EOSC Future

D3.2b

EOSC Architecture and Interoperability Framework

The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement number 101017536

Version 1.0
June 2023

D3.2b / EOSC Architecture and Interoperability Framework

Lead by **GÉANT**
Authored by Michelle Williams (GEANT), Mark Van de Sanden (SURF), Diego Scardaci (EGI), Paolo Manghi (OpenAIRE), Licia Florio (NORDUnet), Klaas Wierenga (GÉANT)

Reviewed by John Shepherdson (CESSDA), Daan Broeder (CLARIN), Christos Kanellopolous (GÉANT), Roksana Wilk (Cyfronet) and Athanasia Spiliotopoulou (JNP)

Dissemination Level of the Document

Public

Abstract

This deliverable builds on the EOSC Future objectives and principles proposed in the previous Deliverable 3.2a *EOSC Architecture and Interoperability Framework* [1] and describes the work carried out and the features implemented since its publication, including the EOSC IF Governance and EOSC IF Registry and associated supply and demand user interfaces. It provides an overview of the EOSC Architecture and of the EOSC Interoperability Framework, as well as an example of how the EOSC IF can support cross-discipline and cross-domain interoperability.

Version History

Version	Date	Authors/Contributors	Description
V0.1	01/05/2023	Authored by Michelle Williams (GEANT), Mark Van de Sanden (SURF), Licia Florio (NORDUnet), Paolo Manghi (OpenAIRE), Mark Dietrich (EGI), Gavin Farrell (EMBL/ELIXIR)	Initiation – Proposed ToC – First draft
V0.2	19/05/2023	Michelle Williams (GEANT), Mark Van de Sanden (SURF), Paolo Manghi (OpenAIRE), Licia Florio (NORDUnet), Diego Scardaci (EGI), Klaas Wierenga (GÉANT)	Issued for internal Review
V0.3	25/05/2023	Michelle Williams (GEANT), Mark Van de Sanden (SURF), Licia Florio (NORDUnet), Gavin Farrell (EMBL/ELIXIR), Christos Kanellopolous (GÉANT), Mark Dietrich (EGI)	Edits and comments
V0.4	02/06/2023	Roksana Wilk (Cyfronet), Athanasia Spiliotopoulou (JNP)	Edits and comments
V1.0	20/06/2023	Michelle Williams (GEANT), Mark Van de Sanden (SURF), Paolo Manghi (OpenAIRE), Licia Florio (NORDUnet), Diego Scardaci (EGI), Klaas Wierenga (GÉANT), Ron Dekker (TGB), Mike Chatzopoulos (ATHENA)	Final Version submitted to EC

Copyright Notice

Table of Contents

Table of Figures

Table of Tables

Glossary

EOSC Future project Glossary is incorporated by reference: https://wiki.eoscfuture.eu/x/JQCK

List of Abbreviations

Acronym	Definition
AAI	Authentication and Authorisation Infrastructures
AARC	Authentication and Authorisation for Research and Collaboration
AISBL	Association internationale sans but lucratif
API	Application Programming Interface
DOI	Digital Object Identifier
EGI	European Grid Infrastructure
EIAB	EOSC Interoperability Advisory Board
EIAC	EOSC Interoperability Area Chairs
EPOT	EOSC Portal Onboarding Team
ERIC	European Research Infrastructure Consortium
ESFRI	European Strategy Forum on Research Infrastructures
EUDAT	European Data Infrastructure
FAIR	Findable, Accessible, Interoperable, Reusable
HLR	High Level Roadmap
HPC	High Performance Computing
IF	Interoperability Framework
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OCRE	Open Clouds for Research Environments
OpenAIRE	Open Access Infrastructure for Research in Europe
PRACE	Partnership for Advanced Computing in Europe
RI	Research Institute
SME	Small and medium-sized enterprise
TCB	Technical Coordination Board
TRL	Technology Readiness Level
UML	Unified Modelling Language
Acronym	Definition
AAI	Authentication and Authorisation Infrastructures
AARC	Authentication and Authorisation for Research and Collaboration
AISBL	Association internationale sans but lucratif
API	Application Programming Interface

1. Executive Summary

This Deliverable D3.2b *EOSC Architecture and Interoperability Framework* presents the overall EOSC architecture as it has evolved over time and describes the implemented EOSC IF Governance and EOSC IF Registry and associated supply and demand user interfaces. It builds on the work carried out by the Work Package 3 team and reported in D3.2a *EOSC Architecture and Interoperability Framework* [1] and D3.3a *Architecture and Interoperability Guidelines for Operational Services of the EOSC-Core* [2] . This work is introduced and summarised in Section 2 of this document.

Deliverable 3.2a *EOSC Architecture and Interoperability Framework [1]* was a comprehensive document and its scope was (i) to define the EOSC Interoperability Framework (EOSC IF) governance based on the output of the *EOSC Interoperability Framework Report* [3] ; and (ii) to start to define more in detail the EOSC Architecture based on the EOSC Future project's Description of Action and to further analyse various types of interoperability and integration models.

Additional to the Architecture aspects of Deliverable 3.2a, Deliverable D3.3a *Architecture and Interoperability Guidelines for Operational Services of the EOSC-Core* [2] went into further detail and described the architecture of the EOSC-Core Technical Platform (referred to as the EOSC Platform) as a set of internal services that allow EOSC to operate, and with which the researcher-facing services and research products in the EOSC-Exchange can be integrated as desired to take advantage of their capabilities.

Since the publication of D3.2a, the EOSC Future project has progressed the EOSC Architecture from a high-level to a detailed architecture using the C4 abstraction-first model. This detailed architecture, described in Section 3 and in the Appendices to this document, includes context-level diagrams, container-level diagrams, and component-level diagrams developed using the Icepanel tool.

WP3 has made significant progress in implementing the EOSC Interoperability Framework (EOSC IF) and its associated governance. The EOSC Interoperability Framework (EOSC IF) aspires to be a framework of policies, guidelines, standards, and best practices that enable different systems, applications, or technologies to communicate and exchange information with each other. Interoperability in this context refers to the ability of systems within the EOSC Platform and across scientific disciplines to work together, exchange data, and use each other's services, ideally in a seamless and efficient manner.

Section 4 provides and overview of the EOSC IF, introduces the EOSC IF Governance and Registry and describes how these have been developed and implemented. This includes establishing the EIAC (EOSC Interoperability Advisory Committee) and EIAB (EOSC Interoperability Architecture Board) as governance forums, developing the EOSC IF Registry, and feeding into the UI for discovering Interoperability Guidelines in the EOSC Marketplace.

The initial focus has been on getting ready to incorporate the EOSC-Core Interoperability Guidelines into the registry, with plans to include thematic and horizontal guidelines in the future. A dedicated area [4] of the EOSC Portal has been created which collects all information regarding the context and procedures, as well as guidance notes relating to the EOSC Interoperability Framework in a public website. It is now possible for a Provider to systematically register an Interoperability Guideline to the EOSC IF by way of the Provider Dashboard.

The EOSC IF Registry and Governance will continue to evolve based on community feedback and specific use cases. The aim is to promote the concept of proposing Interoperability Guidelines for inclusion in the EOSC IF and improve their utility through iteration and user feedback. Effort will be assigned to the EOSC Interoperability Guideline Profile, which will be tested and expanded to include relevant resource types.

The key benefit of a functional Interoperability Framework for EOSC would be that it could facilitate the cross-discipline collaboration of researchers, providers, and research communities, who have been working on interoperability within and across their infrastructures for many years based on their own community best practices and standards. A real-life example of how cross-discipline interoperability problems can be solved, based on the scenario of the COVID-19 pandemic, and how lessons learned in the process can be shared for the benefit of the wider community, is provided in Section 5.

Section 6 describes how the EOSC IF Registry as deployed will be extended to deliver a proof of concept, to support machine-actionable configuration described in the EOSC-IF Registry, demonstrating the capability with a use case that exposes the OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) interface.

Next steps, summarised in Section 7, include initiating activities to identify interoperability references from research product and service resources, adding Interoperability Guideline references to data resource profiles, and iterating the EOSC IF Registry to enable machine-initiated and machine-enabled interoperability and service composability. The Interoperability Guideline review procedure will be improved, and the end-state Governance structure will be determined as part of the EOSC-INFRA2023-01-05 project.

The transition of relevant parts of the Architecture and Interoperability Framework will be specified to the EOSC Association, operational partners, or relevant consortia. This document and related deliverables (D3.3b, D4.1b and D5.1b) will be utilised in the production of an EOSC Reference Architecture document that will be published before the end of the EOSC Future project.

2. Introduction

2.1 Background

This deliverable builds on the work done by the team and reported in *D3.2a EOSC Architecture and Interoperability Framework* and *D3.3a Architecture and Interoperability Guidelines for Operational Services of the EOSC-Core*.

Deliverable 3.2a *EOSC Architecture and Interoperability Framework* was a comprehensive document and its scope was (i) to define the EOSC Interoperability Framework (EOSC IF) governance based on the output of the EOSC Interoperability Framework Report [3] produced by the EOSC Executive Board FAIR and Architecture Working Groups; and (ii) to start to define more in detail the EOSC Architecture based on the EOSC Future project's Description of Action and to further analyse various types of interoperability and integration models across the landscape, focusing on the aspects related to interoperability in the EOSC-Core and EOSC-Exchange platforms.

Deliverable 3.2a defined an outline governance process for the EOSC IF alongside draft procedures for proposing and registering an Interoperability Guideline. It also reported on the analysis done in Work Package 3 (WP3) to collect science use cases to identify requirements and gaps in the EOSC-IF. It further explained how WP3 Working Groups were chartered to produce interoperability outcomes to support the requirements of the science use cases identified. Deliverable 3.4 *Working Group Outputs and Mediated Requirements* goes on to describe in detail the outcome of the Working Groups that were chartered by EOSC Future WP3, so these outcomes will not be repeated in this deliverable (D3.2b). Further descriptions of the interoperability between research collaborations and between e-Infrastructures and research collaborations were also included in that document.

Since the first deliverable was produced, the EOSC IF Governance has been launched, comprising the EOSC Interoperability Advisory Board (**EIAB**), and the EOSC Interoperability Area Chairs (**EIAC**). A dedicated area [4] of the EOSC Portal has been created which collects all information regarding the context and procedures, as well as guidance notes relating to the EOSC Interoperability Framework in a public website. It is now possible for a Provider to systematically register an Interoperability Guideline to the EOSC IF by way of the Provider Dashboard, which populates the **EOSC IF Registry** database, which is itself an integrated component of the EOSC Resource Catalogue. The governance procedures for considering newly proposed Interoperability Guidelines are initiated by a review procedure that is overseen by the well-established EOSC Portal Onboarding Team (EPOT). The EOSC IF Registry will host EOSC-Core and EOSC-Exchange Interoperability Guidelines (which can be either thematic or horizontal guidelines), and a capability is being developed that will support executable configuration templates, enabling machine-readable Interoperability Guidelines that will also be part of the Registry. This capability will be demonstrated with a use case around research product repositories exposing an OAI-PMH interface.

As noted, the **EOSC IF Registry** is implemented as a component of the EOSC Resource Catalogue. Its core feature is the implementation of a new EOSC Profile: the EOSC Interoperability Guideline Profile, which was defined in the scope of the conceptual work around the EOSC IF. The new profile's metadata schema describes Interoperability Guidelines, where the guideline documents themselves are hosted in public repositories and maintained within their related community (declared via a dedicated attribute in the schema which requires the document's DOI). The advantage of creating a registry of Interoperability Guideline *metadata* is that the registry does not have to become a repository in its own right and means that users can continue to utilise preferred repositories for publication of Guidelines as appropriate to their needs. This new profile makes it

possible to reference APIs, machine-executable code (or transformation services hosted in repositories such as Git, for example) within the metadata as well as within the document itself.

Since D3.2a was submitted, the EOSC Profiles v4.0.0 [5] have been extended to encompass the *EOSC Multi-Provider Catalogue Profile*, which facilitates automated onboarding of resources from external catalogues via a dedicated API, as well as profiles to onboard Training Resources and the Interoperability Guidelines described above. In addition, the existing Service Profile[1] has been enhanced to allow definition of a "sub profile" for a Service that describes a Data Source capability of a given service. Future sub-profiles are planned to enhance the Service Profile to describe compute and data storage services in more detail. Further information on this will be provided in Deliverable 3.3b (due to be submitted at the end of June 2023).

Additional to the Architecture aspects of Deliverable 3.2a, Deliverable D3.3a *Architecture and Interoperability Guidelines for Operational Services of the EOSC-Core* [2] went into further detail and described the architecture of the **EOSC-Core Technical Platform** (referred to as the EOSC Platform) as a set of internal services that allow EOSC to operate, and with which the researcher-facing services and research products in the EOSC-Exchange can be integrated as desired to take advantage of their capabilities. D3.3a included the definition of the interactions and interfaces both between components and towards the external world, to implement the use cases that the platform should support, as described in the project High Level Roadmap (HLR). User communities have been directly involved in this effort in line with co-design principles.

D3.3a further introduced the **C4 methodology** which was chosen by the EOSC Future Technical Coordination Board (TCB) after considering the currently available architectural methodologies and related tools, to address the requirement to provide a functional description of the EOSC Architecture. The adoption of the C4 methodology allowed the project to create a consistent, integrated architectural model that also enabled visualisation of the software architecture. The deliverable illustrated the container diagram of the EOSC Platform and included views that highlighted logical connections between components and standards/interfaces adopted to enable components communication; it also included a component diagram for the "Catalogue and Marketplace" container, illustrating the contained components and their interactions with external systems.

After the publication of D3.3a, the team continued work on the items reported below:

- Further refining the logical and standards/interfaces views of the EOSC Platform.
- Developing a component diagram for each of the EOSC Platform containers (e.g. all diagrams presented in this deliverable will be redrawn using C4).
- Developing process view diagrams to show how the different components of the platform work together to implement a use case (e.g. the service provider on-boarding process, AAI interactions, etc.).
- Enriching descriptions in all diagrams.

2.2 Scope of this Deliverable

Deliverable *D3.2b EOSC Architecture and Interoperability Framework* presents the overall EOSC architecture as it evolved over time and in more detail using the C4 methodology and describes the EOSC IF Governance and Registry as implemented, referring back to the original proposals made in Deliverable 3.2a; the C4 methodology is expanded in Appendix A. Further details of the EOSC Platform architecture such as data flows between components and EOSC-Core Interoperability Guidelines will be described in D3.3b.

A governance process has been defined which applies the inclusion criteria (Appendix C) for becoming part of the EOSC IF. Appendix D describes the procedures that have been put in place for requesting inclusion in the EOSC Interoperability Framework, while section 4.1 describes the governance that oversees the process, which is designed to ensure that a fair and transparent assessment is applied to all requests, whether from within the project or from the wider scientific community.

[1] At the time of writing, the project is transitioning the previously defined "Resource Profile" as a "Service Profile" in order to clarify its use to describe only services. Functionality related to the existing "Resource Profile" will be deprecated in the future.

