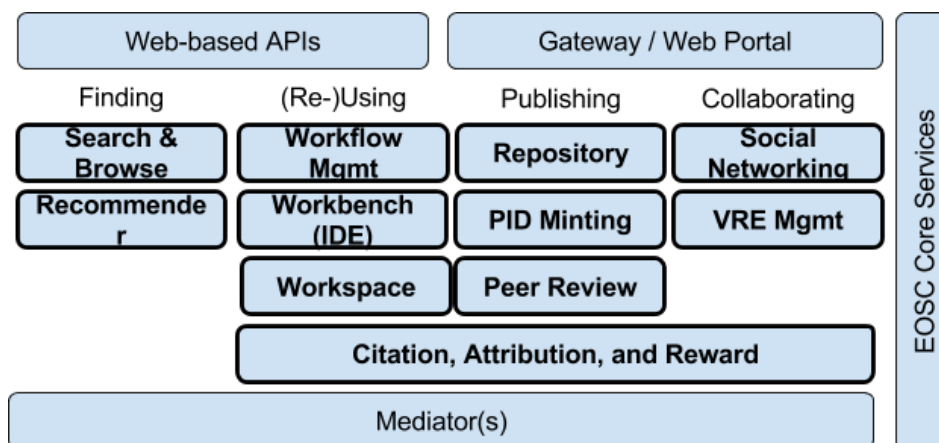


Figure 10. EOSC Services for Researchers

Search & Browse is a class of services envisaged to enable Researchers to find amongst the available research artefacts (e.g. datasets, papers, methods, workflows) those matching their interests. This class shall include, at least, a service dedicated to support the needs of a Researcher willing to seamlessly search over the entire assets space made available via EOSC.

Examples of existing services belonging to this class are the Google-like service, [EUDAT B2Find](#), and the OpenAIRE Search.

Recommender is a class of services envisaged to provide Researchers with objects of potential interest by predicting their preferences. This class shall include at least a cross-cutting recommender service aiming at suggesting suitable artefacts existing in the whole artefacts space.

³¹ Please notice that in this figure and in the ones that will follow the layering between services is for graphical reason, it is not meant to represent service layering.

Workflow Management is a class of services envisaged to enable Researchers to define and execute scientific workflows, i.e. series of data manipulation and computation steps.

Examples of existing services possibly falling in this class and/or exploitable to realize services in this class include WFMS (e.g. [Galaxy](#), [KNIME](#), Taverna).

Workbench is a class of services envisaged to enable Researchers to build new research artefacts out of their research activities. The workbenches Researchers are looking for range from environments enacting the collaborative editing of papers, algorithms, programs, notebooks, protocols, etc. up to advanced working environments enacting them to (re-)use and combine artefacts of every genre.

Examples of existing services belonging to this class are the collaborative editing environments (e.g. [ShareLaTeX](#)), the code development environments (e.g. RStudio, Eclipse), and the annotation environments (e.g. [Hypothesis](#)).

Workspace is a class of services envisaged to enable Researchers to conveniently store own artefacts. The access to this storage space can be granted to groups of actors. This enable sharing of artefacts with selected co-workers.

Citation, Attribution, and Reward is a class of services taking care of promoting and supporting actions aiming at maximising the citation (e.g. automatically keeping track and producing citation statements whenever an artefact is used in the context of a new artefact), attribution (e.g. automatically recording the specific contribution every single researcher is giving in the context of the production of a given artefact) and reward (e.g. collecting indicators on the impact and contribution per researcher).

Repository is a class of services taking care of the long-term storage and availability of the artefacts deposited. Specific service instances exist depending on the typologies of artefacts they support (e.g. paper vs data vs software vs mixed repository, pdf paper repository vs “beyond paper” repository), the coverage (e.g. domain specific vs generalist), the preservation practices, and other features making the homologues services diverse.

Examples of existing services possibly falling in this class and/or exploitable to realize services in this class include generalist repositories (e.g. Zenodo, b2share, figshare, Dryad), domain-specific repositories (e.g. EMBL-EBI data repository), software repositories (e.g. CPAN, CRAN, git/github).

PID Minting is a class of services supporting the association of persistent identifiers to research artefacts. This facility might be paired with the Repository facility discussed above.

Examples of existing services belonging to this class are [B2Handle](#) (ePIC PID) and DataCite DOI.

Social Networking is a class of services supporting social networking interaction patterns in research contexts, e.g. the usage of posts to announce the availability of an artefact, the usage of like or comments to give positive or negative feedback to published artefacts, etc.

VRE Management is a class of services supporting the creation and management of virtual web accessible working environments dedicated to designated communities of cases.

An example of existing service belonging to this class is the D4Science VRE Management System.

4.3. Services for Research Admins

Figure 11 illustrates a set of services that are envisaged to support Research Administrators’ activities described in Section 2.2.2.

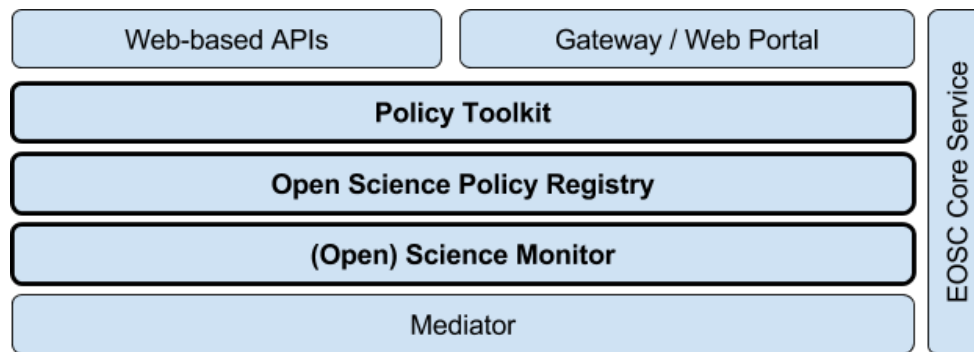


Figure 11. EOsc Services for Research Admins

Policy Toolkit is a collection of services, part of a larger set of instruments that include also resources, tools, and approaches, supporting the management of various policy issues (e.g. Intellectual Property Rights (IPR) and exploitation, Technical provisions regarding machine-readability, interoperability issues) as interpreted within the scoping and/or implementation procedures of different funders, publishers, libraries and other research and community stakeholders. The policy toolkit offers a comprehensive collection of functionalities that can be used to establish machine-readability of policies.

Open Science Policy Registry is a space where policies relating to aspects of open science can be submitted, assessed, and stored. The Registry provides combined facilities to store machine-readable policies, evaluate compliance with EOsc principles of engagement, and produce metrics and reports in support of the OS Monitor. It supports the EOsc externally and internally: as a service for users to assess EOsc compliance and compatibility and as a reporting tool to produce metrics on the progress of open science in Europe.

An example of existing services belonging to this class are the JISC SHERPA/ROMEO services for journals and funders policies.

(Open) Science Monitor is a set of services for supporting Research Performing Organisations (RPOs), Research Funding Organisations (RFOs) and Government Bodies to measure:

- levels of compliance with European Union’s laws, regulations and policies regarding research and research results dissemination;
- Open Science Resources’ (i.e. research artefacts, educational resources, research collaboration, citizen science) levels of openness, trustworthiness and FAIRness that cover each stage of the research lifecycle;
- excellence of science, which includes quantitative and qualitative metrics of different levels (bibliometrics, webometrics, scientometrics, etc);
- impact of science on society and economy.

Examples of existing services belonging to this class are JISC Monitor, consisting of Monitor UK which focuses in reporting APC’s and Monitor local for compliance, and OpenAIRE Gold for FP7 activities.

4.4. Services for Third-party Service Providers

Figure 12 illustrates a set of services that are envisaged to support Third-Party Service Providers’ activities described in Section 2.2.3.

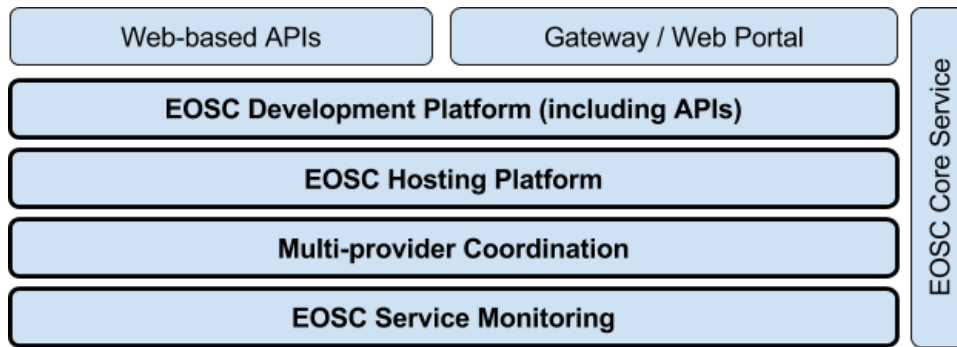


Figure 12. EOSC Services for Third-party Service Providers

EOSC Development Platform (including APIs) is a set of services implementing a platform simplifying the development of services exploiting EOSC Services. This platform should be (a) open and extensible to easily add new facilities made available whenever a new EOSC Services start appearing and (b) cater for diverse exploitation models, e.g. to make it possible for user to use only the subset of facilities they need. These services should add value on top of the web-based APIs, e.g. it should offer “abstractions” on top of the APIs.