The final deliverables from the technical delivery work packages of the EOSC Future project are interrelated, and can be considered in the following context:

- *D4.1b Back-Office design, functional and technical specifications* [30] and *D5.1b Front-Office Design, Functional and Technical Specifications [31]* provide details of the functional architecture of their respective components.
- *D3.3b Architecture and Interoperability Guidelines for Operational Services of the EOSC-Core* will look more in detail at the EOSC Platform architecture and present in-depth descriptions of the EOSC Profiles and EOSC-Core interoperability guidelines for services in the scope of T3.2.
- *D3.4b Working Group Outputs and Mediated Requirements* will describe in detail the EOSC Future Working Group outputs along with the challenges encountered along the way and how these have been addressed.

3. Overview of the EOSC Architecture

In deliverable D3.2a, the EOSC Architecture was described at a high-level, identifying its different elements (i.e. EOSC-Core, EOSC-Exchange, EOSC Interoperability Framework, EOSC Support activities and Science Clusters and Communities). The concept of the EOSC Architecture has evolved from an EOSC platform comprising a set of services offering EOSC capabilities at European level, to an architecture model identifying:

- EOSC platform stakeholders,
- EOSC Platform services,
- EOSC Providers systems,
- Interconnections between services within the EOSC platform, and
- Connections between EOSC provider systems and the EOSC Platform to make resources available through EOSC.

The EOSC architecture model will be described in detail in the next sections.

3.1 Defining the EOSC Platform at the Contextual Level

As described on the C4 model website [6] : "the C4 model is an "abstraction-first" approach to diagramming software architecture, based upon abstractions that reflect how software architects and developers think about and build software". The aim of the C4 model is to define readable architecture diagrams.

In the C4 model the **context level** draws the system as a box in the centre surrounded by its users and the other systems that it interacts with. In the context of EOSC, the EOSC Platform is the system and is surrounded by its various stakeholders that make use of or offer resources through the Platform. The contextual level of the EOSC Platform is in turn described at two different levels:

- **EOSC Overview** – places the EOSC Platform in the centre surrounded by the main actors (i.e. consumer, service provider, research product provider and external catalogue owner) as its users.
- **EOSC Stakeholder view** – identifies the different types of stakeholders within the various actor domains identified.

A more detailed description of the C4 methodology is provided in Appendix A.

3.1.1 EOSC Overview

This diagram provides a high-level overview of the EOSC Platform and its main actors. The EOSC Platform is the engine that enables EOSC operations. It consists of services such as the comprehensive Resource Catalogues, the Marketplace, the Infrastructure Proxy for the EOSC Core services, the AAI Registry supporting the AAI Federation, the Helpdesk, the Monitoring, the Accounting, and the Order Management System. It also identifies the main actors of the EOSC Platform as:

- Consumer – searches for/discovers/orders/accesses/uses resources made available through the platform;
- Service Provider –offers services through the platform;

- Research Product provider – uses the platform offering and manages research products to support Research and Open Science;
- External Catalogue Owner – uses the platform to offer the services and resources made available through its own catalogue through EOSC.

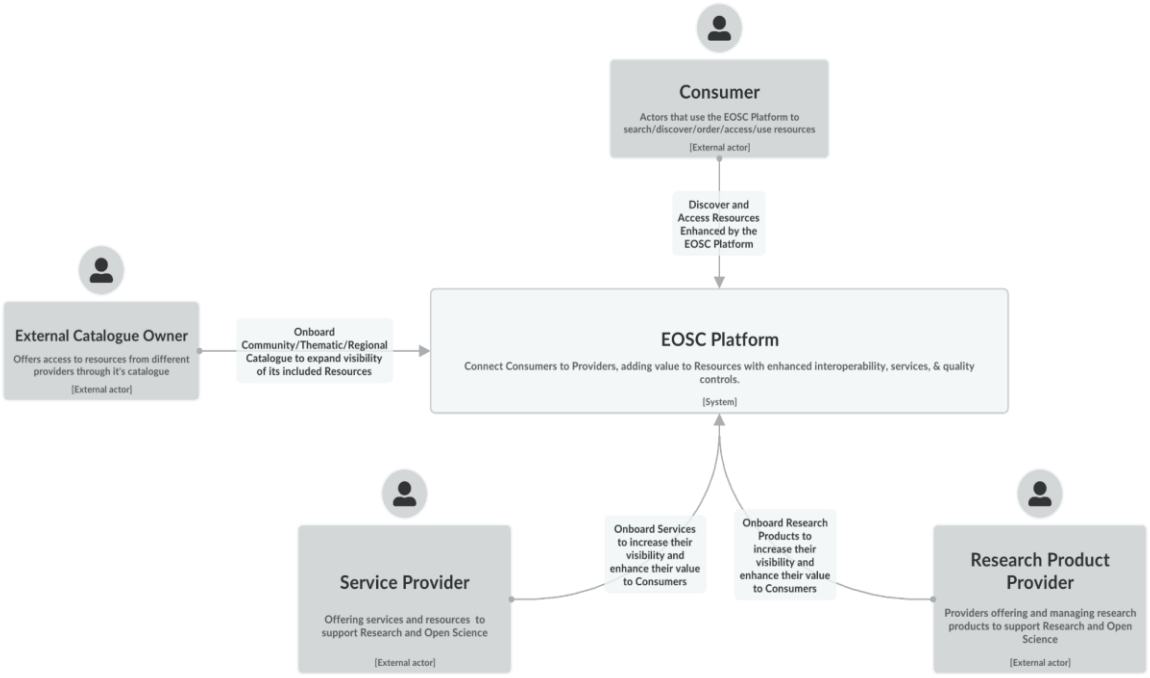


Figure 3.1: High-level contextual overview of the EOSC Platform and its external stakeholders

3.1.2 EOSC Stakeholder Diagram

The main scope of the EOSC Stakeholder diagram is to identify the main stakeholders, and actors, of the EOSC Platform. Deliverable D10.1 *EOSC Future Stakeholder Engagement and Communication Strategy and Plan* [7] identified in section 2 the users as *"who consume EOSC data or resources and those that contribute to the co-design of EOSC Future solutions by providing resources, sharing requirements, or testing solutions. Stakeholders have therefore been grouped into three categories: consumers, providers, facilitators/intermediaries."*

From an architectural point of view, the focus has been on those stakeholders who consume EOSC data and resources (i.e. Consumers) and on those who make resources available through EOSC (i.e. Providers or External Catalogue Owners). At the time of writing D10.1 at M6 (September 2021) of the project, the External Catalogue Owners were not yet identified as a separate category stakeholder. While developing the EOSC Platform architecture diagrams in collaboration with external catalogue owners (e.g. INFRA-EOSC-05-2018-2019, INFRA-EOSC-05-2018-2019 [8] projects [9] and INFRA-EOSC-04-2018 [10] projects [11]) it became clear that it was necessary to include external catalogue owners as a separate category of providers who make resources available through EOSC because their user journey of enriching EOSC Catalogue with new services differs too much from a single provider use case to be included within the same stakeholder type.

Types of Consumers who search for/discover/order/access/use resources made available through the platform include:

- *Research Communities* – research groups, networks, laboratories, or representative bodies at the disciplinary level that can act as a conduit such as academic associations, learned societies and Research Infrastructures.
- *Researchers* – individual researchers.
- *Citizen Scientists* – Individuals as citizen scientists, including through platforms like Zooniverse.
- *Commercial Entities* – Private companies interested in using the EOSC services.

The Types of Providers who offer resources (e.g. services and research products via data sources) through the platform are:

- *e-Infrastructures* – EGI, EUDAT, GÉANT, OpenAIRE, PRACE, national e-Infrastructures, HPC and cloud initiatives.
- *Research Infrastructures* – ESFRI science clusters, thematic clouds, ERICs, EIROforum RIs, AISBLs, *national Research Infrastructures, national labs.*
- *Private sector* – commercial service providers, business organisations, SMEs, start-ups, innovators
- *Others* – organisations, community groups or other stakeholder categories including, for example, open-source communities.

Types of External Catalogue Owners include those connecting the following regional and/or thematic catalogues to the EOSC platform to onboard and make resources available through them in EOSC:

- *Regional external catalogues* – External catalogues connected to the EOSC Platform to make regional and/or national resources available through EOSC.
- *Thematic external catalogues* – External catalogues connected to the EOSC Platform to make thematic resources available through EOSC.

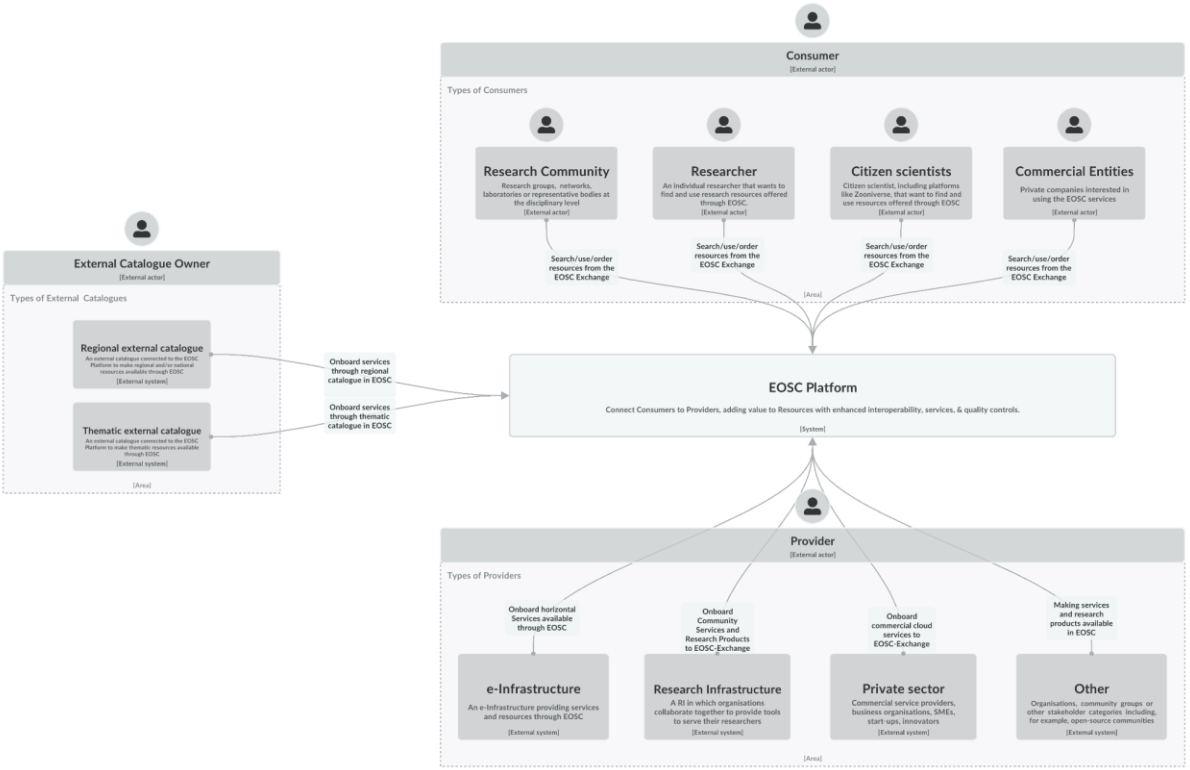


Figure 3.2: High-level contextual overview of the EOSC Platform and its external stakeholders

3.2 Defining the EOSC Platform at the Container Level

In the C4 model, the container diagram shows the high-level shape of the software architecture and how responsibilities are distributed across it. It also shows the primary technologies in operation and how the containers communicate with one another. The container-level diagram should be a simple, high-level technology-focussed diagram that is useful for software developers and support/operations staff alike.

In the context of architectural modelling of the EOSC Platform, at the container-level the individual components inside the EOSC Platform system box defined at the context level are identified and described. Because of the complexity of the Overall EOSC architecture, see Diagram 1, Appendix B, this view is split into three diagrams illustrating the 3 major components comprising the architecture, i.e. EOSC Platform, EOSC Exchange and EOSC Provider, which are described separately in the next sections.

3.2.1 EOSC Platform

This diagram provides a technical overview of the EOSC Platform and its containers. It shows services such as the comprehensive Resource Catalogue and the Marketplace, the Core Infrastructure Proxy for the EOSC Core services, the AAI Registry supporting the AAI Federation, the Helpdesk, Monitoring, Accounting, Execution Framework, Interoperability Framework Registry, Configuration Management System, and Collaboration Tools (e.g. wiki, ticket and email system).

At the container level, the interconnections between the containers within the platform are described including the standards used to integrate the containers.

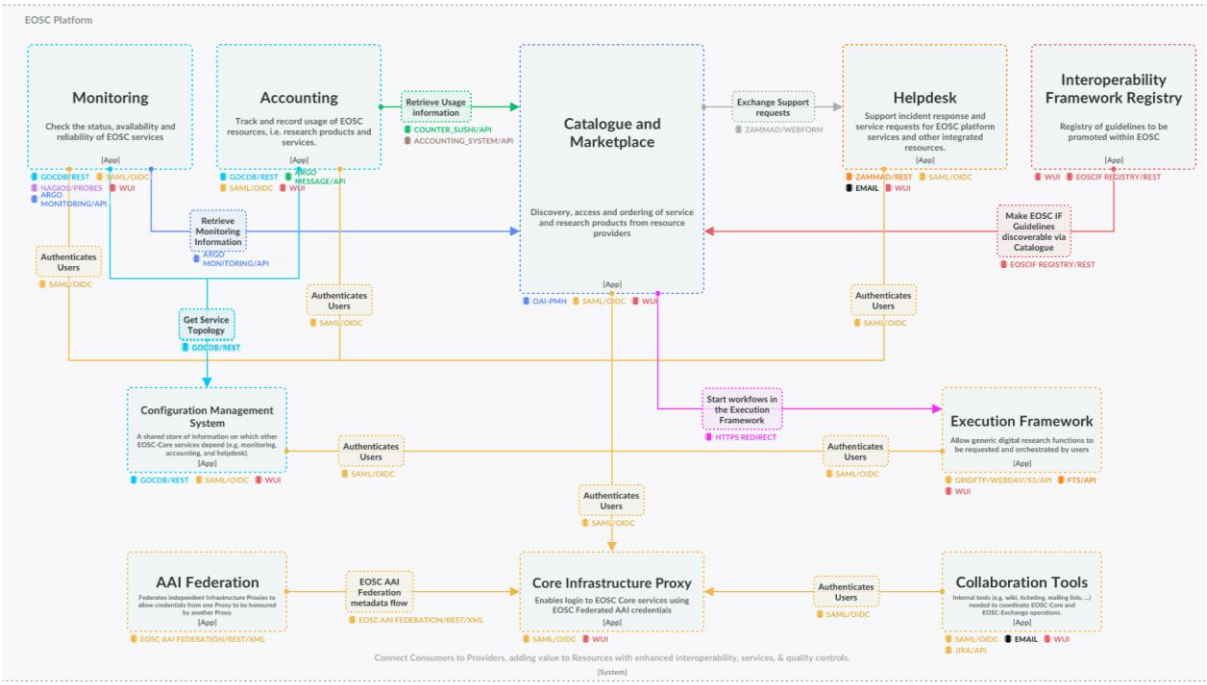


Figure 3.3:EOSC Platform describing the platform containers and the interconnections between the containers

Table 3-1 provides an overview of the EOSC Platform containers and the interconnections between them, the direction in which the information flow is being initiated and the Interface/Protocol used for interaction.

Table 3-1: Overview of EOSC Platform components and interconnections with other EOSC Platform components

EOSC Platform container	Direction	Target container	Information flow	Interface/ Protocol
Catalogue and Marketplace	←	EOSC Platform: Monitoring	Retrieve Monitoring Information	ARGO MONITORING/API
	←	EOSC Platform: Accounting	Retrieve Usage Information	Research Products accounting: COUNTER_SUSHI/API, Service accounting: ACCOUNTING_SYSTEM/ API
	→	Provider: Helpdesk	Exchange Support requests	ZAMMAD/WEBFORM
	←	EOSC Platform: Interoperability Framework Registry	Make EOSC IF Guidelines discoverable via Catalogue	EOSCIF REGISTRY/REST

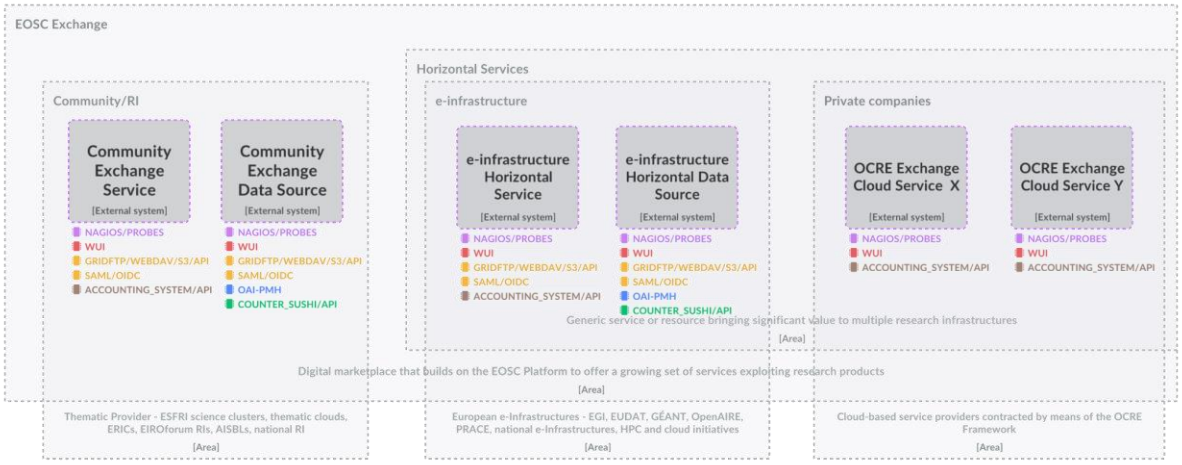
EOSC Platform container	Direction	Target container	Information flow	Interface/ Protocol
	→	EOSC Platform: Core Infrastructure Proxy	Authenticates Users	SAML/OIDC
Helpdesk	←	EOSC Platform: Catalogue and Marketplace	Exchange Support Requests	ZAMMAD/WEBFORM
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Accounting	→	EOSC Platform: Catalogue and Marketplace	Retrieve Usage Information	Research Products accounting: COUNTER_SUSHI/API, Service accounting: ACCOUNTING_SYSTEM/ API
	→	EOSC Platform: Configuration Management System	Get Service Topology	GOCDB/REST
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Monitoring	←	EOSC Platform: Catalogue and Marketplace	Retrieve Monitoring Information	ARGO MONITORING/API
	→	EOSC Platform: Configuration Management System	Get Service Topology	GOCDB/REST
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Configuration Management System	←	EOSC Platform: Monitoring	Get Service Topology	GOCDB/REST
	←	EOSC Platform: Accounting	Get Service Topology	GOCDB/REST
	→	EOSC Platform: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Core Infrastructure Proxy	←	EOSC Platform: Services (ALL)	Authenticates Users	SAML/OIDC
	→	EOSC Platform: AAI Federation	EOSC AAI Federation metadata flow	EOSC AAI FEDERATION/REST/XML
AAI Federation	→	EOSC Platform: Core Infrastructure Proxy	EOSC AAI Federation metadata flow	EOSC AAI FEDERATION/REST/XML
Interoperability Framework Registry	→	EOSC Platform: Catalogue and Marketplace	Make EOSC IF Guidelines Discoverable via Catalogue	EOSCIF REGISTRY/REST
Collaboration Tools	→	EOSC Platform: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Execution Framework	→	EOSC Platform: Catalogue and Marketplace	Start workflows in the Execution Framework	HTTPS REDIRECT

3.2.2 EOSC Exchange

The EOSC Exchange has been depicted as the set of services storing and exploiting FAIR data (e.g. research products) and encouraging its reuse within EOSC. The set of services and research products are made available through different types of providers:

- Community/RI – make community-based services and research products available through EOSC;
- e-infrastructures – make horizontal services and research products available through EOSC;
- Private companies – commercial cloud providers who make horizontal resources available to EOSC, for example through the OCRE Framework.

The service and research products made available through e-infrastructures and the OCRE Framework are indicated as horizontal services. Horizontal services are generic services or resources bringing significant value to multiple research infrastructures.



shows an abstract view of the EOSC Exchange and Horizontal services as a section of the EOSC Exchange. Exchange resources have been defined as either thematic resources offered by Research Communities and/or horizontal resources offered by e-infrastructures and/or private companies, such as for example those offered through the OCRE Cloud Framework. The diagram also relates the resources to the platform.

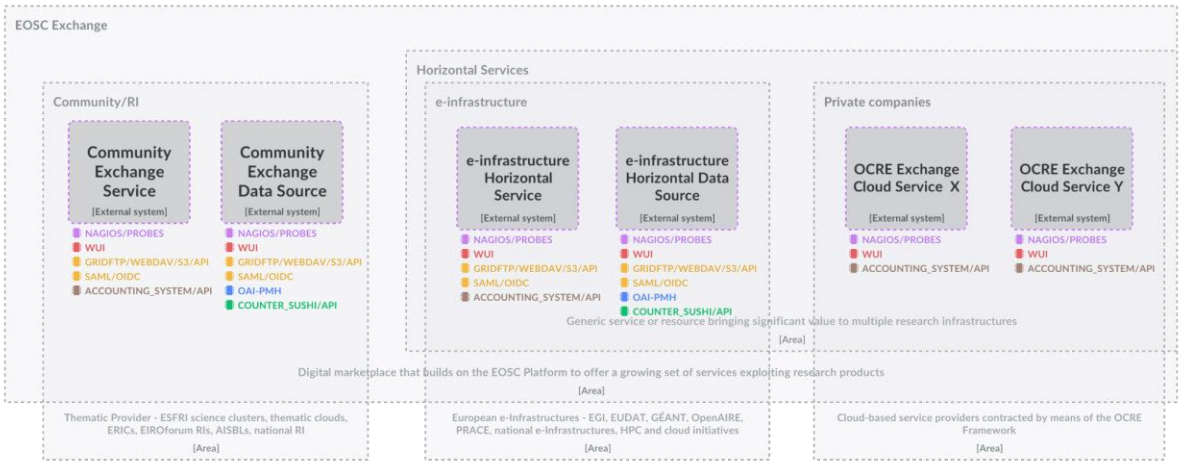


Figure 3.4: EOSC Exchange showing horizontal services and resources

Table 3-2 provides an overview of resources made available through the EOSC Exchange and their interconnections and APIs.

Table 3-2: Overview of EOSC Exchange resources and their interconnections with the EOSC Platform

Provider system	Direction	Target container	Information flow	Interface/ Protocol
Community Exchange Service	←	EOSC Platform: Monitoring	Monitor QoS	NAGIOS/PROBES
	→	EOSC Platform: Accounting	Publish Resource Usage	ACCOUNTING_SYSTEM/API
	↔	EOSC Platform: Execution Framework	Support Data Transfers	GRIDFTP/WEBDAV/S3/API
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Community Exchange Data Source	←	EOSC Platform: Monitoring	Monitor QoS	NAGIOS/PROBES
	→	EOSC Platform: Accounting	Publish Resource Usage	COUNTER_SUSHI/API
	←	EOSC Platform: Catalogue and Marketplace	Harvest metadata	OAI-PMH
	↔	EOSC Platform: Execution Framework	Support Data Transfers	GRIDFTP/WEBDAV/S3/API
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
e-Infrastructure Exchange Service	←	EOSC Platform: Monitoring	Monitor QoS	NAGIOS/PROBES
	→	EOSC Platform: Accounting	Publish Resource Usage	ACCOUNTING_SYSTEM/API
	↔	EOSC Platform: Execution Framework	Support Data Transfers	GRIDFTP/WEBDAV/S3/API
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
e-infrastructure Exchange Data Source	←	EOSC Platform: Monitoring	Monitor QoS	NAGIOS/PROBES
	→	EOSC Platform: Accounting	Publish Resource Usage	COUNTER_SUSHI/API
	←	EOSC Platform: Catalogue and Marketplace	Harvest metadata	OAI-PMH
	↔	EOSC Platform: Execution Framework	Support Data Transfers	GRIDFTP/WEBDAV/S3/API
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC

Provider system	Direction	Target container	Information flow	Interface/ Protocol
OCRE Exchange Cloud Service	←	EOSC Platform: Monitoring	Monitor QoS	NAGIOS/PROBES
	→	EOSC Platform: Accounting	Publish Resource Usage	ACCOUNTING_SYSTEM/API

3.2.3 EOSC Provider

An EOSC Provider is an organisation making resources available through EOSC which can be either services or research products or both. To make resource products discoverable through EOSC, providers must first onboard the data source and then onboard the research products through harvesting of the data source, as research products cannot be onboarded directly.

An EOSC Provider organisation is in itself a complex actor consisting of different systems Figure 3.5 shows the components and IT interfaces of a Provider organisation that need to be connected to the various EOSC Platform capabilities to enable them to make resources available through EOSC.

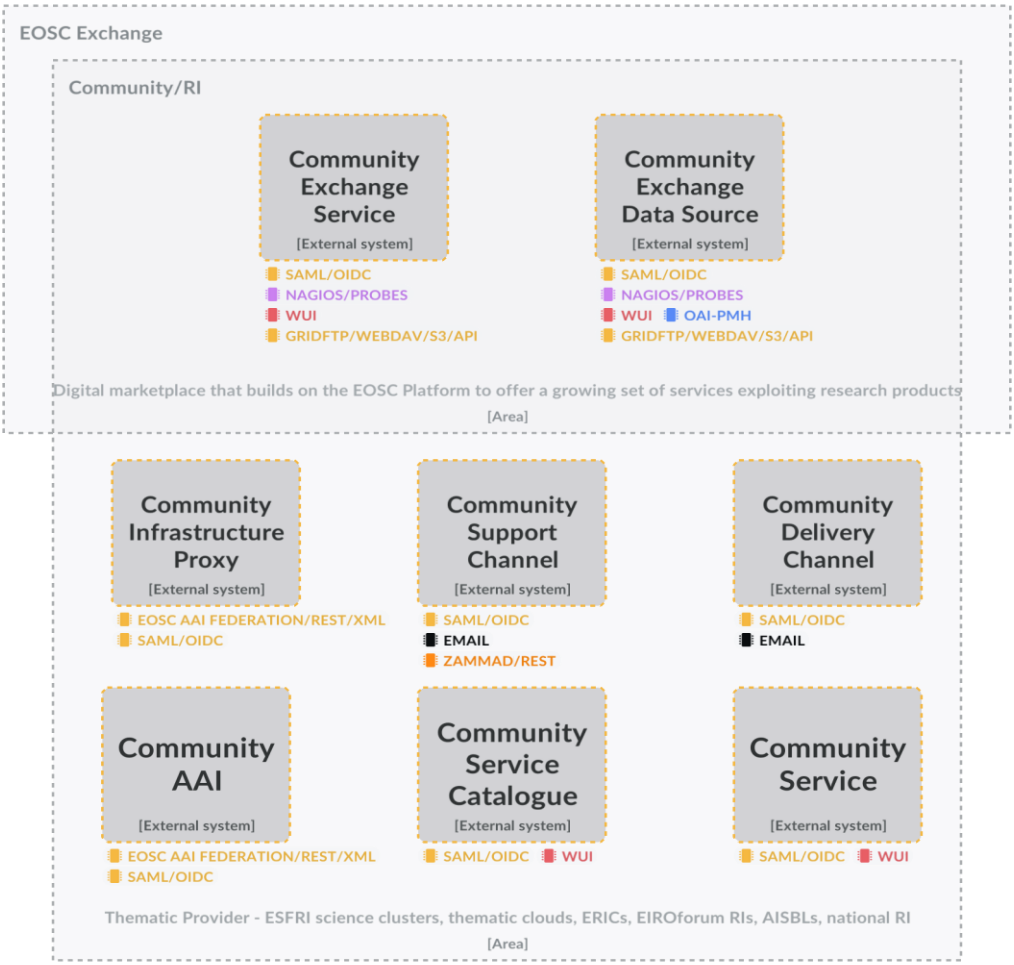


Figure 3.5: EOSC Provider organisation containers that need to be connected to the EOSC Platform

Table 3-3 provides an overview of EOSC Provider systems, their interconnection and APIs with the EOSC Platform, both with local systems and with thematic provider systems.

Table 3-3: Overview EOSC Provider containers and their interconnections with the EOSC Platform

Provider systems	Direction	Target container	Information flow	Interface/ Protocol
Exchange Service	←	EOSC Platform: Monitoring	Monitor QoS	NAGIOS/PROBES
	→	EOSC Platform: Accounting	Publish Resource Usage	ACCOUNTING SYSTEM API
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Exchange Data Source	←	EOSC Platform: Monitoring	Monitor QoS	NAGIOS/PROBES
	→	EOSC Platform: Accounting	Publish Resource Usage	COUNT
	←	EOSC Platform: Catalogue and Marketplace	Harvest metadata	OAI-PMH
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Service Catalogue	→	EOSC Platform: Catalogue and Marketplace	Onboard Services into EOSC	PROVIDER PORTAL/REST
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Support Channel	←	EOSC Platform: Helpdesk	Exchange support requests	EMAIL, ZAMMAD/REST
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Delivery Channel	←	EOSC Platform: Catalogue and Marketplace	Exchange Order Requests	EMAIL
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Infrastructure Proxy	→	EOSC Platform: AAI Federation	EOSC AAI Federation metadata flow	EOSC AAI FEDERATION/REST/ XML
	←	Other: Provider AAI	Authenticates Users	SAML/OIDC
	←	Provider: Services (ALL)	Authenticates Users	SAML/OIDC
Provider AAI	→	EOSC Platform: AAI Federation	EOSC AAI Federation metadata flow	EOSC AAI FEDERATION/REST/ XML

3.3 Defining the EOSC Catalogue and Marketplace at Component Level

In the C4 model, the component diagram zooms into and decomposes each of the containers further to identify their major building blocks and their interactions. In the context of architecture modelling of the EOSC Platform, the component level looks inside the individual containers defined at the container level of the EOSC Platform.