EOSC Hosting Platform is a set of services enabling the deployment of service developed by Third-Party Service Providers onto EOSC nodes (cf. Sec. 3.4).

Multi-provider Coordination is a set of services aiming at coordinate the set of EOSC Services (and their providers) exploited by an Third-Party Service Provider in the context of a specific service.

EOSC Service Monitoring is a set of services enacting an EOSC Service Provider to monitor the EOSC Services he/she is willing to exploit/exploiting to implement its own service. These services realise a dashboard Third-Party Service Providers can configure to control the compliance of the target EOSC Services behaviour with respect to the agreed SLAs as well as any other constraint worth monitoring for the health of service the Third-Party Service Provider is responsible for.

Examples of existing services belonging to this this class are the one available in ARGO (<http://argoeu.github.io/>).

4.5. Services for EOSC Suppliers

Figure 13 illustrates a set of services that are envisaged to support EOSC Suppliers’ activities described in Section 2.2.4.

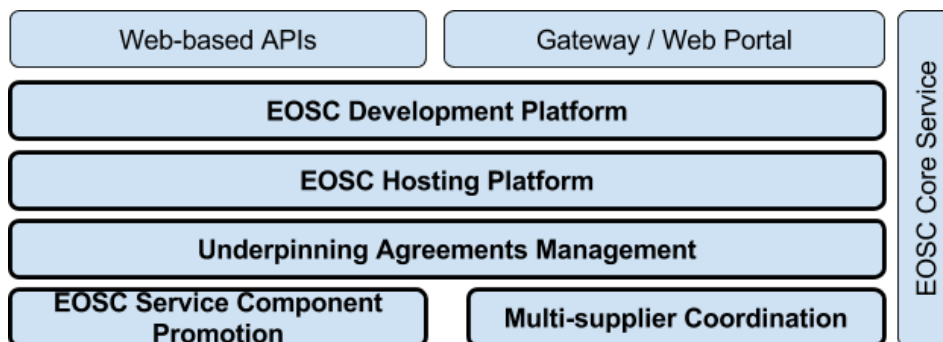


Figure 13. EOSC Services for EOSC Suppliers

EOSC Development Platform is a set of services implementing a platform simplifying the development of EOSC Services (or EOSC Service Components). This platform should be (a) open and extensible to easily add new facilities made available whenever a new EOSC Services start appearing and (b) cater for diverse exploitation models, e.g. to make it possible for user to use only the subset of facilities they need. These services should add value on top of the web-based APIs, e.g. it should offer “abstractions” on top of the

APIs.

EOsc Hosting Platform is a set of services enabling the deployment of EOsc Service onto EOsc nodes (cf. Sec. 3.4).

Underpinning Agreement Management is a set of services enabling the definition, storage, access, and monitoring of Underpinning Agreement governing the exploitation of a specific EOsc Service Component in the context of an EOsc Service.

EOsc Service Component Promotion is a set of services enabling Service Component owners (EOsc Service Component Supplier(s), Data (Service) Supplier(s)) to make their services known to the EOsc Users.

Multi-supplier Coordination is a set of services enabling to develop and manage a service delivery plan across more suppliers.

4.6. Services for EOsc System Managers

Figure 14 illustrates a set of services that are envisaged to support EOsc System Managers' activities described in Section 2.2.5³².

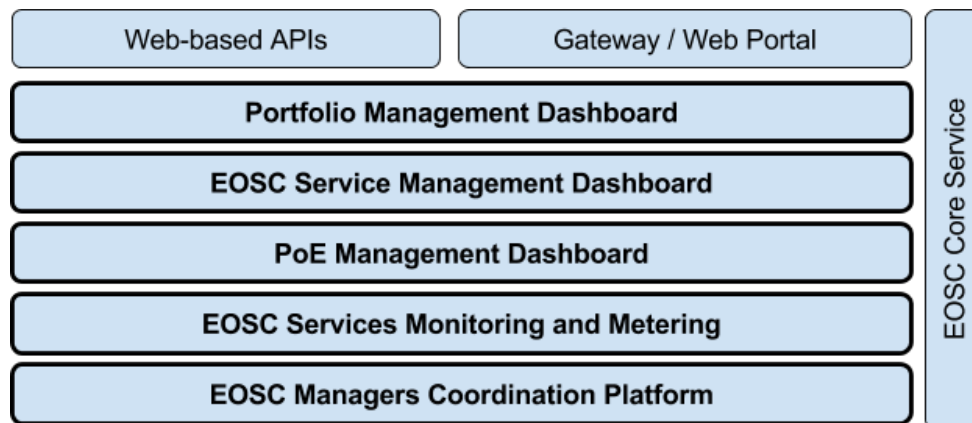


Figure 14. EOsc Services for EOsc System Managers

Portfolio Management Dashboard is a set of services enabling the various EOsc Managers to collaboratively develop and maintain the EOsc Portfolio. Such a Dashboard should provide the managers with a comprehensive picture of the “performance” of the currently offered services (e.g. use and users, incidents) as well as of the candidate services (e.g. audience, costs).

EOsc Service Management Dashboard is a set of services EOsc Service Providers are provided with to manage their service, e.g. be informed on incidents and requests for changes.

PoE Management Dashboard is a set of services enabling to define, develop and monitor the Principles of Engagement / Rules of Participation. Such services include a repository of the established rules (possibly described for human-consumption and machine-consumption), a compliance checker to assess the extent an entity (e.g. EOsc Service) is matching the established rules, an automatic alert system informing the managers whenever an entity is no longer adhering to the established rules.

EOsc Services Monitoring and Metering is a set of services enabling EOsc Managers to be constantly acquainted on the “performances” of the currently offered EOsc Services.

EOsc Managers Coordination Platform is a set of services enabling a smooth and timely communication and collaboration among the EOsc Managers. Such services include a repository of the documents and any other material worth sharing, various communication channels for sending messages to selected members and/or groups, ticketing systems enabling to assign tasks and monitor their execution.

³² These services are likely to be revised once both the EOsc governance model and the service management frameworks will be further defined and agreed.

5. CONCLUSIONS AND NEXT STEPS

This deliverable has reported the result of the work performed in Task 5.1 “EOSC Overall Architecture”. The outcome obtained so far is the result of a stepwise and iterative process of understanding and revision steps performed in the first 12 months of the project with the contribution of WP5 participants and many others from different work packages.

Quite early in performing this task it became clear that there were a number of aspects to be addressed before laying down the foundational aspects of the architecture. Among them, it was needed to clearly understand and make assumptions on a variety of complementary concepts and processes, like that of EOSC service, EOSC added-value services, EOSC catalogue, etc. Some of these concepts were also discussed in other Tasks/WPs. This required the set-up of collaboration and coordination activities with these Task/WPs that are planned to be continued and improved in the second period of the project. The Glossary in Appendix A was actually initiated to respond to the need of using an as much as possible consistent terminology across WP5 tasks. For the time being this terminology has been discussed exclusively among the participants of this work package. The plan for the second period is to transform this Glossary into a project resource by collecting feedback and terminology items also from different WPs.

We decided to derive the initial classes of EOSC services by starting with the identification of the roles of the expected users of the EOSC system. For each of these roles we have then identified its functionality needs and, consequently, the required services. Different approaches to the identification of EOSC services could have been taken. An alternative solution might have been to start directly with the proposal of a list of services based on the experience of the Task 5.1 members avoiding the lengthy investigation and characterization of the user roles. We have preferred the former analytic approach since it provides evidence on why certain services are needed and it facilitates more systematic iterative extensions and refinements.

The architecture produced so far should be interpreted as a description of classes of services that the EOSC system has to provide to support the identified activities. It is not yet a full reference architecture nor it is meant to be a concrete architecture for the EOSC. The architecture envisaged in this deliverable is intended to provide a grounding for further discussions and should enable these discussions to be based on a more solid basis.

The identification and specification of a concrete architecture will require further work to develop a better understanding of what services that are already provided by the existing systems (e.g. Research Infrastructures, service providers) can be brought into EOSC to contribute to its development, and to arrive at a consensus on the multiple aspects involved in its definition, like the core services to enable existing systems to interoperate, the federated delivery models chosen, the appropriate APIs, etc.