Because of the complexity of the EOSC Catalogue and Marketplace, see Diagram 2, Appendix B, this diagram is split into 2 views, i.e. EOSC Catalogue and Marketplace – Internal view and EOSC Catalogue and Marketplace – Provider view, which are described separately in the next sections.

3.3.1 EOSC Catalogue and Marketplace – Internal View

To explain the component level within the EOSC Platform architecture a detailed description of the Catalogue and Marketplace is provided, showing the internal connections between the Catalogue and Marketplace components as well as the connections with the other capabilities of the EOSC Platform, see Figure 3.6.

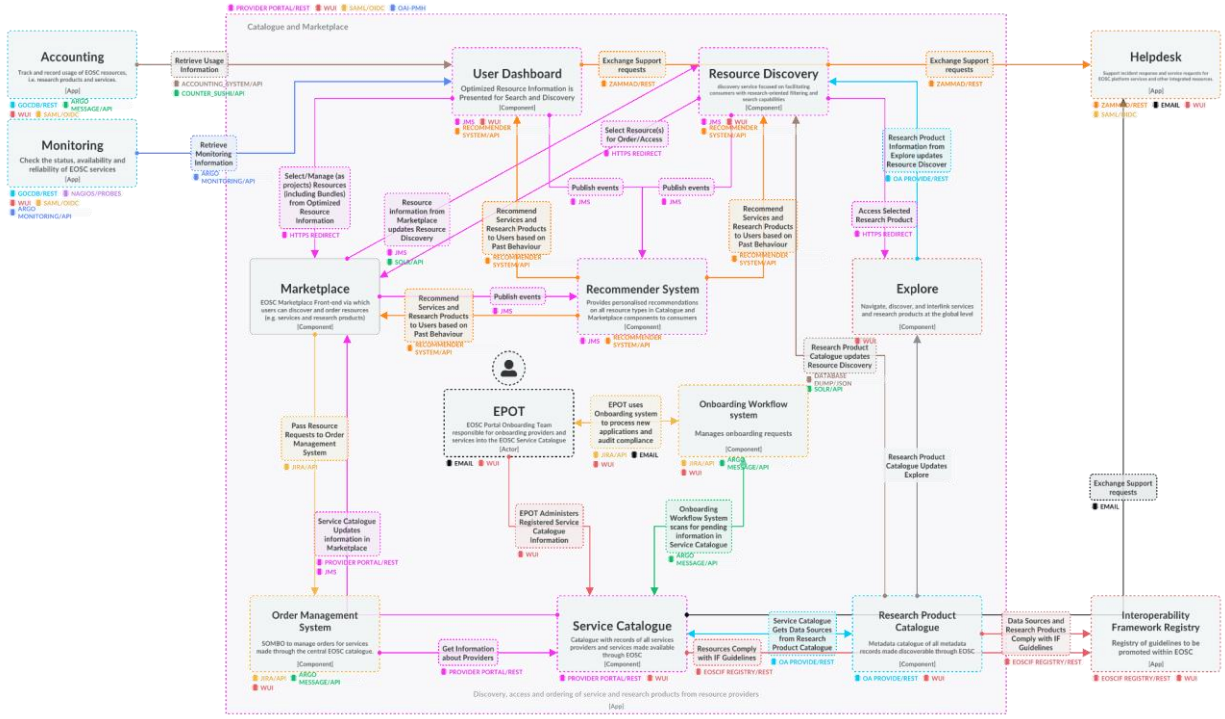


Figure 3.6: EOSC Catalogue and Marketplace internal component level diagram

Table 3-4 provides an overview of the EOSC Catalogue and Marketplace components and the interconnections between them, the direction in which the information flow is initiated and the Interfaces/APIs used for interaction.

Table 3-4: Overview the EOSC Catalogue and Marketplace components and the interconnections with other components and EOSC Platform containers

Catalogue and Marketplace components	Direction	Target Catalogue and Marketplace component or EOSC Platform container	Information flow	Interface/ Protocol
User Dashboard	→	Marketplace	Select/Manage (as projects) Resources (including Bundles) from Optimised Resource Information	HTTPS REDIRECT

Catalogue and Marketplace components	Direction	Target Catalogue and Marketplace component or EOSC Platform container	Information flow	Interface/ Protocol
	→	Recommender System	Publish events	JMS
	←		Recommend Services and Research Products to Users based on Past Behaviour	RECOMMENDER SYSTEM/API
	→	EOSC Platform: Helpdesk	Exchange Support requests	ZAMMAD/REST
	←	EOSC Platform: Monitoring	Retrieve Monitoring Information	ARGO MONITORING/API
	←	EOSC Platform: Accounting	Retrieve Usage Information	Research Products accounting: COUNTER_SUSHI/API, Service accounting: ACCOUNTING_SYSTEM/API
	→	EOSC Platform: Core Infrastructure Proxy	Authenticates Users	SAML/OIDC
Resource Discovery	←	Marketplace	Resource information from Marketplace updates Resource Discovery	JMS, SOLR/API
	→		Select Resource(s) for Order/Access	HTTPS REDIRECT
	→	Recommender System	Publish events	JMS
	←		Recommend Services and Research Products to Users based on Past Behaviour	RECOMMENDER SYSTEM/API
	←	Research Product Catalogue	Research Product Catalogue updates Resource Discovery	DATABASE DUMP/JSON, SOLR/API
	→	Explore	Access Selected Research Product	HTTPS REDIRECT
	←		Research Product Information from	OA PROVIDE/REST

Catalogue and Marketplace components	Direction	Target Catalogue and Marketplace component or EOSC Platform container	Information flow	Interface/ Protocol
			Explore updates	
	→	EOSC Platform: Helpdesk	Exchange Support requests	ZAMMAD/REST
	→	EOSC Platform: Core Infrastructure Proxy	Authenticates Users	SAML/OIDC
Marketplace	←	User Dashboard	Select/Manage (as projects) Resources (including Bundles) from Optimised Resource Information	HTTPS REDIRECT
	→	Resource Discovery	Resource information from Marketplace updates Resource Discovery	JMS, SOLR/API
	←		Select Resource(s) for Order/Access	HTTPS REDIRECT
	→	Recommender System	Publish events	JMS
	←		Recommend Services and Research Products to Users based on Past Behaviour	RECOMMENDER SYSTEM/API
	→	Order Management System	Pass Resource Requests to Order Management System	JIRA/API
	←	Service Catalogue	Service Catalogue Updates information in Marketplace	PROVIDER PORTAL/REST, JMS
	→	Provider: Infrastructure Proxy	Authenticates Users	SAML/OIDC
Recommender System	←	User Dashboard	Publish events	JMS
	→		Recommend Services and Research Products to Users based on Past Behaviour	RECOMMENDER SYSTEM/API

Catalogue and Marketplace components	Direction	Target Catalogue and Marketplace component or EOSC Platform container	Information flow	Interface/ Protocol
	←	Resource Discovery	Publish events	JMS
	→		Recommend Services and Research Products to Users based on Past Behaviour	RECOMMENDER SYSTEM/API
	←	Marketplace	Publish events	JMS
	→		Recommend Services and Research Products to Users based on Past Behaviour	RECOMMENDER SYSTEM/API
Explore	←	Resource Discovery	Research Product Information from Explore updates	OA PROVIDER/REST
	→		Access Selected Research Product	HTTPS REDIRECT
	→	Research Product Catalogue	Research Product Catalogue Updates Explore	OA PROVIDE/REST
EPOT (EOSC Portal Onboaring Team) (Actor)	↔	Onboarding Workflow system	EPOT uses Onboarding system to process new applications and audit	JIRA/API, WUI, ARGO MESSAGE/API
	←	Service Catalogue	EPOT Administers Registered Service Catalogue Information	WUI
Onboarding Workflow System	↔	EPOT	EPOT uses Onboarding system to process new applications and audit	JIRA/API, WUI, ARGO MESSAGE/API
	↔	Service Catalogue	Onboarding Workflow System scans for pending information in Service Catalogue	ARGO MESSAGE/API

Catalogue and Marketplace components	Direction	Target Catalogue and Marketplace component or EOSC Platform container	Information flow	Interface/ Protocol
	→	EOSC Platform: Core Infrastructure Proxy	Authenticates Users	SAML/OIDC
Order Management System	←	Marketplace	Pass Resource Requests to Order Management System	JIRA/API
	→	Service Catalogue	Get Information about Providers	PROVIDER PORTAL/REST
	→	EOSC Platform: Core Infrastructure Proxy	Authenticates Users	SAML/OIDC
Service Catalogue	→	Marketplace	Service Catalogue Updates information in Marketplace	PROVIDER PORTAL/REST, JMS
	←	EPOT	EPOT Administers Registered Service Catalogue Information	WUI
	↔	Onboarding Workflow System	Onboarding Workflow System scans for pending information in Service Catalogue	ARGO MESSAGE/API
	←	Order management System	Get Information about Providers	PROVIDER PORTAL/REST
	↔	Research Product Catalogue	Service Catalogue Gets Data Sources from Research Product Catalogue	OA PROVIDE/REST
	→	EOSC Platform: Interoperability Framework Registry	Resources Comply with IF Guidelines	EOSCIF REGISTRY/API
	→	EOSC Platform: Helpdesk	Exchange Support requests	EMAIL
	→	EOSC Platform: Core Infrastructure Proxy	Authenticates Users	SAML/OIDC
Research Product Catalogue	→	Resource Discovery	Research Product Catalogue updates Resource Discovery	DATABASE DUMP/JSON, SOLR/API

Catalogue and Marketplace components	Direction	Target Catalogue and Marketplace component or EOSC Platform container	Information flow	Interface/ Protocol
	→	Explore	Research Product Catalogue Updates Explore	OA PROVIDE/REST
	↔	Service Catalogue	Service Catalogue Gets Data Sources from Research Product Catalogue	OA PROVIDE/REST
	→	EOSC Platform: Interoperability Framework Registry	Data Sources and Research Products Comply with IF Guidelines	EOSCIF REGISTRY/API

3.3.2 EOSC Catalogue and Marketplace – Provider View

As described in section 3.2.3, a provider organisation is a complex actor consisting of different systems. This section describes how the provider systems are connected to different components of the EOSC Catalogue and Marketplace. Figure 3.7 shows the providers systems and the EOSC Catalogue and Marketplace components involved, as well as the connections and the Standards/APIs to be used.

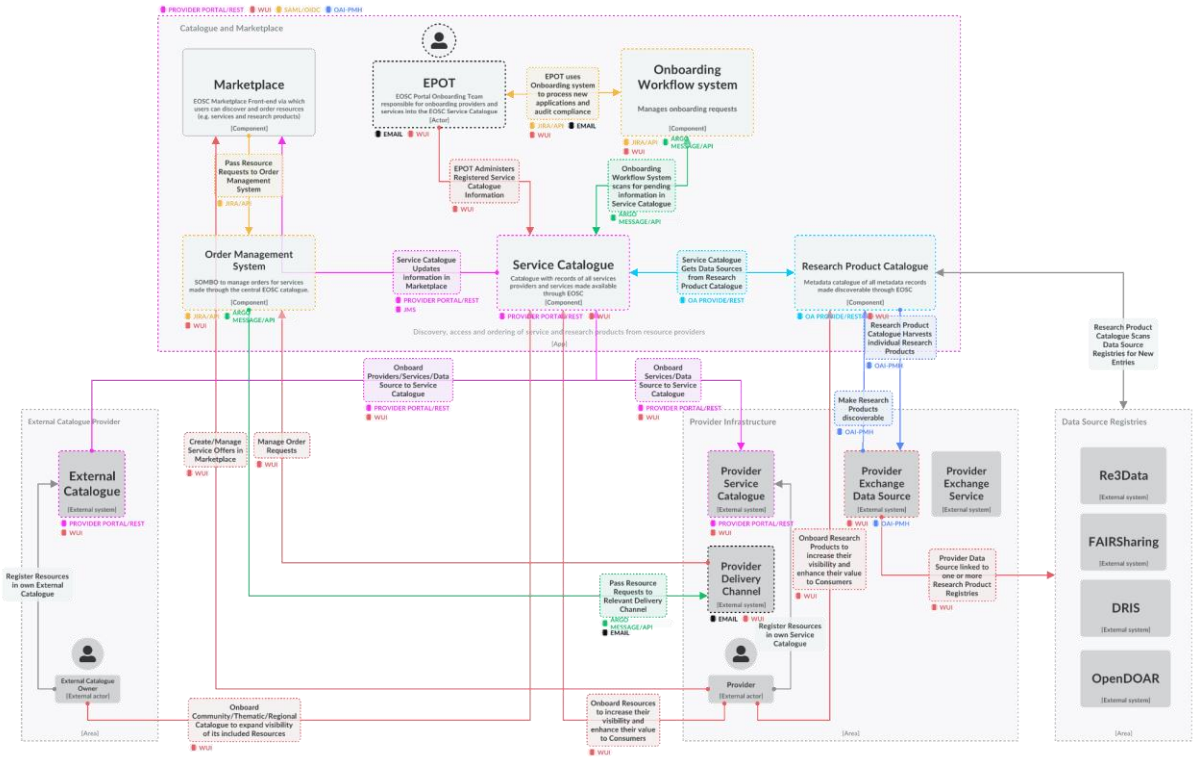


Figure 3.7: EOSC Catalogue and Marketplace component level diagram

Table 3-5 provides an overview of the components of the EOSC Catalogue and Marketplace that are involved in the interactions between the systems of an EOSC Provider, the direction in which the information flow is initiated, and the Interface/Protocol required for interaction.

Table 3-5: Overview of the Provider Systems connected to the EOSC Catalogue and Marketplace components

Provider Infrastructure Systems	Direction	Target Catalogue and Marketplace component and/or Provider Infrastructure service	Information flow	Interface/ Protocol
Provider (Actor)	→	Service Catalogue	Onboard Resources to increase their visibility and enhance their value to Consumers	WUI
	→	Provider Service Catalogue	Register Resources in own Service Catalogue	Provider local methods
	→	Research Product Catalogue	Onboard Research Products to increase their visibility and enhance their value to Consumers	WUI
	→	Marketplace	Create/Manage Service Offers in Marketplace	WUI
Provider Service Catalogue	→	Service Catalogue	Onboard Services/Data Source to Service Catalogue	WUI, PROVIDER PORTAL/REST
	←	Provider (Actor)	Register Resources in own Service Catalogue	Provider local methods
Provider Exchange Data Source	→	Data Source Registries (External)	Provider Data Source linked to one or more Research Product Registries	WUI
	→	Research Product Catalogue	Make Research Products discoverable	OAI-PMH
	←		Research Product Catalogue Harvests individual Research Products	
Provider Exchange Service	The onboarding of a provider exchange service is done via the Provider (actor) and Provider's own Service Catalogue.			
Provider Delivery Channel	→	Order Management System	Manage Order Requests	WUI
	←		Pass Resource Requests to Relevant	EMAIL, ARGO MESSAGE/API

Provider Infrastructure Systems	Direction	Target Catalogue and Marketplace component and/or Provider Infrastructure service	Information flow	Interface/ Protocol
			Delivery Channel	

Table 3-6 provides an overview of the components of the EOSC Catalogue and Marketplace that are involved in the interactions between the services of an External Catalogue Owner, the direction in which the information flow is initiated, and the Interface/Protocol required for interaction.