As outlined earlier in the document, the current list of service classes reflects the state-of-the-art ability to support Open Science. The current list is still far from being able to fully implement the Open Science view. The plan for the next months is to progressively enrich this first architecture version by taking into account progress in the shaping of Open Science by multiple existing and emerging initiatives. In particular, in the next couple of months we intend to collect feedback from other relevant initiatives external to the EOSCpilot (e.g. eRosa, INSTRUCT, SKA).

The services dedicated to Federated Service Management will also be revised to align them to the content of Deliverable 5.3 “EOSC Federated Service Management Framework” (M18). A similar approach will be taken when the content of Deliverable 2.5 “EOSC Rules of Engagement” will be finalised (M19).

In the second half of the second period, we will also start mapping existing services listed in some of the relevant EU and National catalogues, e.g. in eInfraCentral, with the presented service classes. This will be a first exercise to understand what might be the potential implementation of the identified architecture under the assumption that the listed services could all become EOSC services (i.e. be compliant with the EOSC Principles of Engagement).

Finally, we also plan to start the identification of a clear methodology (including processes) for regulating the evolution of the architecture. The EOsc architecture will be an evolving one, due not only to the progresses towards Open Science, but also to the expected changes of its component systems. The introduction of a methodology compliant with the governance mechanisms has been suggested during the discussion had so far with project stakeholders as a necessary instrument to preserve the consistency of the architectural choices. The definition of this methodology will be addressed towards the end of the second year when the governance structures will be more mature.

REFERENCES

- Andreozzi, S.; Berg, A.; Bot, J.; Collier, I.; Cordewener, B.; Holsinger, S.; Ferreira, N.; Koumantaros, K.; Meyer, K.; Vianello, D. (2018) ***EOSC Service portfolio***. EOSCpilot Deliverable D5.2 <https://eoscipilot.eu/content/d52-eosc-service-portfolio>
- Andreozzi, S.; Hormia-Poutanen, K.; Wood, J.; Ayris, P.; Cotter, S.; Luyben, K.; Manola, N.; Méndez, E.; Rossel, C.; Vignoli, M. (2017) ***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
- Appleton, O. (2012) ***Business models for Federated e-Infrastructures***. FedSM Deliverable D3.1, http://fitsm.itemo.org/sites/default/files/FedSM-D3.1-Business_models-v1.0.pdf
- Ayris, P.; Berthou, J-Y; Bruce, R.; Lindstaedt, S.; Monreale, A.; Mons, B.; Murayama, Y.; Södergård, C.; Tochtermann, K.; Wilkinson, R. (2016) ***Realising the European Open Science Cloud. First report and recommendations of the Commission High Level Expert Group on the European Open Science Cloud***. https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf
- Boardman, J. and Sauser, B. (2006) ***System of Systems - the meaning of "of"***, 2006 IEEE/SMC International Conference on System of Systems Engineering, Los Angeles, CA, 2006, pp. 6 doi: [10.1109/SYSOSE.2006.1652284](https://doi.org/10.1109/SYSOSE.2006.1652284)
- Duan, Y.; Fu, G.; Zhou, N.; Sun, X. Narendra, N. C. and Hu, B. (2015) ***Everything as a Service (XaaS) on the Cloud: Origins, Current and Future Trends*** IEEE 8th International Conference on Cloud Computing, New York City, NY, 2015, pp. 621-628. doi: [10.1109/CLOUD.2015.88](https://doi.org/10.1109/CLOUD.2015.88)
- Hienola, A.; Matthews, B.; Robertson, D.; Castelli, D.; Bicarregui, J.; Parland-von Essen, J.; Laaksonen, L.; Bijsterbosch, M.; Dovey, M.; Scott, M.; Pineda, O.; Oster, P.; Kontro, S.; Sorvari, S.; Girona, S.; Andreozzi, S.; Sartzetakis, S.; Bassler, U.; Beckmann, V.; Legre, Y. (2017) ***Draft Governance Framework For the European Open Science Cloud***. EOSCpilot Deliverable D2.2 <https://eoscipilot.eu/content/d22-draft-governance-framework-european-open-science-cloud>
- European Commission, Directorate-General for Research and Innovation (2016) ***Open innovation, open science, open to the world - A vision for Europe***. European Commission Doi: [10.2777/061652](https://doi.org/10.2777/061652)
- European Commission (2017) ***EOSC Declaration***. October 2017. https://ec.europa.eu/research/openscience/pdf/eosc_declaration.pdf
- Maier, M. W. (1996) ***Architecting Principles for Systems-of-Systems***. INCOSE International Symposium, 6: 565–573. [10.1002/j.2334-5837.1996.tb02054.x](https://doi.org/10.1002/j.2334-5837.1996.tb02054.x)
- Nielsen, M. (2012) ***Reinventing Discovery: The New Era of Networked Science***. Princeton University Press
- Ross-Hellauer T. ***What is open peer review? A systematic review*** [version 1; referees: 1 approved, 3 approved with reservations]. F1000Research 2017, 6:588 (doi: [10.12688/f1000research.11369.1](https://doi.org/10.12688/f1000research.11369.1))
- Terrovitis, M.; Tsiavos, P.; Oster, P.; Sandberg, M.; Gheller, C.; Pineda, O.; Seger, P. (2017) ***Draft Stakeholder Map***. EOSCpilot Deliverable D2.1 <https://eoscipilot.eu/content/d21-draft-stakeholder-map>
- Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak A. et al. (2016) ***The FAIR guiding principles for scientific data management and stewardship***. Scientific Data, 3:160018. [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

ANNEX A. GLOSSARY

The terms reported below and their definitions are part of a larger glossary that is expected to evolve during the project lifetime. The continuously updated version of the Glossary is available at <https://docs.google.com/document/d/1Gq73IU2tjVnXLIJpXLH9b4GBO6EGijpldUtGYHbMm0/edit>

Architecture: The fundamental organization of a software system embodied in its components, their relationships to each other, and to the environment, and the principles guiding its design and evolution.

Data (Service) Supplier: Any *EOSC Supplier* making available its own data (by means of a *Service*) to enable an *EOSC Service Provider* to operate an *EOSC Service*.

End-user: Every *EOSC User* being a *Researcher*, *Research Administrator*, or a *Third-party Service Provider*. Actors playing this role are the primary consumer of *EOSC Services*.

EOSC Compatible Service: Any service worth having in an *Service Catalogue* without being an *EOSC Service*. Compatibility implies only that such services can coexist with the *EOSC Services* within the catalogue context. Because of the inclusive willingness of the catalogue, no technical requirement is currently envisaged aiming at guaranteeing that these services can technically coexist / interoperate with *EOSC services* in application scenarios.

EOSC Core Service: An *EOSC Service* enacting the *EOSC System* and the implementation and delivery of the rest of *EOSC Services*. These services include *EOSC Service Catalogue* for service discovery, identity provisioning services for individuals and digital artifacts, federated authentication and authorization services, monitoring, accounting and billing, and service level agreement management.

EOSC Node: *EOSC Nodes* are the “organisational pieces” of the *EOSC System* called to contribute to the provisioning of one or more *EOSC Services*. Because of this, *EOSC Nodes* are: (a) the “places” where *EOSC Services* and/or *EOSC Service Component* reside; (b) operated by an *EOSC Service Provider* and/or and *EOSC Supplier*.

EOSC Portfolio: See *EOSC Service Portfolio*.

EOSC Principles of Engagement (PoE): Policies, processes, and roles governing the behaviour and the duties of *EOSC Users* when using *EOSC Services*.

EOSC Service: The ways to provide value to *EOSC End-user* through bringing about results that they want to achieve. *Services* usually provide value when taken on their own – unlike the specific service components of which they are composed. *EOSC services* are supplied by an *EOSC Service Provider* in accordance with the *EOSC Principles of Engagement for Service Providers*. [based on FitSM vocabulary] *EOSC Services* are approved by the *EOSC Service Portfolio Management Committee*.

EOSC Service Catalogue: The list of live / ready-to-use *EOSC Services* offered by the *EOSC System* with relevant information about these services. More Catalogue instances can exist, each tailored to serve the needs of a designated community. [FitSM vocabulary revised]

EOSC Service Component: It is a logical part of an *EOSC Service* that provides a function enabling or enhancing the implementation of the service. If components are services on its own and are expected to be exploited independently of the service they are contributing to, then the *Service Component* is also an *EOSC Service*. [FitSM vocabulary revised]

EOSC Service Component Supplier: Any organisation that provides an *EOSC Service Provider* with an *EOSC Service Component* that is need to provide the *EOSC Service*. [FitSM vocabulary revised]

EOSC Service Developer: Any *EOSC Supplier* making available its technical solution to enable an *EOSC Service Provider* to operate an *EOSC Service*.