Table 3-6: Overview of the External Catalogue owner services to onboard resources into EOSC

External Catalogue Infrastructure	Direction	Target Catalogue and Marketplace component and/or Provider Infrastructure service	Information flow	Interface/ Protocol
External catalogue Owner (Actor)	→	Service Catalogue	Onboard Community/Thematic/Regional Catalogue to expand visibility of its included Resources	WUI
	→	External Catalogue	Register Resources in own External Catalogue	Provider local methods
External Catalogue	→	Service Catalogue	Onboard Providers/Services/Data Sources to Service Catalogue	WUI, PROVIDER PORTAL/REST
	←	External Catalogue Owner (Actor)	Register Resources in own External Catalogue	Provider local methods

3.4 Linking the EOSC Architecture to the Code Level

The fourth level of the C4 model is the Code level and aims to link the architecture components to code, using for example UML class diagrams, entity relationship diagrams or similar. According to the C4 model website [12] , the fourth level is seen as "an optional level of detail and is often available on-demand from tooling such as IDEs". In the Icepanel tool [13] , which is used to model the EOSC Architecture, the fourth layer of the C4 model is implemented through links from architecture model objects to external information sources, such as for example code repositories.

In the context of the EOSC Architecture modelling, the links option is used to link the architecture object models to the code repositories, the live service and the entry of the architecture object in the EOSC Core portfolio, see Figure 3.8.

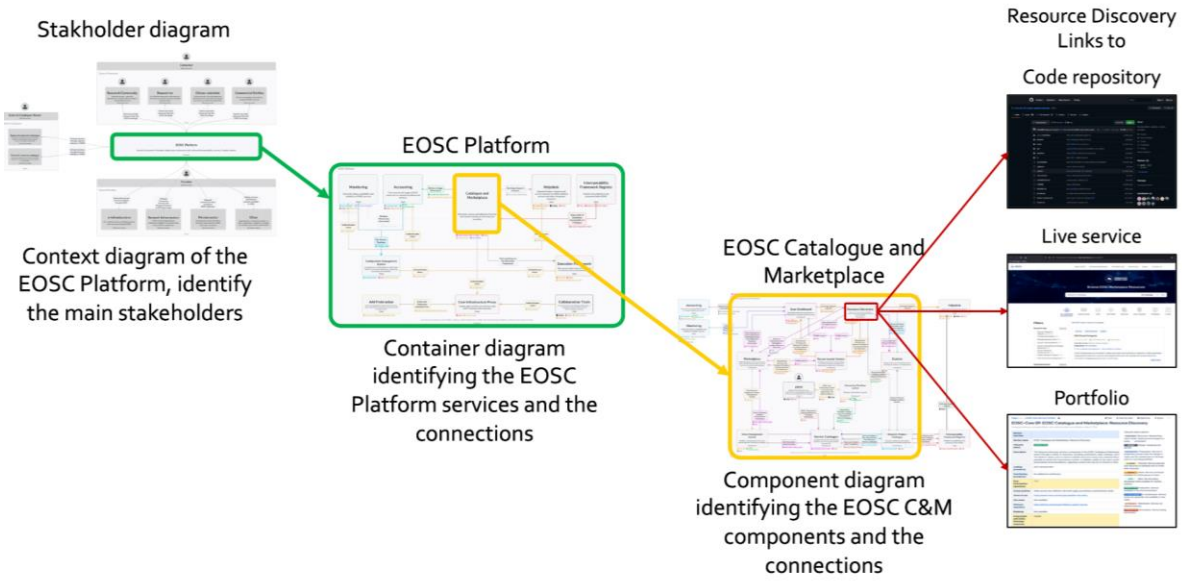


Figure 3.8: EOSC Architecture in the 4 layers of the C4 model

In Table 3-7, an example is provided showing how an architecture object is linked to external information sources.

Table 3-7: Example of architecture object links to external information sources

Architecture object	Reference name	Reference link
EOSC Catalogue and Marketplace: Resource Discovery	Code repository	https://github.com/cyfronet-fid/eosc-search-service
	Service URL	https://search.marketplace.eosc-portal.eu/search/all?q=*
	Portfolio	https://wiki.eoscfuture.eu/display/EOSCSMS/EOSC-Core+SP%3A+EOSC+Catalogue+and+Marketplace%3A+Resource+Discovery

3.5 Authenticating Users in EOSC

Interoperability at the level of the Authentication and Authorisation Infrastructures (AAI) between and across community services has been on the radar of research collaborations for several years. Thanks to the work done by the previous AARC projects, and the work carried out within the EOSC Architecture Task Force [14] (operated by the EOSC Association), the EOSC Future project could make significant progress in this area.

As the architecture building blocks for the AAIs were agreed and there was already significant experience gained within the research and e-infrastructures in operating AARC-compliant AAIs, work could focus on federating these AAIs and on establishing the EOSC AAI Federation as defined in the *EOSC AAI Architecture* report [15] . The purpose of the EOSC AAI Federation is to enable the AAI operated by the Research Infrastructures in the Science Clusters and the e-Infrastructures' AAIs to fully interoperate both with each other and with core services without the need for bilateral agreements and in line with the *AARC Blueprint Architecture and Guidelines* [16] .

More details concerning the technical architecture and achievements relating to the AAI for EOSC are reported in Deliverable 7.3 [17] *EOSC Federated Authorisation and Authentication Activities*. The harmonisation efforts for the AAI are carried out by groups external to the EOSC Future project that determine best practice.

In the context of the EOSC architecture, it is important for communities, RIs and e-infrastructures, to understand the communication flow between the different components involved in a federated infrastructure. The reference documentation for this consists of the EOSC AAI Architecture document and the *AARC Blueprint*

Architecture [18] , which is the basis of the EOSC AAI Architecture. A simplified diagram of a general authentication flow for an EOSC Exchange Service is provided below in Figure 3.9:

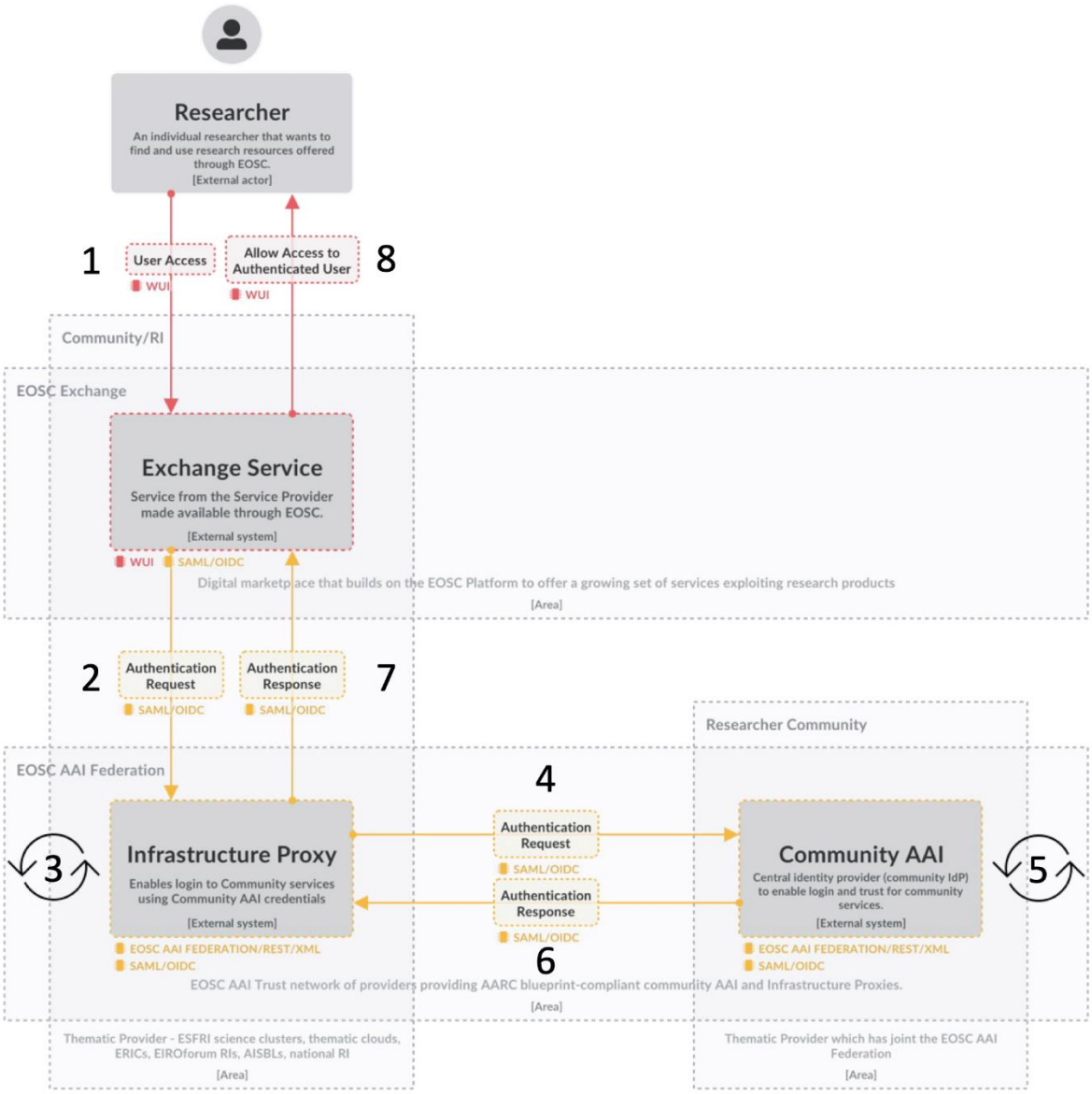


Figure 3.9: User access flow diagram of a Researcher accessing a community exchange service

Table 3-8 describes the authentication communication flow between the different components: from the Researcher accessing a service and needing to log in, to them selecting and logging in via their community AAI, to access being granted.

Table 3-8: Authentication communication flow between the different components

Step	Architecture object	Direction	Target architecture object	Description	Interface/ Protocol
1	Researcher (Actor)	→	Exchange Service	Researcher access a Community Service made available through the Exchange and wants to log in	WUI

2	Exchange Service	→ auto-redirect	Infrastructure Proxy	Community Service redirects the authentication request to the Infrastructure Proxy	SAML/ OIDC
3	Infrastructure Proxy	↻		The Researcher selects his/her own Identity Provider	WUI
4	Infrastructure Proxy	→ auto-redirect	Community AAI	The Researcher is redirected from the Community Infrastructure Proxy to his/her own Community AAI to authenticate his/herself	SAML/ OIDC
5	Community AAI	↻		The Researcher authenticates his/herself at his/her own AAI service	WUI
6	Community AAI	→ auto-redirect	Infrastructure Proxy	The AAI service of the Researcher sends the authentication response to the Community Infrastructure Proxy of the Service (Community Exchange Service) via Authentication Response	SAML/ OIDC
7	Infrastructure Proxy	→ auto-redirect	Exchange Service	The Community Infrastructure Proxy sends the authentication response to the requesting Service (Community Exchange Service) via Authentication Response	SAML/ OIDC
8	Exchange Service	→ auto-redirect	Researcher (Actor)	Depending on the authentication response the Researcher is authenticated and receives authenticated access to the Service (Community Exchange Service) via Allow Access to Authenticated User	WUI

4. Overview of the EOSC Interoperability Framework

The EOSC Interoperability Framework (EOSC IF) aspires to be a framework of policies, guidelines, standards, and best practices that enables different systems, applications, or technologies to communicate and exchange information with each other. Interoperability in this context refers to the ability of systems within the EOSC Platform and across scientific disciplines to work together, exchange data, and use each other's services, ideally in a seamless and efficient manner. The EOSC IF is designed to include documentation relating to technical specifications, protocols, regulations, standards and interfaces that facilitate the exchange of data and services between systems and implicitly includes all standards, formats, and guidelines used in science, provided they meet the inclusion criteria defined by the EOSC IF governance.

A high-priority goal of the EOSC IF is to promote the integration and interoperability of different systems, applications, and technologies, which is particularly important in complex environments involving diverse technologies and complex data sharing requirements, such as in the case of the European Open Science Cloud.

The key benefit of a functional Interoperability Framework for EOSC would be that it could facilitate the cross-discipline collaboration of researchers, providers, and research communities, who have been working on interoperability within and across their infrastructures for many years based on their own community best practices and standards.

The EOSC IF has been designed to comprise a wide range of components, which could be human readable (such as guidelines, domain standards, schemas, policy frameworks, or any other related community formats to facilitate interoperability with EOSC) or machine readable (e.g. APIs or configuration templates). The EOSC IF

consists of governance and services for the purpose of proposing, accepting, registering and promoting EOSC Interoperability Guidelines, which apply to the EOSC-Exchange and the EOSC-Core.

Deliverable 3.2a *EOSC Architecture and Interoperability Framework* introduced the concept of EOSC IF Governance as a structure committed to applying working principles observed across existing bodies that are understood to support fairness and transparency, but without attempting to introduce its own quality mark or certification model. That deliverable proposed a design that would encompass tools and procedures for the purposes of allowing a body (a community or an organisation) to propose guidelines for inclusion in the EOSC IF and of reviewing and ratifying such proposals, as well as for communicating to the scientific community the eventual inclusion of said guidelines and where to find them. During the course of 2022, subsequent to an initial community consultation, the EOSC IF Governance was implemented, and the relevant governing bodies were chartered. Important lessons were learned along the way, which are discussed in the following sections along with the results observed for each governing body.

4.1 The EOSC IF Governance Model

Deliverable 3.2a *EOSC Architecture and Interoperability Framework* proposed that a pragmatic, but robust approach should be implemented as the EOSC IF Governance, comprising the following aspects:

- An independent group, similar to an editorial board, that would assess requests for inclusion in the EOSC IF against specified Inclusion Criteria [19] , and would make recommendations for inclusion or exclusion of proposed guidelines.
- An overarching body that has oversight of and the responsibility for the EOSC IF as a whole.
- A body to own the EOSC-Core Interoperability Guidelines.

This model was implemented as described in detail in the published *Charter for the EOSC Interoperability Advisory Board and EOSC Interoperability Area Chairs* [20] . The proposed approach was then subjected to a community consultation [21] to allow the scientific and research community to ask questions and offer feedback on the concepts as designed.

The governing bodies' responsibilities are:

EOSC Interoperability Advisory Board (EIAB): the overarching goal of the EOSC Interoperability Advisory Board is to oversee the evolution of the EOSC IF; its outputs are based on the recommendations of the EIAC and result in endorsement of new guidelines but may also include recommendations for the deprecation of guidelines. Additional to the endorsing (or otherwise) of proposed guidelines, its oversight of the EOSC IF contributes to the practical and tangible aspects of establishing interoperability for EOSC services and resources. The EIAB's role is performed by the EOSC Future Technical Coordination Board.