EOSC Service Level Agreement (SLA) It is a documented agreement between an *EOSC User* and an *EOSC Service Provider* that specifies the *EOSC Service* to be provided and the service targets that define how it

will be provided. [FitSM vocabulary revised]

EOsc Service Portfolio It is a list of all the *EOsc Services* offered by *EOsc Service Providers* including those in preparation, live and discontinued. [FitSM vocabulary revised]

EOsc Service Provider It is an organisation that manages and delivers an *EOsc Service* to “customers”. Customers are the *EOsc System User*. [FitSM vocabulary revised]

EOsc Supplier Every *EOsc User* exploiting one or more *EOsc Services* to make available *EOsc Service Components* enacting the development and provisioning of *EOsc Services*. Suppliers include *EOsc Service Component Suppliers*, *Data (Service) Suppliers*, and *EOsc Service Developers*.

EOsc System The overall IT system realising *EOsc*.

EOsc System Manager Every *EOsc System User* being an *EOsc System Owner*, *EOsc System Top Manager*, or an *EOsc Service Provider*.

EOsc System Owner Every *EOsc System User* responsible / accountable for the establishment and maintenance of *EOsc System*. Actually this is a collective role played by a committee. By liaising with the rest of *EOsc System Managers* (namely *EOsc System Top Managers* and *EOsc Service Providers*) this role is called to steer the *EOsc System* by setting the key goals and directions. Among its tasks there is the definition of the *EOsc Service Portfolio*.

EOsc System Top Manager Every *EOsc System User* responsible / accountable for the overall operation of *EOsc System*. Actually this is a collective role played by a committee. By liaising with the rest of *EOsc System Managers* (namely *EOsc System Owners* and *EOsc Service Providers*) this role is called to put in place actions aiming at guaranteeing that the *EOsc System* is behaving according to the goal and directions established by the *EOsc Owners*.

EOsc System User Every actor (human or machine) exploiting an *EOsc Service*. An *EOsc System User* might be further specialised in roles including (a) *End-users*, i.e. a *Researchers*, *Research Administrator*, and *Third-party Service Provider*, (b) a *Supplier*, i.e. *EOsc Service Component Supplier*, *Data (Service) Supplier*, *EOsc Service Developer*, (c) an *EOsc System Manager*, i.e. *EOsc System Owner*, *EOsc System Top Manager*, *EOsc Service Provider*.

Reference architecture The abstract architectural elements in the domain independent of the technologies, protocols, and products that are used to implement the domain.

Research Administrator A role played by an *EOsc System User* willing to inform their organization on research activities of interest by relying on *EOsc Services*. Research activities of interest for an organisation include research projects funded by the organization and *Research Administrators* might be called to collect the results of these projects to to produce metrics and indicators allowing to assess the past (impact) and shape the future (trends).

Researcher A role played by an *EOsc User* willing to perform his/her research activity by relying on *EOsc Services*.

Third-party Service Provider A role played by an *EOsc User* willing to develop and operate a *Service* by relying on one or more *EOsc Services*.

Underpinning Agreement It is a documented agreement between an *EOsc Service Provider* and an *EOsc Service Supplier* that specifies the underpinning *EOsc service component(s)* the *EOsc Service Provider* is provided with by the *EOsc Service Supplier*, together with the related service targets [FitSM vocabulary revised]

ANNEX B. DETAILED DESCRIPTION OF ACTIVITIES TO BE SUPPORTED BY THE EOOSC SYSTEM

B.1. Activities performed by Researchers

The following table details the activities identified in [Sec. 2.2.1](#).

Table 1. Activities performed by Researchers

Name and Description	Responsible actor ³³
<p>Finding research artefacts. This is the class of actions researcher perform to discover research outcomes of potential interest.</p> <p>There are two major typologies of discovery actions: “<i>push-oriented</i>” and “<i>pull-oriented</i>”. The former typology envisages an active role of the user that by specifying a query or browsing through the available items tries to find those fitting its purpose. The latter typology envisages an active role of finding “<i>system</i>”/ “<i>service</i>” that proactively provides Researchers with items of potential interests (recommender systems).</p> <p>Discovery actions are possibly affected by <i>domain-specific practices</i>. It is of paramount importance to provide researchers with various discovery options promoting diverse trade-off between “<i>precision</i>” and “<i>recall</i>”, e.g. researchers interested in exploratory discovery tasks over a wide information space have diverse expectations than researchers interested in discovery artefacts in a specific repository by using a domain-specific terminology/ontology.</p> <p>Similar arguments occur when considering the typology of artefacts researchers are looking for.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • <i>Discovery by keywords</i>, e.g. the researcher specifies some words characterising the artefact he/she is looking for and get back the published artefacts matching the specified keywords according to some similarity metrics; • <i>Browsing</i>, e.g. the researcher explores the collection of published artefacts and moves across them by relying on explicit and/or derived links; • <i>Faceted-search</i>, e.g. the researcher explores the collection of published artefacts (that have been organised according to several faceted classification systems) by applying multiple filters. 	<p>Researcher (R), EOOSC Service Provider (I)</p>
<p>Accessing research artefacts. This is the class of actions researchers perform to obtain an artefact of potential interest.</p> <p>There are two typologies of access worth being considered: “<i>remote access</i>” and “<i>local / in-situ access</i>”. The former typology is certainly oriented to reach the</p>	<p>Researcher (R), EOOSC Service Provider (I)</p>

³³ The RACI Matrix is expected to be used to explain responsibilities. The following values are used:

- R (Responsible) those who do the work to achieve the task.
- A (Accountable) the one ultimately answerable for the correct and thorough completion of the deliverable or task, and the one who delegates the work to those responsible.
- C (Consulted) Those whose opinions are sought, typically subject matter experts; and with whom there is two-way communication.
- I (Informed) Those who are kept up-to-date on progress, often only on completion of the task or deliverable; and with whom there is just one-way communication.

<p>widest audience since it envisages a web-based/network protocol-based way to get the artefact of interest. The latter requires that the requester moves where the artefact is based. For the case of local / in situ access the best the EOSC system can do is to provide researchers with information on how access is going to take place.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • <i>To move the artefact of interest</i> (or a subset of it) from the current location to another one considered more suitable for the researcher tasks. The primary artefact to consider here are datasets. 	
<p>(Re-)using research artefacts/outcomes. This is the class of actions Researchers perform in order to create new research outcomes by exploiting existing ones.</p> <p>The scope is very wide because of the variety of the typologies of artefacts to be (re-)used, the diversity of (re-)use patterns, the variance of (re-)use goals, etc. With the emergence of the Open Science approach these activities will progressively be performed more and more collaboratively.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • <i>Collating</i> a number of new or re-used datasets to produce a new dataset; • <i>Processing</i> or <i>mining</i> input data, according domain specific algorithms, in order to generate new products; • Combining a set of existing services into a user-defined workflow; • <i>Repeating, replicating, repurposing</i> a research activity; 	<p>Researcher (R), EOSC Service Provider (I)</p>
<p>Depositing research artefacts. This is the class of actions researchers perform in order to store and preserve the artefacts they own.</p> <p>These actions are key in order to enact others to access a research outcome (e.g. publication, data, tool, workflow, research object) also in the future. They are usually performed by relying on repository services taking care of storing the research outcome and guaranteeing that it remains available in the short and long period. Researchers may expect different functionalities from these repositories depending on the nature of the research outcome to be stored and on the preservation practices to be implemented.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • <i>Depositing a research outcome</i> and associating with it appropriate metadata sets (e.g. for retrieval, preservation, policy management); • <i>Depositing successive versions</i> of the research outcome; 	<p>Researcher (R), Research Admin (I), EOSC Service Provider (I)</p>

<p>Publishing research artefacts. This is the class of actions researchers perform in order to communicate and make available their research outcomes to peers and to other stakeholders that may be interested in exploiting them.</p> <p>This class of actions are part of the activities dedicated to scholarly communication. In the new Open Science vision that demands for repeatability, reproducibility, and repurposability, this form of communication must go well beyond the paper and publisher-centered one that has been used by researchers since now. It must evolve into a “comprehensive and social” communication offering to the interested consumers the possibility to access and consume the entire set of capabilities, processes and procedures that have been exploited in producing the research outcome. In order to make this possible all the contextual elements must be linked to the research outcome. Researches must be enabled to do it even if with the progress of technologies and, in particular of information mining, these links can be collected and derived automatically.</p> <p>In order to support Open Science publishing should also assure the necessary level of preservation of the research outcomes, i.e. the possibility of accessing and using it along the time.</p> <p>The publishing of research outcomes may also include peer-review processes to ensure the quality what is submitted to publication. This processes can be under the control of a publisher and a set of nominated reviewers or it can take new forms like the emerging open-peer review [Ross-Hellauer, 2016].</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • <i>Linking published research outcomes</i> (e.g. publications with related datasets and tools). 	<p>Researcher (R), Research Admin (I), EOSC Service Provider (I)</p>
<p>Collaborate to address a common challenge. This is the class of actions researchers perform in order to engage others to jointly work on a research endeavour.</p> <p>These actions include the rich and multifaceted set of activities researchers perform in web accessible “virtual laboratories / collaboratories” where collaboration is taking place across the borders of organisations and countries.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • <i>Create and operate a “virtual research environment”</i> to share research artefacts with co-workers willing to contribute to a research activity; • <i>Openly share a research idea / issue / text / algorithm, etc.</i> to collect contribution from the community in the large / the crowd; • <i>Annotate research artefacts.</i> 	<p>Researcher (R)</p>

<p>Open and track requests for enhancements. This is the class of actions researchers perform in order to provide EOSC Managers with requests for changes wrt the currently offered services.</p> <p>Requests for changes can be of various genre ranging from the report of a malfunction to the request for a completely new service, the capability- / capacity-oriented reinforcement of an existing service. The primary stakeholders responsible for replying to such requests are the EOSC Managers.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Inform service provider that a service quality should be enhanced; • Inform EOSC managers that an existing service should be developed to support new features; 	<p>Researcher (R), EOSC System Owner (I), EOSC System Top Manager (I)</p>
---	---

B.2. Activities performed by Research Admins

The following table details the activities identified in [Sec. 2.2.2](#).

Table 2. Activities performed by Research Admins

Name and Description	Responsible actor
<p>Monitoring research outcomes. This is the class of actions research administrators perform in order to be informed on the list of artefacts resulting from a research activity they promote and the impact these artefacts have produced (e.g. citation metrics, acknowledgments).</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Identify and select the appropriate EOSC Services for the organization (i.e. Services that are suitable for their organization). This includes repositories and CRIS, where liaison with the organization’s ITs and library (if any) could be beneficial. Collect feedback from organization researchers. • Put the processes and policies in place to ensure <u>all</u> research outcomes will be identified. This could be liaised with the organization’s library and legal office. Interact with EOSC System Owner and/or EOSC System Top Manager for issues related to machine readable policies and monitoring policies compliance. • Identify all research outputs linked to a specific funding scheme, organization, person/researcher. • Contextualize the research outcomes: retrieve all relevant information (metadata), ensuring provenance is preserved. 	<p>Research Admin (R)</p>

<p>Assessing research impact. This is the class of actions on research analytics that research admins perform in order to assess the impact of these artefacts.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Formulate quantitative and qualitative metrics for short and long term research impact. • Analyze results and discover trends and correlations relevant to: excellence of science, policies, use of e-Infrastructure, impact to society and economy. • Produce reports for organization strategic boards. Visualize results and share in private/public fora. • Define new indicators and metrics as these come from the new data analysis. 	Research Admin
<p>Operating Communication hubs. This is the class of actions Research Admins perform in order to establish communication channels with Third-Party affiliates, as well as to encourage/promote the development of community-driven networks both for researchers and the general public/citizens. To ensure efficiency in these activities, they interact and coordinate efforts with EOSC System Top Manager, EOSC Service Provider, and library liaisons.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Operate and maintain a/multiple communication hub(s) for supporting research communities, citizen scientists and citizens, funding agencies, RIs and e-infrastructures that they can use to: <ul style="list-style-type: none"> ○ organise, advertise and participate in congresses and events (internal, external); ○ exchange information regarding research activities and beyond; ○ support and strengthen scholarly communication drivers such as (open) peer review; ○ have better control on particular administrative procedures and identify gaps and what needs to be strengthened; ○ provide guidance through specialised services catalogues that accommodate research lifecycle, sensitive data, discipline-specific needs and enhance research discovery and performance; ○ provide technical support for training activities through collaborative workshops, webinars and through relevant online tools and services ○ provide a collection of various means of communication to facilitate research collaborations; ○ support interdisciplinary and international communication channels. 	Research Admin (R)

<p>Open and track requests for enhancements. This is the class of actions Research Admins perform in order to provide EOSC with requests for changes and/or additions with respect to the currently offered services. Research Admins interact with EOSC Managers, i.e. the primary actors to reply to these requests, to ensure that the quality of the EOSC services dedicated to them is maintained.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • provide feedback on the current services offered by (a) suggesting improvements that follow Open Science research trends, including policymaking; (b) identifying areas where a potential new service/tool could be beneficial; (c) Informing/ submitting tickets about system and/or service errors; 	<p>Research Admin (R), EOSC System Top Manager (C), EOSC Service Provider (A)</p>
--	---

B.3. Activities performed by Third-Party Service Providers

The following table details the activities identified in [Sec. 2.2.3](#).

Table 3. Activities performed by Third-party Service Providers

Name and Description	Responsible actor
<p>Finding EOSC Services. This is the class of actions Third-Party Service providers perform to get acquainted with the EOSC System services that they might usefully exploit in developing their own.</p> <p>Two main models of discovery of these service offering can be envisaged: autonomous and brokered. In the autonomous model the provider, by interfacing with the EOSC Services Catalogue, tries to identify the set of capabilities and resources needed to implement the service he/she is called to / willing to realize. In the brokered model, the provider interfaces with an EOSC specific service by providing it with a “wish list” and gets back one or more proposal for suitable services to use.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Find services with a <i>terms of use</i> compatible with the one planned for the new third-party service to be developed; • Find services operated by a <i>known provider</i>; • Find a service offering a specific functionality among those <i>published by a specific community</i>; 	<p>Third-Party Service Provider (R), EOSC Service Provider (C)</p>
<p>Accessing EOSC Services. This is the class of actions Third-Party Service providers perform to get access to the EOSC Services they are interested in.</p> <p>The access considered here is the programmatic one, i.e. the service provider develop her planned service what is needed for the service s/he is planning to offer to establish a recognised connection with one or more EOSC Services.</p>	<p>Third-Party Service Provider (R), EOSC Service Provider (C)</p>

<p>Using EOSC Services. This is the class of actions Third-Party Service Providers perform to actually exploit the facilities offered by one or more EOSC Services.</p> <p>The usage patterns are the programmatic ones and focuses on the APIs and protocols the target EOSC Service offer to make it possible for the service provider to build upon the offered facilities to implement the envisaged / planned service.</p>	<p>Third-Party Service Provider (R), EOSC Service Provider (C / I)</p>
<p>Managing the SLA. This is the class of actions Third-Party Service Providers perform to establish the Service Level Agreement regulating the provisioning of the target EOSC Service(s). Depending on the target EOSC Service it might possible or not to “tune” / “negotiate” the agreement, i.e. there are EOSC Service(s) whose provisioning is established a-priori / independently of the user. In addition to the “definition” of the agreement, this class of actions includes any agreement-related activity, e.g. the notification of any infringement of the agreement by any of the parts.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Reviewed SLAs at planned intervals; • Evaluate service performance against service targets defined in SLAs; 	<p>Third-Party Service Provider (R / A), EOSC Service Provider (R / A)</p>
<p>Providing feedback on EOSC Services. This is the class of actions Third-Party Service Providers perform to provide any EOSC Service Provider with feedback resulting from the concrete exploitation of the service in their own settings. These activities should be performed in a systematic way thus to guarantee a constant liaison between the consumption side and the provisioning side. In particular, EOSC managers are provided with valuable indicators helping to define the EOSC portfolio / offering (e.g. discontinue a “bad” service, reconsidering the SLA characterising a service) and to maintain the provisioning process to guarantee the planned quality level.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Participate regular service reviews with the service providers; • Inform service providers about unsatisfactory service level issues; 	<p>Third-Party Service Provider (R), EOSC Service Provider (I), EOSC System Owner (I), EOSC System Top Manager (I)</p>

<p>Open and track requests for enhancements. This is the class of actions Third-Party Service Providers perform in order to provide EOSC Managers with feedback and requests for changes wrt the currently offered services. These actions are not systematic, rather they occur whenever a change in the provided EOSC Service is expected, e.g. enhance the service “capacity”, reconsider the API or the supported standard(s). Actions are triggered by the Third-Party Service Provider and are primarily managed by the EOSC Service Provider that is in charge to support the planned enhancement. The EOSC Service Provider is responding to the requests by liaising with both the EOSC System Top Manager and EOSC System Owner.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Inform EOSC managers that an existing service should be developed to support new features; • Inform EOSC managers that need for service capacity EOSC service will change. • Create a request to modify EOSC architecture (or to accept an exception to EOSC architecture), namely to ask architecture managers that an existing service should be developed to support new features or to propose completely new services; 	<p>Third-Party Service Provider (R), EOSC Service Provider (A), EOSC System Top Manager (C), EOSC System Owner (I)</p>
---	--

B.4. Activities performed by EOSC Suppliers

The following table details the activities identified in [Sec. 2.2.4](#).