EOSC Interoperability Area Chairs (EIAC): the overarching goal of the EOSC Interoperability Area Chairs (EIAC) is to perform the initial assessment of proposed guidelines, making recommendations for inclusion/exclusion to the EIAB. The role was designed o be performed by the EOSC Future WP3 task leads, calling in the assistance of experts where required. In practice this has proven difficult to implement because of the additional accountability this entails over and above delivery of the activities within the tasks and work package itself.

In view of this, the EIAC members have appointed deputies empowered to perform the guideline reviews, who operate in the guise of an **EOSC IF Editorial Board** that supports the EIAC in reviewing the guidelines and best practices proposed as EOSC Interoperability Guidelines against the acceptance criteria.

Following a pilot of the process that covered some initial ad hoc requests for volunteers from the EIAC panel to review proposed/draft interoperability guidelines, it was decided that a more formal procedure should be implemented, including a series of biweekly meetings to:

- Look ahead at the pipeline of forthcoming Interoperability Guideline candidates that are planned to be proposed for inclusion in the EOSC IF, having received prior notification from the responsible author/s.
- Assign the review of guidelines received for consideration to specified individuals and subject matter experts.

- Determine the outcome of the review of each proposed guideline and decide whether an endorsement should be made to the EIAB.

It was further determined that these meeting should be supported by a mechanism for tracking the guidelines in their various stages from draft to review. A *Guidance Note for Reviewing EOSC Interoperability Guidelines* [22] was published for the purposes of ensuring a consistent approach to the review of all proposed Guidelines, and for helping individuals asked to participate in the review process to understand what is required of them.

In addition to the governing bodies, Deliverable 3.2a also defined a series of governance processes, including:

- **A process for submitting, consulting, and accepting guidelines/other frameworks** has been agreed and published (refer to Appendix C). This procedure is incorporated into the EOSC Onboarding procedures and is overseen by the EOSC Onboarding Team (who provide support and guidance to the EIAC), while the EIAC participates in the EPOT for the purposes of managing the review of proposed EOSC Interoperability Guidelines. This procedure was initiated alongside the M25 release which provided the necessary capability to the EOSC Portal (refer to section **Error! Reference source not found.**), and requires Interoperability Guidelines proposed for inclusion in the EOSC to be reviewed according to a published set of inclusion criteria to onboard an interoperability guideline (refer to Appendix B).
- **A structured proposal template providing information about the EOSC Interoperability Guideline:** a template was initially proposed in Deliverable 3.2a, which was refined in response to a community consultation and the feedback received from a broad range of contributors. The resulting data model was accepted into the EOSC Profiles as the *EOSC Interoperability Guideline Profile* [23] and now forms the basis of the onboarding procedure, collecting the necessary (but coarse-grained) metadata at the point where an Interoperability Guideline is onboarded via the EOSC Providers Portal. The metadata relating to each onboarded Interoperability Guideline is stored in the EOSC IF Registry and reviewed by the EPOT. The resulting EOSC Interoperability Guideline metadata is made visible to users by way of the EOSC Marketplace and elements of the metadata make it possible for users to filter their search. The metadata profile was designed to be sufficiently lightweight in order to limit the complexity of onboarding Interoperability Guidelines to EOSC and will evolve in line with specific requirements and constructive feedback received from the scientific community gathered during the normal course of operations (refer to Appendix E).
- **A registry for accepted guidelines and frameworks**: Deliverable 3.2a proposed that a Registry should be introduced where artefacts will be curated using an agreed EOSC Profile extension for EOSC Interoperability Guidelines. Work has been carried out to implement such a registry; further details relating to its implementation can be found in section **Error! Reference source not found.**.

Initially, the focus has been on reformulating the EOSC-Core Interoperability Guidelines into public-facing documents (to operate as Core<>Exchange interoperability guidelines, rather than the comprehensive Core<>Core versions that had initially been published). In the time where the onboarding functionality was still being designed and developed for the Provider UI, it was necessary to allow the proposal and onboarding procedure to be instigated by way of an email to the eiac@eosc-future.eu inbox, which would be acted upon by the EIAC chair and, subsequent to EIAB approval, manual onboarding to the EOSC IF Registry by the EPOT. This manual process resulted in the first Interoperability Guideline, the *EOSC Security Operational Baseline* [24] , being loaded to the EOSC IF Registry database in March 2023.

Task 3.1 also created a set of templates [25] for use by other parties when proposing Interoperability Guidelines for inclusion in the EOSC IF:

EOSC-Core Interoperability Guideline Template: to be used when describing specifications for the purposes of interoperating with EOSC-Core services. The template informs readers that EOSC-Core guidelines should include context and a description in order to provide technical instructions to Providers wishing to integrate their services and/or resources (or be interoperable) with one or more EOSC-Core Services (EOSC-Core Infrastructure Proxy, Monitoring, Accounting, Helpdesk, Order Management, Metrics, service and resource registries, Provider Portal, Marketplace, PIDs for services, Knowledge Base registry), and is designed to be used by operators of EOSC-Core Services to provide guidance to Service and Resource Providers who wish to benefit from the functionalities that the EOSC-Core services offer.

EOSC-Exchange Interoperability Guideline Template: this template was developed to apply to two different contexts. The guidance informs readers that an **EOSC-Exchange Interoperability Guideline (thematic/community)** refers to the context where thematic or community services and resources interoperate with each other, usually at a domain (e.g. Cluster-specific) level. It is understood that this type of Interoperability Guideline is created by scientific community experts to:

- provide guidance to their communities,
- highlight the importance of their interoperability efforts at the Europe-wide EOSC level, and
- increase awareness of thematic and community accomplishments to facilitate/increase interoperability relating to specific domains.

An **EOSC-Exchange Interoperability Guideline (horizontal)** on the other hand refers to the context where services and resources interoperate with each other across communities and infrastructures. These are important at the EOSC Exchange level to help EOSC facilitate cross-linking of services and resources while staying domain agnostic. These are of wider applicability than the thematic/community guidelines and have a much wider scope than the EOSC-Core Interoperability Guidelines. It is understood that this type of Interoperability Guideline is created by scientific community experts to:

- provide guidance across communities,
- highlight their importance at the Europe-wide EOSC level,
- increase awareness of thematic and community accomplishments to facilitate/increase interoperability across communities and domains.

It should be noted that it is assumed horizontal/thematic/community Interoperability Guidelines are already established and published by an Infrastructure or Scientific Domain/Community who act as the owner or custodian of the Guidelines, and therefore it is unlikely that an author might intend to write a completely new Interoperability Guideline for submission to the EOSC Interoperability Framework. However, it is conceivable that an existing Interoperability Guideline or collection of Guidelines would benefit from being structured into an EOSC Interoperability Guideline, therefore the template was provided on that basis.

To further complement the support offered to users (with a view to optimising engagement with the EOSC IF), a *Guidance Note for Writing an EOSC Interoperability Guideline* [26] document was published.

Both templates are publicly available [25] .

In summary, in support of the governance and the implementation of the EOSC IF Registry, a range of documentation was published for use by the various users and stakeholders:

- Dedicated information area of the EOSC Portal relating specifically to the EOSC IF: https://eosc-portal.eu/eosc-interoperability-framework
- A set of guidance notes for use by guideline reviewers:
  - Guidance note: Writing a Guideline for the EOSC Interoperability Framework https://doi.org/10.5281/zenodo.7929869
  - Guidance note: Reviewing a Guideline for the EOSC Interoperability Framework https://doi.org/10.5281/zenodo.7929899
  - Guidance note: Onboarding an Interoperability Guideline to EOSC https://doi.org/10.5281/zenodo.7929833
  - Guidance note: Assigning an Interoperability Guideline to an EOSC Service Profile https://doi.org/10.5281/zenodo.7929855
- EPOT Procedure: Onboard an Interoperability Guideline: https://wiki.eoscfuture.eu/display/EOSCOB/EPOT+Procedure-14%3A+Onboard+an+Interoperability+Guideline
- EPOT Procedure: Audit an Interoperability Guideline: https://wiki.eoscfuture.eu/display/EOSCOB/EPOT+Procedure-15%3A+Audit+an+Interoperability+Guideline

4.2 EOSC Interoperability Framework: Interoperability Registry

The EOSC IF Registry is designed to create a supporting structure for the Governance of the EOSC IF, provide a tool for Providers to utilise when proposing Interoperability Guidelines for inclusion in the EOSC IF, and essentially establish a library of guidelines, specifications, policies, and so on. Its technical capability was demonstrated in the beta EOSC Portal in the M18 release, deployed to production in the M22 release (limited to EPOT Admin users), and extended to include a UI for Providers and within the Marketplace which provided the necessary capability to the EOSC Portal as part of the M25 release. As referenced in section 4.1 above, this online technical capability allows both onboarding and discovery, as follows:

Onboarding:

- Users (Provider Representatives) to onboard Interoperability Guidelines to the EOSC Interoperability Framework's Registry using the EOSC Provider Portal user interface. This implements the capability and process for proposing an Interoperability Guideline for inclusion in the EOSC IF.
- Users (Provider Representatives) to assign interoperability guidelines to their Services in order to assert compliance to those guidelines.
- Users (EOSC Onboarding Team Representatives) to review Providers' proposals to onboard new Interoperability Guidelines to the EOSC IF using the EOSC Provider Portal EPOT Admin user interface.

Discovery:

- The Interoperability Guidelines are discoverable via the Marketplace, via a dedicated search filter, which has been provided as a 'beta' service for the purposes of ensuring that users feel comfortable with giving feedback that will help to evolve the feature set to improve utility based on real-life scenarios and use cases.
- The Service Profile in the Marketplace has been updated to show the 'assigned guidelines'.
- Users (all) can view EOSC Interoperability Guideline profiles (and refer to the registered guideline documentation) within the EOSC Marketplace, which has been launched as a beta service in order to gain initial feedback from the user community regarding its design, navigation and content.
- The Interoperability Guideline Profile in the Marketplace will be updated to enable a user to see which Services support the Guideline (planned for summer 2023).

The Interoperability Guideline documents themselves are created outside the remit of Work Package 3 and Task 3.1, which had the task of designing and implementing the EOSC IF Governance. Interoperability Guidelines written in respect of achieving interoperability with the EOSC Core Services are created by the EOSC Core Service operators. The EOSC IF Registry, enabled by its related user interfaces, will allow Interoperability Guidelines as defined by the scientific community to be onboarded to the EOSC IF.

Appendix D and Appendix F provide more information regarding the features delivered so far.

4.3 Challenges encountered in implementing the EOSC IF Registry and the EOSC IF Governance

The most significant challenges encountered during the initiation of the EOSC IF and its governance relate to:

- A general expectation that the EOSC Interoperability Framework would be a fully comprehensive, tangible knowledge base of interoperability guidelines that would from 'day one' enable wide-scale, seamless technical interoperability across the whole scientific community.
- Similarly, balancing an expectation that the EOSC Interoperability Guideline Profile data model would be sufficiently lightweight so as not to impose an overhead on the Provider onboarding and maintaining it, but sufficiently so that it would meet all FAIR criteria and incorporate metadata of interoperability objects that would satisfy all users of all communities.
- Establishing a baseline of prerequisite knowledge that would ensure that Guideline documents would be on the one hand sufficiently comprehensive that a user could understand and implement their specifications without further complexity, no matter what part of the community they are from (i.e., whether an e-Infrastructure manager, a Provider of data services for researchers, or a researcher

looking to ensure interoperability of their research outputs with other communities), while being succinct enough to ensure that the user is not overwhelmed by referenced information.

- Sourcing an editorial board with both the breadth of knowledge and time capacity that is necessary to review the documentation in sufficient detail and with a quick turnaround; utilising WP3 Task Leads and their deputies has proven challenging, particularly because of a lack of time given their involvement in the implementation of project-specific activities, but also in terms of the potential for conflict of interest in the sense where there would be a risk of the same person or group of people being requested to review and/or approve a guideline that they have also written.
- Requests for 'one size fits all'-style templates which, given the heterogeneity of the community that the EOSC IF is intended to cater for, is an unrealistic target.
- A significant post-M25 challenge foreseen is that as the functionality rolls out to production, the Registry will initially appear to be 'empty', and motivating Guideline owners and custodians to register their Guidelines whilst Service Providers are not yet in the habit of assigning them creates a 'chicken or egg' dilemma, which will need careful management during the promotion of the new features.
- An enduring challenge is presented in the context where new infrastructure components are developed (with associated Interoperability Guidelines), their adoption may initially be low, and therefore the perception of their maturity may be lacking, which can result in low community uptake, which creates a vicious cycle of low adoption = low maturity = low adoption, and so on.

Furthermore, Deliverable 3.2a recommended that the initial scope of the Interoperability Registry should incorporate the 'Main Standards and Interfaces as a Starting Point for the EOSC', which captured a combination of services, protocols, standards and APIs that it would not make sense to include in the EOSC IF as standalone artefacts, i.e., it was determined that it is necessary to provide relevant context as to how such protocols, standards and APIs would be deployed in solving a given and specified interoperability problem. Some of the objects listed within the original table have been incorporated by in-progress Interoperability Guidelines or are considered 'candidates' to be the topic of further outreach and consultation activities with the wider community.

5. EOSC IF Supporting Cross-Discipline Interoperability

The EOSC IF offers the scientific community access to practical guidance and direction on how to improve the interoperability elements of their work (e.g. data/tools/services). This improved interoperability in turn can extend results and output for use beyond the boundaries of other scientific domains and enable reuse in further studies. For example, the COVID-19 pandemic has had a lasting and global impact in accelerating scientific research on the virus by several Research Infrastructures (RIs). Most notably, this work has not been constrained to the life sciences domain but has also been more widely the subject of research in the social sciences [27] , in the high energy particle physics domain [28] and beyond. The level of scientific collaboration that has taken place around COVID-19 is a prime example of what EOSC aims to achieve in terms of its stated goal of facilitating cross-domain scientific data and services usage. The EC has also funded the Beyond COVID (BY-COVID) project [29] , which also acknowledges COVID as an important cross-domain topic of relevance to interoperability, open science and the overarching mission of EOSC.

In order to use the COVID scenario as an example of the potential for interoperability-related learning and problem solving to be shared with a wider community, a COVID Metadata guideline is currently under development to show how the EOSC IF with its associated governance can provide a forum to promote guidelines deriving from a specific activity to the wider scientific community. The work of this project, as summarised in Appendix G, was undertaken as a cross-Cluster science use case project under EOSC Future WP6. This work involved direct interviews with RI experts on metadata usage in order to produce a framework for contextual metadata surrounding COVID, especially within research in the domains represented by the participating RIs. This work included six RIs – BBMRI, EU-OPENSCREEN, CESSDA, CLARIN, ECRIN, and EATRIS - the majority of these operating in the social sciences or life sciences fields. This work will be written up and developed into a more formalised EOSC Interoperability Guideline to help facilitate interoperability around COVID research. An added advantage of this project is that it will produce recommendations that are not associated with a specific software component or service, and are the result of broad community representation and discussion in the project.