Table 4. Activities performed by EOSC Suppliers

Name and Description	Responsible actor
<p>Managing the EOSC Service Component Provisioning. This is the class of actions EOSC-Service-Component Suppliers perform to provide EOSC Service Providers with the target EOSC Service Component. It includes the activities aiming at guaranteeing the availability and performance of the specific component according to the agreement established by the underpinning agreement.</p>	<p>EOSC Service Component Supplier (R), EOSC Service Provider (I)</p>
<p>Publishing a future/possible EOSC Service Component. This is the class of actions potential forthcoming EOSC Service Component Suppliers perform in order to advertise the service they own to the EOSC Managers. Suppliers prepare a careful description of the service both from a functional and non-functional perspective thus to enable the management to take an informed decision on service exploitation. In particular, non-functional related information particularly relevant include potential clients / market, capacity, and use cases.</p>	<p>EOSC Service Component Supplier (R), EOSC Service Provider (I), EOSC System Top Manager (I), EOSC System Owner (I)</p>
<p>Managing the Data (Service) Provisioning. This is the class of actions Data (Service) Suppliers perform to provide EOSC Service Providers with their data. These actions include the adaptation of data provisioning approaches to the exploitation settings (e.g. changes in formats and protocols to be supported), continuous update of the characterisation of the service capacity (mainly the data coverage, extent, size, and other quality related features), immediate alert of any (planned) service interruption, repair service malfunctions.</p>	<p>EOSC Data (Service) Supplier (R)</p>

<p>Managing the Underpinning Agreement. This is the class of actions EOSC Service Component Suppliers perform to establish the agreement regulating the provisioning of their own services in the context of the provisioning of an EOSC Service. Depending on the specificities of the provided component, there are cases where the underpinning agreement is the result of a negotiation between the supplier and the EOSC Service Provider and other cases where the agreement is defined unilaterally from the supplier or the EOSC Service Provider.</p>	EOSC Service Component Supplier (R)
<p>Developing an EOSC Service. This is the class of actions EOSC Service Developers perform to develop the <i>technical solution</i> underlying an EOSC Service. These actions requires the EOSC platform, i.e. the environment in which the software is deployed and executed to realise the EOSC Service.</p> <p>Examples of typical actions in this class are:</p> <ul style="list-style-type: none"> • Design/revise the technical solution to support customer needs; • Create a <i>service transition package</i> necessary to take the service into use. 	EOSC Service Developer (R), EOSC Service Provider (A)

B.5. Activities performed by EOSC System Managers

The activities ***EOSC System Managers*** are expected to perform via EOSC depends on their responsibility in managing the EOSC System, i.e. whether they are playing the role of ***EOSC System Owner, EOSC System Top Manager, or EOSC Service Provider.***

In particular, the following activities and responsibilities are envisaged.

Table 5. Activities performed by EOSC System Managers

Name and Description	Responsible actor
<p>Governing the EOSC system. This is the class of actions EOSC System Owner(s) perform to develop, control and regulate the EOSC system as a whole. These activities are performed in accordance with the policies and decisions taken from the entire EOSC Governance.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Creation, storage and discovery of machine readable “EOSC policies”; • Monitoring uptake and compliance of EOSC Services with EOSC policies; 	EOSC System Owner (R), EOSC System Top Manager (C)
<p>Operating the EOSC system. This is the class of actions EOSC System Top Manager(s) perform to make the EOSC system working as planned.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Assign one individual accountable for the overall EOSC system with sufficient authority to exercise his/her roles. • Define goals and policies. • Conduct management reviews at planned intervals. 	EOSC System Top Manager (R), EOSC System Owner (A)

<p>Managing the EOSC Portfolio. This is the class of actions EOSC Managers perform in order to define and maintain the EOSC Portfolio.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Manage and Maintain the EOSC service portfolio: <ul style="list-style-type: none"> ○ Add a service to the EOSC service portfolio. ○ Update a service to the EOSC service portfolio. ○ Retire a service to the EOSC service portfolio. • Manage the design and the transition of new/changed service: <ul style="list-style-type: none"> ○ Create and Approve a Service Design and Transition Package (SDTP). ○ Update a Service Design Transition Package. • Manage the organisational structure involved in delivering the EOSC services. <p>The SDTP describes the core documentation of the service, from when it is first suggested as a possibility to when it is finally retired.</p> <p>The EOSC service portfolio is the basis for the EOSC service Catalogues.</p>	<p>EOSC System Owner (A), EOSC System Top manager (R), EOSC Service Provider (C)</p>
<p>Managing the EOSC Service. This is the class of actions each EOSC Service Provider is called to perform in order to make a specific EOSC Service available for use and behaving like expected. Independently of the service specificities, the management of a service include the continuous monitoring, the maintenance of the enabling technology, the constant curation of the data made available by the service, etc.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Curate the content the service give access to; • Manage users (e.g. grant access, assign role); 	<p>EOSC Service Provider (R), EOSC System Top manager (I)</p>
<p>Managing Catalogue(s). This is the class of actions EOSC Managers are called to perform in order to define and maintain a Catalogue Service. Several typologies of catalogues can be envisaged yet any key typology is the EOSC Services Catalogue.</p> <p>Examples of typical actions in this class include:</p> <ul style="list-style-type: none"> • Maintain the EOSC Catalogue(s): <ul style="list-style-type: none"> ○ Add a new service to EOSC Catalogue(s) ○ Update a service to EOSC Catalogue(s) ○ Remove a service to EOSC Catalogue(s) 	<p>EOSC Service Provider (R),</p>

<p><i>Managing the requests for changes.</i> This is the class of actions EOsc Managers are called to perform to respond to the solicitations received from the EOsc End-users making use of the EOsc system and its Services as well as from other actors (e.g. EOsc Governance).</p> <p>Examples of typical actions in this class are:</p> <ul style="list-style-type: none">• Manage changes (including emergency changes)<ul style="list-style-type: none">○ Record a Request For Change (RFC).○ Classify a Request For Change (RFC).○ Evaluate a Request For Change (RFC).○ Approve a change.○ Implement a change.○ Perform a post implementation review.• Maintain the list, descriptions and step-by-step workflows and well-known and recurring changes (standard changes).• Maintain the schedule of change.	EOsc Managers (R)
---	-------------------

ANNEX C. EOSC STAKEHOLDERS AND THEIR ROLES IN THE EOSC SYSTEM

EOSCpilot D2.1 [Terrovitis et al. 2018] identifies a number of stakeholders that are expected to contribute to EOSC, both in terms of governance and provision. Moreover, D2.2 [Hienola et al. 2017] decided to concentrate on three stakeholder roles: **Providers**, **Consumer**, **Decision-makers** - “understanding that different stakeholders can play multiple roles, or different roles are different points in the research lifecycle or within their organisation”.

This Appendix aims at clarifying what are the relationships among actor roles envisaged for the needs of defining the EOSC System and the stakeholder primary roles discussed in WP2.

Table 6. Associating EOSC Stakeholders Primary Roles with EOSC System Actor Roles

Stakeholder Primary Role in D2.2	Relationship with actor roles in Sec. 2.2
<p>Provider: Provides services, data or other resources (e.g. scientific instruments, training) into EOSC. The typical stakeholders playing this role are: e-Infrastructures, Service Providers, Enterprise, Academic Institutions and Research Libraries, Research Infrastructures, Outputs from VRE, and Other H2020 Projects.</p>	<p>The stakeholder Provider role primarily matches the Supplier and the EOSC Service Provider envisaged in this deliverable. In fact, actors playing these roles provides resources to EOSC. In particular, among the Suppliers there are (a) the EOSC Service Component(s) Suppliers that provide services and components into EOSC that are exploited to develop and operate an EOSC Service; (b) the Data (Service) Supplier that provides data into EOSC; (c) the EOSC-Service Developer that provides technical solutions into EOSC. The EOSC Service Providers provide EOSC Services into EOSC.</p> <p>Moreover, End-users may play the role of Provider too. In fact, when exploiting EOSC Services they might provide resources into EOSC, e.g. when a Researcher deposits and publishes a dataset into an EOSC Repository he / she is actually providing the dataset into EOSC.</p>
<p>Consumer: Will make use of services, data, or other resources from EOSC. The typical stakeholders playing this role are: Learned Societies, Research Communities, Scientific and Professional Associations, Research Infrastructures, Research Producing Organisation, e-Infrastructures, VRE, and Other H2020 Projects, Academic Institutions and Research Libraries, Enterprise, and the General Public.</p>	<p>The stakeholder Consumer role matches all the EOSC System Users envisaged in this deliverable. In fact, every System User is expected to rely on the envisaged EOSC Services (cf. Sec. 4) to perform his/her tasks. Typical consumers are the End-User when being a Researcher, a Researcher Admin, or a Third-party Service Provider. In addition to them, System Managers and Suppliers are expected to make use of EOSC Services too.</p>
<p>Decision-maker: Will be involved in the strategic direction, compliance and funding of EOSC. The typical stakeholders playing this role are: National, Regional or Local Government Agencies, Research Funding Bodies.</p>	<p>The stakeholder Decision-maker role does not have a primary match with the envisaged EOSC System roles. Rather it is primarily connected with the EOSC Governance boards envisaged in D2.2. The decision and strategic directions approved by the EOSC Governance boards will be put in place by the EOSC System Owner, the EOSC System Top Manager as well as by the EOSC Service Providers.</p>

In addition to the stakeholders' primary roles discussed above, three additional roles have been envisaged in D2.2 [Hienola et al. 2017]: **Intermediary, Funder, Policy-maker**.

Table 7. Associating EOSC Stakeholders Supplementary Roles with EOSC System Actor Roles

<i>Stakeholder Supplementary Role in D2.2</i>	<i>Relationship with actor roles in Sec. 2.2</i>
Intermediary: Many stakeholders (including e-infrastructures, research infrastructures, VREs etc.) will consume services from some providers to provide value added services to other consumers. This is a sub-role for both Stakeholder Provider and Consumer.	The stakeholder Intermediary matches the role of Third-party Service Provider as well as EOSC Service Provider . In fact, a Third-party Service Provider make use of EOSC Services to provide value added services to a designated community. An EOSC Service Provider make use of EOSC Service Components provided by EOSC Suppliers to deliver added value services (EOSC Services).
Funder: Provides funding for research on a local, national or international level. This is a sub-role of Stakeholder Decision-maker.	The stakeholder Funder role does not have a primary match with the envisaged EOSC System roles. They certainly contribute to the development of the EOSC system by supporting EOSC System Managers, Suppliers, as well as EOSC End-users.
Policy-maker: Regulates policy at a local, national or regional level. This is a sub-role of Stakeholder Decision-maker.	The stakeholder Policy-maker does not have a primary match with the envisaged EOSC System roles. They may match the Research Admin role when interested in exploiting EOSC Services to assess the impact their policies have in EOSC.

EOSCpilot D2.1 [Terrovitis et al. 2018] identifies a number of stakeholders that are expected to contribute to EOSC, both in terms of governance and provision.

This Appendix aims at clarifying what are the actor roles operating in the framework of these stakeholder initiatives.

Table 8. Associating EOSC Stakeholders with EOSC System Actor Roles

<i>Stakeholder³⁴</i>	<i>Roles in EOSC System</i>
European e-Infrastructures: Existing key European e-Infrastructures (EUDAT, EGI, OpenAIRE, GEANT and PRACE) are the natural starting point for engaging EOSC stakeholders. They represent activities that are closely linked to e-Infrastructures, including cloud operations, and they are already organized for further integration.	<p>European e-Infrastructures members / players can play almost all roles in the context of EOSC.</p> <p>E-Infrastructure members can play the role of EOSC End-users. They can be <i>Third-Party service providers</i> willing to develop new facilities for their e-Infrastructures by relying on EOSC offering.</p> <p>E-Infrastructure members can also play the role of EOSC Suppliers. They can contribute their artefacts (be these artefacts services and/or data) to EOSC thus to make it possible for EOSC to offer added value facilities on top of RI artefacts. Such added</p>

³⁴ The content of this part comes from D2.1.

	<p>value facilities can be diverse, they can range from facilities conceptually homologous to that initially provided by the e-Infrastructure yet adapted to serve the needs of a larger and heterogeneous community like that served by EOSC to completely new services resulting from the combination of diverse artefacts provisioned by several providers including e-Infrastructures.</p> <p>e-Infrastructure members can play the role of <i>EOSC Managers</i>. They can be called to play the role of <i>EOSC Service Providers</i> when willing to operate an EOSC Service on behalf of EOSC.</p>
<p>Data / Research Initiatives: There is a variety of organizations and initiatives, e.g., RDA interest and working groups, which constitute already organized research communities with specific stakes in cloud services.</p>	<p>Data / Research Initiative members can play the role of <i>EOSC End-users</i>.</p> <p>Data / Research Initiative members can play the role of <i>EOSC Suppliers</i>.</p>
<p>Cloud providers: Cloud services providers, public and private, are key stakeholders by definition in EOSCPilot. Engaging major companies, that provide services to a wide range of research activities is important for bringing together the needs of the research communities and the offered services, and to address some Data privacy issues.</p>	<p>Cloud providers can play the role of <i>EOSC Suppliers</i>.</p>
<p>Research funders: Funding bodies both on national and EU level are major stakeholder in EOSC, since they support research in all its stages. Despite their different organizational schemes in different countries, EOSC needs to actively engage them to support the future direction of EU cloud infrastructures.</p>	<p>Research funders can play the role of <i>EOSC End-users</i>. Funding excellent research (e.g. ERC), they need to ensure all facilities to support research activities e-infrastructures are in place to reduce grantees burden in producing and sharing research results.</p> <p>Research funders can play the role of <i>EOSC Managers</i>. They need to identify appropriate mechanisms to allocate funding for the use and optimal utilization of the relevant infrastructures, thus participating / advising in the governance for potential certification mechanisms.</p>
<p>Cloud community: The research community working with cloud technologies and participating in EU cloud projects is a valuable ally in designing the framework for cloud services in EU. Identifying key players in this sector is an important challenge.</p>	<p>Cloud community can play the role of <i>EOSC Supplier</i>.</p>

<p>Research Communities and Institutions: Researchers, Research communities and Research institutions as main consumers of cloud services are a basic pillar in EOSC, since they define the consumer side of the cloud ecosystem. This is a broad class of consumers and identifying the key contact points and the right organization scheme to engage them is a complex task. Some of them are already organized at European scale, such as the HPC Centre of Excellence (CoEs).</p>	<p>Research Communities and Institutions can play the role of <i>EOSC End-users</i>.</p>
<p>Research Infrastructures: The Research Infrastructure projects of EC complement the existing EU infrastructure landscape, by providing thematic (vertical) infrastructures in contrast with the horizontal e-infrastructure. RIs are the key starting point for engaging communities since they are already organized in an operational structure and use or provide cloud services.</p>	<p>Research Infrastructure members / players can play almost all roles in the context of EOSC.</p> <p>RI members can play the role of <i>EOSC End-user</i>. They can be <i>Researchers</i> willing to exploit EOSC facilities for their research purposes in parallel with the facilities offered by their own RI. They can be <i>Third-Party service providers</i> willing to develop new facilities for their RI by relying on EOSC offering.</p> <p>RI members can play the role of <i>EOSC Suppliers</i>. They can contribute their artefacts (be these artefacts services and / or data) to EOSC thus to make it possible for EOSC to offer added value facilities on top of RI artefacts. Such added value facilities can be diverse, they range from facilities conceptually homologous to that initially provided by the RI yet adapted to serve the needs of a larger and heterogeneous community like that served by EOSC to “mediators” aiming at providing seamless access to datasets spread across diverse RIs, completely new services resulting from the combination of diverse artefacts provisioned by several RIs.</p> <p>RI members can play the role of <i>EOSC Managers</i>. They can be called to play the role of <i>EOSC Service Providers</i> when willing to operate an EOSC Service on behalf of EOSC.</p>
<p>Policy makers: Policy makers affect cloud infrastructures in profound ways, even when they do not act as funders. For example, regulatory bodies on data privacy, on competition and, of course, on research can shape the future of the cloud ecosystem in the EU. EOSCpilot has to identify the most closely related ones and investigate the more productive way to engage them.</p>	<p>Policy makers can play several roles including <i>Research Admins</i>. However, the scope of the stakeholder identified by D2.1 goes well beyond actors to be served by the EOSC System.</p>

ANNEX D. TYPICAL SCIENTIFIC SCENARIOS

All the EOSCpilot demonstrators are working together to develop their domain-specific use cases with Service Suppliers participating in the project. In doing so they aim at validating whether EOSC services that exploit these supplied services can meet the case requirements.

This section aims at further illustrating the activities performed by the actors with the identified roles (and hence to clarify requirements) by exploiting two concrete examples extracted from the EOSCpilot Science Demonstrators workflows. For each workflow, the different EOSC actors involved and their relationships are identified.

The Service Pilots task (Task 5.4) aims to support the development of the Science Demonstrators by helping them to meet the goals stated in their work plans.

At the time of writing this deliverable, all Science Demonstrators have been developed by exploiting a wide range of services:

- Network connectivity services provided by GÉANT to lower the gap between different countries in Europe and its neighbouring countries by helping the development of research and education networking;
- IaaS, data federating services (e.g. Onedata), data distribution services (e.g. CVMFS) and data preservation services (e.g. EUDAT B2SAFE);
- High level services such as D4Science Virtual Research Environment-as-a-Service or the repository service Invenio.

To a greater or lesser extent, all Science Demonstrators (SDs) (that have been active until the writing of this deliverable) make use of different combinations of services from the above categories with the aim of providing a user facing service in the form of a portal or a frontend to allow their respective science communities to solve their problems or give them specific added value. The end users will need no knowledge whatsoever of the details of EOSC services involved, and nor should they.