The following key elements will be addressed in the EOSC IF Guideline from META-COVID:

- The planned EOSC Interoperability Guideline on COVID metadata aims to provide a comprehensive approach to leveraging contextual metadata. It will address the confusion surrounding the definition of "contextual metadata" by comparing it to other types of metadata, particularly provenance. The guideline will clearly define contextual metadata by splitting it into two components: "direct output/data provenance" and "contextual research activity."
- Essential elements of contextual metadata, including research project, research activity, research programme, and research process, will be precisely defined in the guideline. The Guideline will establish the relationships between these elements and explain how they are interconnected.
- To facilitate practical implementation, the Guideline will build upon existing metadata schemas in various domains and RIs. No major adaptations or extensions will be necessary to begin the process. It will outline how contextual metadata, whether explicitly or implicitly available, can be assessed and utilised to support interoperability. Technical prerequisites such as APIs, knowledge graphs, and NLP will be described.
- The Guideline will feature use cases that demonstrate the benefits of incorporating contextual metadata. It will also provide guidance on improving the documentation and management of contextual metadata in the future, particularly if existing metadata schemas used in RIs are adapted or updated.
- Overall, the guideline will serve as a valuable resource for researchers across diverse scientific domains, offering arguments, examples, and practical advice on harnessing the power of contextual metadata in the context of COVID.

6. Towards Machine-Readable Interoperability

The EOSC Interoperability Framework aims to provide a common framework to describe the compatibility capabilities of EOSC resources across research infrastructures' borders. By introducing the ability of assigning resources to EOSC Interoperability Guidelines, the EOSC Exchange is endowed with an interoperability overlay that enables the discovery and composition of resources based on their ICT interoperability features. EOSC Interoperability Guidelines are sets of instructions for EOSC Providers to follow (aka implement) to enable given functionality/interoperability between EOSC services and/or EOSC research products; for example, combinations of metadata profiles/schemas, file formats, exchange formats, API signatures, and protocols, together with instructions on how these should be adopted for interoperability.

The degree of composability and automation enabled (e.g. service orchestration) depends on the description detail of the resources. Given the current typologies of resources supported by the EOSC, namely services and products, it is envisaged that two composability patterns could apply feasible:

Composability of products and services: indicating how a service can be used to process a dataset. Examples are:

- A service used by a researcher to process a given product, e.g. a researcher discovers interesting Twitter Data and finds EOSC Services capable of processing Twitter Data.
- The service is invoked by another service to process a product (that is, it performs remote invocation of datasets); for example, an orchestrator service allows researchers to discover repository services that can be used to deposit a dataset of interest; in this pattern, repository services are characterised by their adherence to common guidelines for deposition as defined in the (not yet published) *Interoperability Guideline: Research Product Deposition*.

Composability of services and services: indicating how a service can be combined with or reused by another service. Examples are:

- Services are combined or used in the same experimental settings of a researcher: For example, researchers interested in text processing in Cultural Heritage search and find services that are able to process products compatible with EOSC Textual Files guidelines related to that discipline in the EOSC resource catalogue;

- The services are automatically pipelined and executed as steps of a workflow by a workflow engine (e.g., CWL, Galaxy, Taverna, etc.).

The EOSC IF registry enables EOSC Services metadata to refer to unique identifiers and descriptions of Interoperability Guidelines. In doing so, such references could enhance the expression of the compliance of a resource to an Interoperability Guideline as currently implemented, and therefore enable EOSC IF-driven discovery and, potentially, forms of resource composability, whether manual (developers to perform the job) or automated (orchestrators performs the necessary assembling based on metadata information). Still, the achievement of an ideal environment of "transparently composable resources across RIs and clusters" will not be achievable in the short term, given the heterogeneous and consolidated practices (where they exist) of the Research Infrastructures and Research Clusters. In this scenario, the top-down imposition of composability solutions is not viable. It would disrupt current practices, imply the dismissal of technologies, and require new major investments, in many cases hardly accountable to RIs and clusters. Accordingly, the EOSC IF aims at an incremental achievement of the long-term cross-disciplinary vision described above. Ensuring resources in the EOSC Exchange can describe/specify their interoperability capabilities, will enable a context where EOSC providers can choose to invest into the implementation of EOSC Interoperability Guidelines based on conscious choices driven by opportunities rather than imposition.

6.1 EOSC IF Registry v1.0

The EOSC IF registry v1.0 creates a "thin" semantic layer on top of the EOSC resource catalogue that allows viewing of EOSC resources in terms of compliance to IF guidelines. At May 2023, the Registry allows the creation (and viewing) of a simple semantic relationship between an EOSC Interoperability Guideline and a Service within the EOSC Catalogue. For example, using this method, it is possible to express a Service as 'isCompatibleWith' EOSC Core Interoperability Guidelines (e.g. EOSC Security Operational Baseline, the EOSC Core Help Desk, etc.) or EOSC Exchange Interoperability Guidelines (e.g. EOSC repository publishing guidelines developed in WP3, EOSC Galaxy Workflow). As such, the registry allows the "annotation" of EOSC resources across research/e-infrastructures' borders with the guidelines they comply with. This mechanism enables an interoperability-driven overlay of EOSC resources across different disciplines and the discovery of resources based on interoperability features.

Initially, the EOSC IF Registry will relate to a set of guidelines intended as human-readable instructions, together with pointers to specifications, that EOSC Providers need to implement to enable given functionality/interoperability between their EOSC services and/or EOSC research products. These guidelines may refer to multiple existing standards, protocols and guidelines etc. (e.g. FAIRsharing DOIs), in such a way that: (i) Providers can understand how to design and configure services to interoperate with a wider community of users; (ii) Users can understand any nuanced requisites needed to implement code to interoperate with such services.

6.2 EOSC IF Registry v2.0

As part of the EOSC Future effort, a proof of concept (TRL5/6) version 2.0 of the registry will be delivered by the end of the project to showcase how the EOSC IF registry can be extended to support forms of EOSC resource orchestration and machine-composability. To this aim, it will handle EOSC Interoperability *configurations*, structured metadata profiles through which providers can describe the actual access parameters of their services for a given IF guideline. The owner of the IF guideline can define a *configuration template*, which is a structured profile (i.e. metadata schema) that defines properties and vocabularies of access parameters. Service providers can, when they declare compatibility with a given Interoperability Guideline via the EOSC Service Provider Portal, also insert the actual parameters required by the configuration template, referred to as *configuration instances (*aka configuration). Such an extension showcases (i) discovery and composability of resources based on the Interoperability Guidelines they are compliant with (e.g. APIs, metadata standards, protocols); and (ii) service-oriented access to services exploiting their configurations (i.e. IF Guidelines access parameters). For example, if the EOSC OAI-OMH protocol is registered as an EOSC Exchange Interoperability Guideline in the EOSC IF registry, the owner (manager) can upload the configuration template that describes the properties to be provided by services complying to OAI-PMH in order to be machine reusable; it may consist of one property "OAI-PMH URL entry point/API", from which all protocol verbs can be invoked. The EOSC Provider of a repository service that is regularly registered to the EOSC Catalogue as a data source may:

1. Specify that the service supports OAI-PMH by adding a reference to the related IF guideline entry in the EOSC IF registry.
2. Fill in the configuration instance for OAI-PMH.

The repository configuration enables other services to discover the repository via the EOSC Catalogue and dynamically access its APIs.

7. Conclusions and Next Steps

7.1 EOSC Architecture

As the EOSC Future project has progressed, the EOSC Architecture has evolved from a high-level architecture of the EOSC MVE and EOSC Platform to a detailed architecture based on the C4 abstraction-first model, consisting of:

- Context-level diagrams describing the EOSC Platform as a single system and identifying its stakeholders.
- Container-level diagrams identifying the EOSC Platform containers and the interconnections between them and indicating the Interfaces/Protocols used for the connections.
- A component-level diagram of the EOSC Catalogue and Marketplace, identifying the components and the interconnections between them, and the Interfaces/Protocols used for the connections.

The architecture diagrams were developed using Icepanel, a software architecture modelling tool in which the elements are defined as objects within the model. This makes it possible to easily use the defined objects at different levels of the C4 model and across different diagrams. the architecture containers and components are linked to code repositories, live services and to the EOSC Portfolio via the link function in Icepanel.

Different diagram flows have been developed to explain the interactions between the various components within the architecture, for example a user access flow to explain the log in sequence initiated by a Researcher, as described in section 3.5. Diagram flows have also been developed to explain the onboarding procedure, as well as for the Helpdesk and the Data Transfer Service.

While use of Icepanel has provided many advantages in developing the EOSC Architecture, it has also presented some challenges. Because EOSC is built through several different projects, for example with the contribution of the FAIRCORE4EOSC project, which develops platform components supporting FAIR in EOSC, the architecture diagrams need to be shared with a broad community. As Icepanel is a commercial product with a user-based licensing model, the current license for the use of the tool is limited to members of the EOSC Future project. To ensure continued access to Icepanel and the architecture model, discussions will be initiated with stakeholders for a potential handover of the architecture model to the EOSC Association.

In the remaining months of the project, EOSC Future will develop an *EOSC Architecture Reference* with the aim of providing a complete overview of the EOSC Architecture and consolidating information which is currently disseminated across different deliverables into a single document. The *EOSC Architecture Reference* document will be aimed at external stakeholders who want to learn about EOSC from an architecture context. It will provide valuable input to support transition to the HORIZON-INFRA-2023-EOSC-01-05 EOSC Architecture and Interoperability Framework[2] call.

[2] https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/horizon-infra-2023-eosc-01-05

7.2 EOSC IF and the EOSC IF Governance Model

Since submitting its initial proposal, WP3 has undertaken significant work to implement the EOSC IF and the associated governance model, including establishing the EIAC and EIAB, supported by dedicated tools and regular meetings. The challenges explained in section 4 feed into the continuing improvement of the procedures and documentation.

An EOSC IF Registry has been implemented In collaboration with the technical work packages – WP4 and WP5 – which enables Providers to onboard proposed Interoperability Guidelines for consideration by the EPOT and EIAB, for potential inclusion in the registry. A UI has been developed to allow Interoperability Guidelines to be discovered in the EOSC Marketplace, based on an initial coarse-grained metadata model (the EOSC Interoperability Guideline Profile) that has been incorporated into the EOSC Profiles. The initial focus has been on preparing the EOSC-Core Interoperability Guidelines for inclusion in the EOSC IF Registry, and the new features will be communicated to the scientific community, with a view to incorporating thematic and horizontal Interoperability Guidelines within the registry.

A set of documentation has been published to provide guidance to the relevant stakeholders and users.

The EOSC IF Registry and EOSC IF Governance as deployed must be evolved in response to specific feedback and requirements obtained from members of the community that these functions are designed to benefit. It should be recognised that the EOSC IF is a pathfinding activity that seeks to bring new and innovative capabilities to the scientific community at large, and specific use cases should be analysed relating to the various stakeholder groups, such as:

- Service Providers
- Data Catalogue Managers
- Resource Providers
- Resource Users

As such, the Interoperability Framework and related Marketplace UI for Interoperability Guidelines will be promoted with a view to 'normalising' the concept of proposing Interoperability Guidelines for inclusion in the EOSC IF, but these aspects should be iterated alongside user feedback and specified use cases in order to improve their utility.

Effort within this evolution and iteration process should be assigned to the EOSC Interoperability Guideline Profile, which should be tested with the stakeholder community, and extended to include relevant and useful EOSC Interoperability Resource Types, in order to make it possible to specify whether a Guideline is primarily an API, an executable configuration template, a protocol, a policy, etc. This will require updates to the Provider Portal and the Marketplace UI.

Additional to iteration and normalising the use of the EOSC IF and 'populating' the Registry:

1. An activity should be initiated to identify interoperability references from research product resources: as we cannot expect to have (in the short term) Data Sources exposing metadata of research products that include proper links to EOSC Interoperability Guidelines, it may be possible to proactively infer interoperability capabilities (metadata formats, exchange formats, file formats) from the metadata of dataset/software products using components already present within the EOSC Architecture and by use of tags;
2. Similarly, it may be possible to identify interoperability references from service resources: for the IF (if any) inferred above, identify which services in the Service registry can be associated with such formats (e.g. the service accepts data input of that format) and explore the utilisation of using tags that are periodically promoted by the EOSC-IF representatives to signpost Interoperability-related materials and resources.
3. Add Interoperability Guideline references to data resources: on both sides (Graph and Catalogue) it should be possible to (manually/automatically) include references to the identified Interoperability Guidelines from the resources; moreover, both service catalogue and Graph APIs should be adapted to enable the discovery of resources based on (among other criteria) possible Interoperability Guidelines, if technically feasible.

4. Based on use case-driven feedback from the proof of concept, iterate v2.0 of the EOSC IF Registry into a feature that is capable of performing machine-initiated and machine-enabled interoperability and service composability.

The Interoperability Guideline review procedure has been identified as suboptimal, as referenced in section 4 above; improvements to the procedures will be made in the remainder of the project and it will be necessary to determine the correct form of the end-state Governance, which is a topic of the EOSC-INFRA2023-01-05 call.

The relevant parts of the Architecture and Interoperability Framework will be subject to a structured transition to either the EOSC Association, the EOSC operational partner or the EOSC-INFRA2023-01-05 (EOSC Architecture and Interoperability Framework) and EOSC-INFRA2023-01-04 (Next Generation Services For Operational And Sustainable EOSC Core Infrastructure) consortia, although to which of these is yet to be specified.

8. Appendix A – Architecture modelling

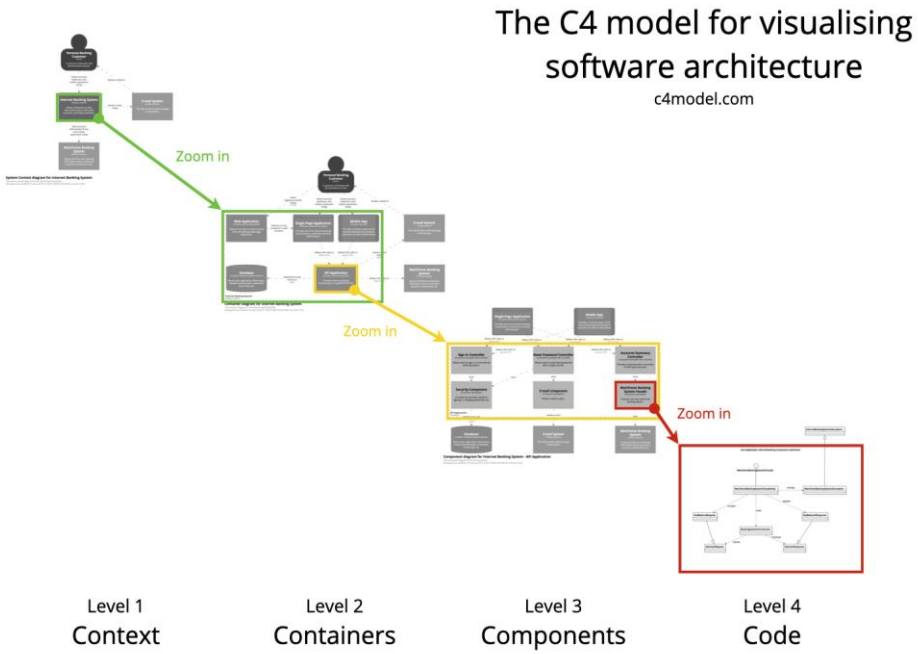
8.1 C4 Model

The C4 model is a lean graphical notation technique for modelling the architecture of software systems. The C4 model was created by the software architect Simon Brown between 2006 and 2011 on the roots of Unified Modelling Language (UML) and the 4+1 architectural view model. The aim of the C4 model is to define readable architecture diagrams.