In order to understand which roles are involved in the final delivery of services resulting from the Science Demonstrators, we consider the successful evolution of the prototypal demonstrators to fully-fledged production-level services. Once this has been achieved, additional actor roles may also be called upon. Their needs will then be analysed in the next period to evaluate whether they require the introduction of new service classes in the architecture model.

The first example is the *EPOS/VERCE Science Demonstrator*, which aims to seamlessly access to EOSC resources by enabling the computational seismologists to carry out data analysis and produce, as a final result, data products that are available under the FAIR principles. The implementation of this Science Demonstrator involves actors playing the following roles:

- The **Researcher** is responsible to conduct research activities, following the FAIR principles directives, by exploiting services supplied by existing computing and data infrastructures represented in the EOSCpilot project. In the context of this specific SD, the Researcher will use the service published in the Service Catalogue to generate synthetic seismograms for public Earth models and earthquake's source simulations (actions: 3, 4, and 5).
- The **Data Scientist** is responsible to upload real observations and Earth models in a distributed third-parties archives, and to compare this data with the simulated data (synthetic seismograms) generated by the Researcher in order to evaluate and refine the Earth models (actions: 6, 7, and 8). For this specific SD, the Researchers and Data Scientists roles are played by the same actor.
- **Service Providers** responsible to provide the needed services (e.g. cloud VMs, analysis tools) for supporting the implementation of the SD pilot.

The diagram shown in the figure below describes how these actors interact each other:

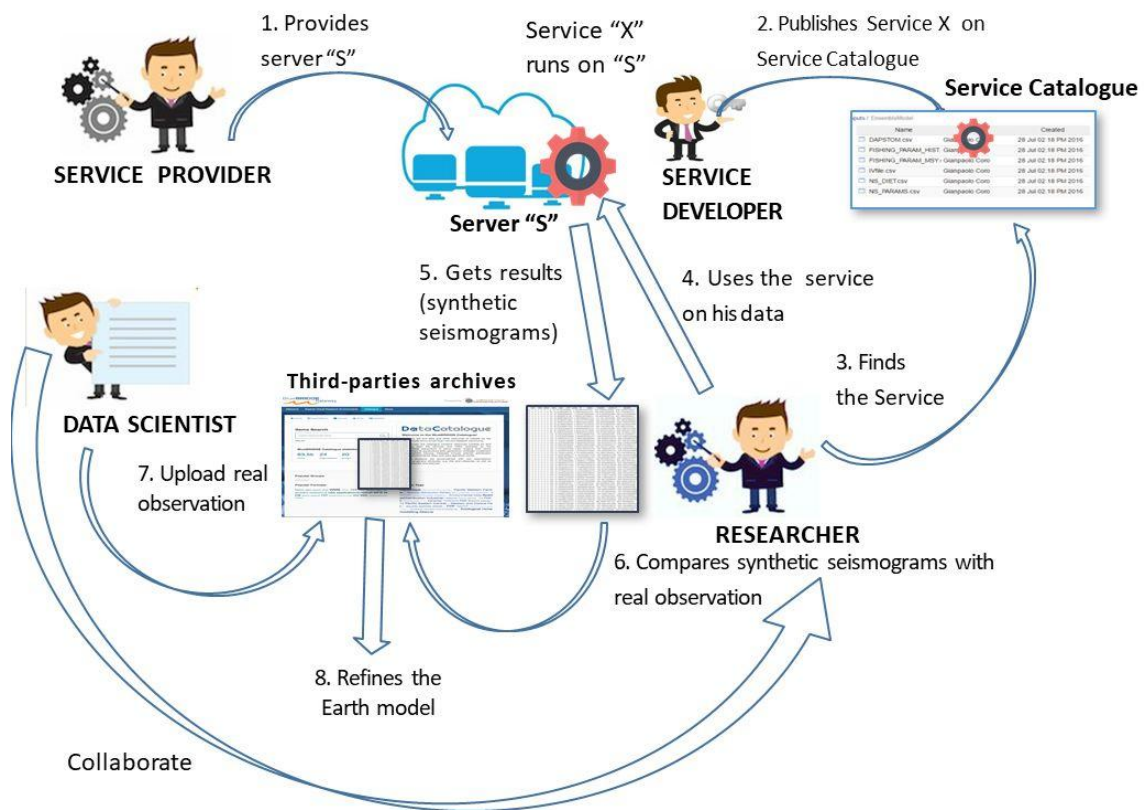


Figure 15. EPOS/VERCE Science Demonstrator

The second example is the **DP-HEP Science Demonstrator**. The overall goal of this Science Demonstrator is to implement a service for the long-term preservation and reuse of HEP data, documentation and associated software. The sources of the data are (potentially) the four main experiments at CERN's Large Hadron Collider (LHC), although it is likely that only data from one experiment will be used. This objective is a collective effort of actors playing the following EOSC roles:

- **Researcher** and any end-user interested in accessing or making use of any of the artifacts preserved by this service. The Researcher will access the data, the software and the documentation to reproduce a simulation. The software will be accessed via CVMFS, while the data and the documentation are stored in EUDAT services.
- **Service Providers** responsible to provide the needed services (e.g. for data preservation and/or distribution) for supporting the implementation of the SD pilot.
- **Community Software Manager** who is someone very familiar with the data, software and documentation and who is able to assess what artifacts need to be preserved to enable reuse over a considered time period. In the case of HEP, the timescale for retaining the data is several decades. For this specific SD, the Community Software Manager ingests the data, the software and the documentation to demonstrate how existing data can be re-used and shared. The Community Software Manager role described in this workflow is specific to this SD pilot, so there is no need to be included in the EOSC users and roles (sec. 2.1).
- **Service Providers** of the B2SHARE and B2SAFE services responsible to operate the service according to the service targets.

The diagram in Figure 16 below, describes how these actors interact each other:

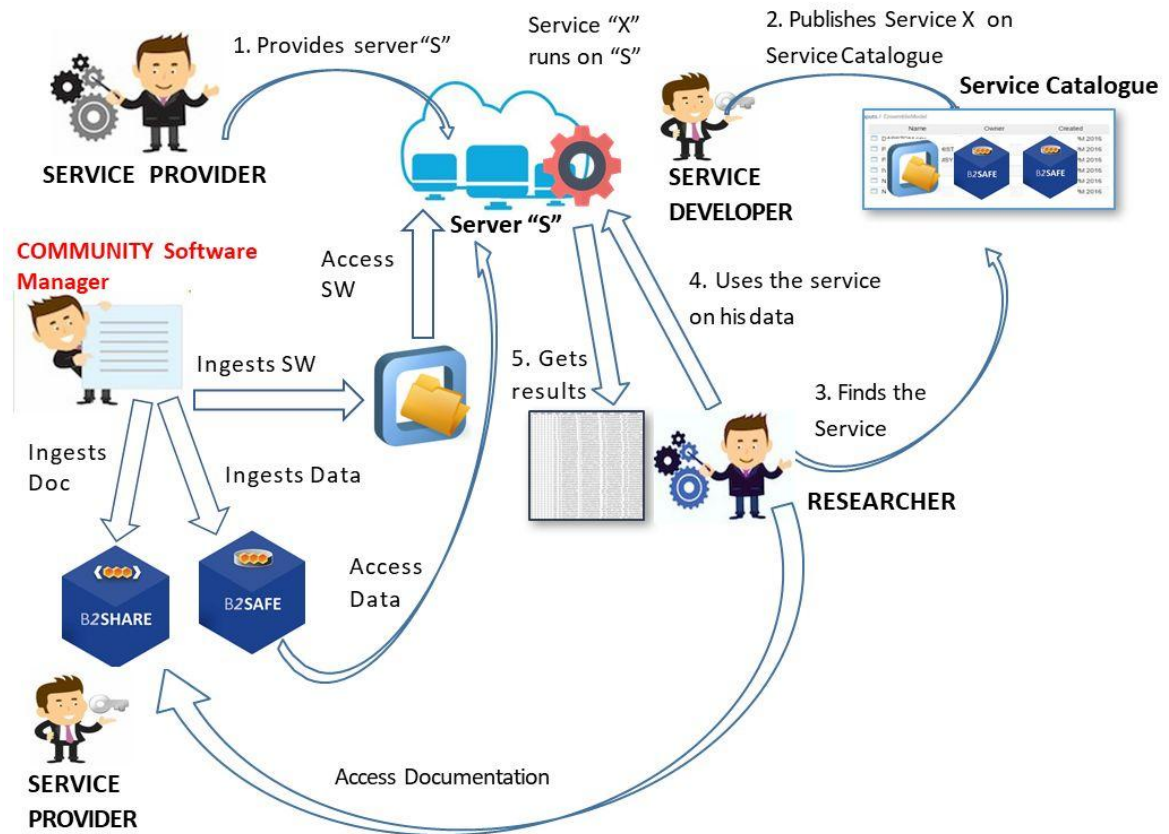


Figure 16. DP-HEP Science Demonstrator