8.2 C4 abstraction-first approach

The C4 model is an "abstraction-first" approach to diagramming software architecture, based upon abstractions that reflect how software architects and developers think about and build software. The small set of abstractions and diagram types makes the C4 model easy to learn and use. Please note that you don't need to use all 4 levels of diagram; only those that add value - the System Context and Container diagrams are sufficient for many software development teams.

8.2.1 C4 abstraction-first approach to software architecture diagramming



8.2.2 Short definition of the C4 abstraction levels

Abstraction level	Description
System context	A System Context diagram is a good starting point for diagramming and documenting a software system, allowing you to step back and see the big picture. Draw a diagram showing your system as a box in the centre, surrounded by its users and the other systems that it interacts with.
Container	The Container diagram shows the high-level shape of the software architecture and how responsibilities are distributed across it. It also shows the major technology choices and how the containers communicate with one another. It's a simple, high-level technology focussed diagram that is useful for software developers and support/operations staff alike.

Component	The Component diagram shows how a container is made up of a number of "components", what each of those components are, their responsibilities and the technology/implementation details.
Code	This is an optional level of detail and is often available on-demand from tooling such as IDEs. Ideally, this diagram would be automatically generated using tooling (e.g. an IDE or UML modelling tool).

8.3 Icepanel collaborative diagramming tool

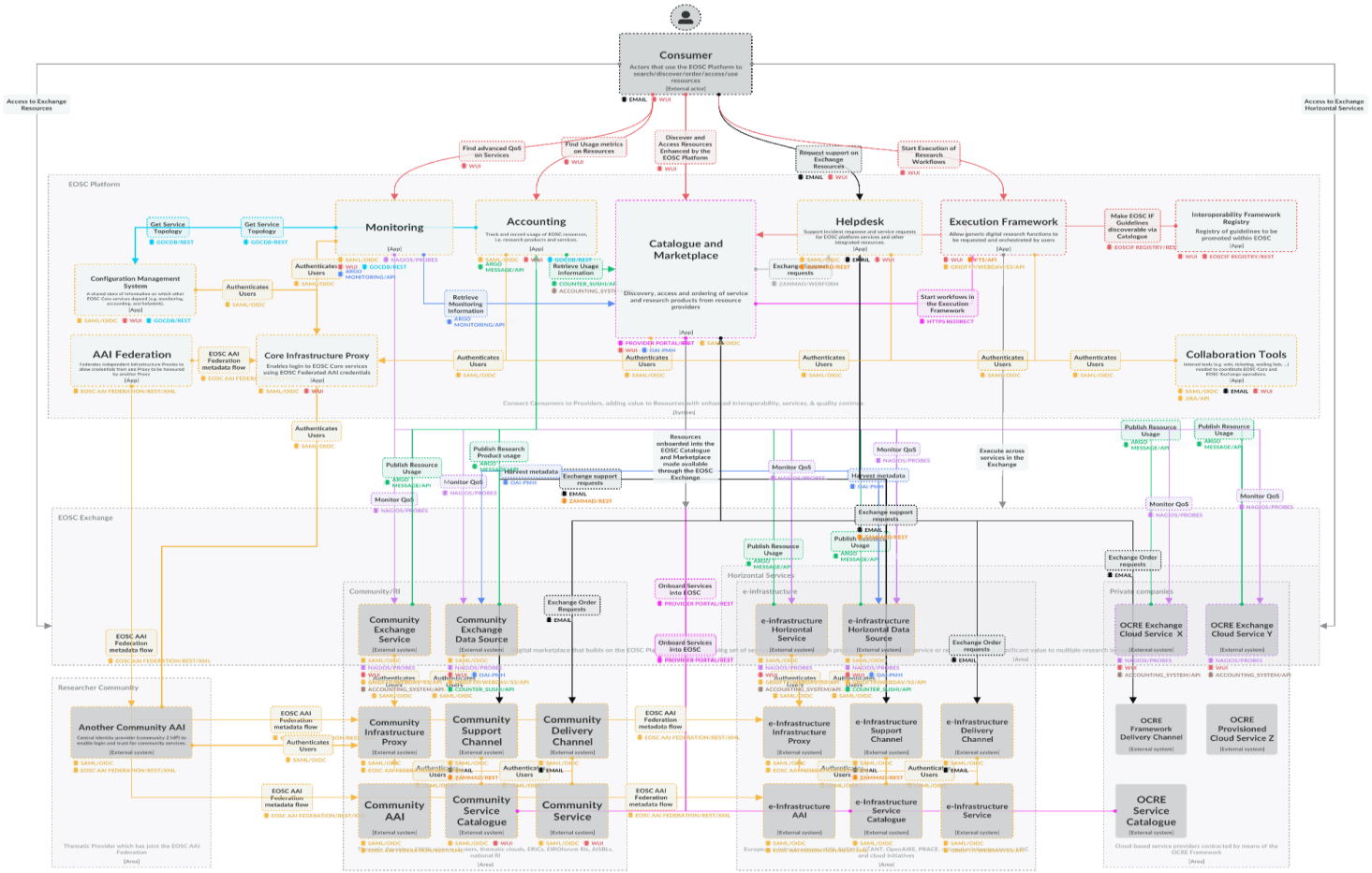
To develop the EOSC Architecture in the C4 model, EOSC Future used the online Icepanel tool. IcePanel is a collaborative diagramming tool that helps software engineering and product teams align on technical decisions. It is an software architecture modelling tool in which the elements are defined as objects within the model. In this way you can easily use defined objects at different levels of the C4 model and across different diagrams. Via the link function in Icepanel the architecture containers and components are linked to code repositories, live service and to the EOSC Portfolio.

Icepanel also offers the capability to define diagram flows to explain interactions between different components within the architecture. Within EOSC Future this capability has been used to develop diagram flows to explain an user access flow, the log in sequence initiated by a Researcher, the onboarding procedure to onboarding services, data sources and catalogues into EOSC and diagram flows to explain usage of the Helpdesk and the Data Transfer Service.

9. Appendix B – Detailed EOSC Architecture Diagrams

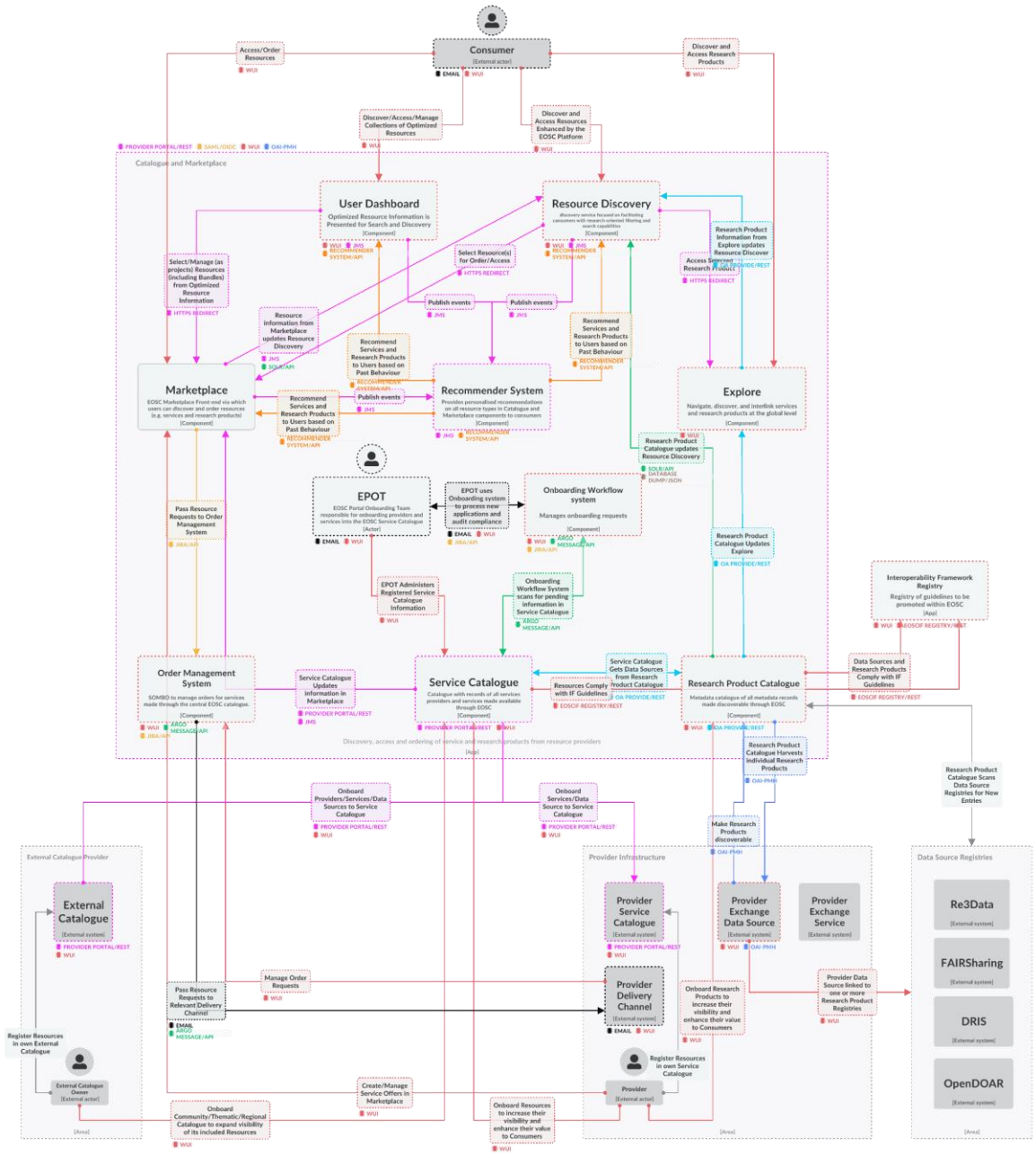
9.1 Diagram 1: EOSC Platform Architecture overall architecture diagram

EOSC Platform overall container-level architecture diagram showing the EOSC Platform, the EOSC Exchange, Providers and the interconnections between the different architecture objects.



9.2 Diagram 2: EOSC Catalogue and Marketplace architecture diagram

EOSC Catalogue and Marketplace component-level architecture diagram showing the EOSC Catalogue and Marketplace with the interconnections between the different architecture objects.



10. Appendix C – Additional Inclusion Criteria to Onboard an Interoperability Guideline to the EOSC Portal

The inclusion criteria are used to validate that providers and resources seeking to join EOSC-Exchange meet and commit to a set of requirements that flow from the EOSC Rules of Participation. Inclusion criteria include is applicable to the procedures of onboarding to the EOSC Portal as a new Provider, and for a Provider to onboard either a service/resource, a community resource catalogue, a research product or a learning resource, and all describe the nuanced requirements that a Provider must meet in order to qualify.

These inclusion criteria were extended to include criteria for onboarding an interoperability guideline. This criteria will evolve as the EOSC-IF matures. The published inclusion criteria[3] appears in section 3.6 of the documentation, is captured below and will be published to the EOSC Portal's Provider Hub[4] to coincide with the release of functionality for the EOSC-IF Registry to support the M25 release:

10.1 Additional inclusion criteria to onboard an interoperability guideline

In addition to meeting the common inclusion criteria to onboard all resource types, these criteria must be met to onboard an interoperability guideline:

- The guidelines documentation and profile (i.e. the metadata defined by the provider) is actively maintained.
- Use a well-supported, commonly used and open public repository to host or publish the guidelines that is capable of version control for future versions of the documentation or software, and that will assign a PID.
- Ensure that the described technical interoperability has been demonstrated in a relevant environment.
- Ensure the guideline is mature, meaning that it is version 1+, or higher, that it is actively maintained and/or that it has evidenced uptake. This refers to at least part of the available software (necessary for) implementing the guideline and that it includes adequate support documentation.
- The guideline must comprise the minimum required information as specified in the templates found at https://eosc-portal.eu/eosc-interoperability-framework.
- The guideline must have a documented and methodical approach to the description of how to implement its recommendations and related software.
- URLs cited in the guideline documentation or the guideline's EOSC-hosted metadata should have a Fully Qualified Domain Name (FQDN).
- Also and specifically for EOSC-Exchange Interoperability Guidelines (thematic and horizontal interoperability guidelines as defined in the EOSC Glossary):
  - The Provider must be able to show evidence of prior utilisation of the guideline in the given communities if asked to do so by the evaluation team. This could be evidenced by descriptions of uptake and/or maturity.
  - The Provider must be able to show evidence of prior consultation with the relevant community/ies if asked to do so by the EPOT.

[3] https://wiki.eoscfuture.eu/display/EOSCSMS/EOSC-Exchange+inclusion+criteria
[4] https://eosc-portal.eu/eosc-providers-hub/how-become-eosc-provider/instructions-onboard-providers-and-resources-eosc

11. Appendix D – Procedures for Inclusion in Interoperability Registry

11.1 EPOT Procedure-14: Onboard an Interoperability Guideline

11.1.1 Trigger

- The EIAC Chair (member of EPOT) checks the EPOT's EOSCOB Kanban board or is notified by email of a new ticket that has automatically been created in the EPOT JIRA project at the point that a Provider submits a new Guideline Profile to be onboarded.
- The EIAC Chair (member of EPOT) finds a new email in the onboarding@eosc-portal.eu inbox that provides notifications on new resource registrations.
- The EIAC Chair (member of EPOT) finds a new Interoperability Guideline with the state **Pending/InActive** in the Provider Portal.
- Identified during regular (weekly) manual keyword searches and filters based on the 'Interoperability Guideline Catalog' dropdown filter in the Catalogue or during audit procedures.

11.1.2 Step-by-step instructions

Step	Provider actions	EPOT actions	Action
0	Identify candidate guideline		The Provider establishes that it has identified a candidate Interoperability Guideline for inclusion in the EOSC Interoperability Framework. The guideline must meet the Inclusion Criteria for EOSC Interoperability Guidelines.
1	The Provider Representative (PR) selects the method to onboard the Guideline		The PR logs on to the Portal and proceeds with onboarding the Guideline. Note: if the Provider has not already been registered at the EOSC Portal, then the PR must follow the Onboard a Provider procedure. The PR is then offered two options to onboard the Guideline: a) via a web interface for each Guideline individually or b) via the Portal Application Programming Interface (API). If the web interface is selected then Step 1a follows, otherwise Step 1b. Note: The first Guideline of the Provider must be entered using the web interface, see Step 1a.

Step	Provider actions	EPOT actions	Action
1a	The PR applies to onboard a Guideline via the web interface		The PR applies for the onboarding of an Interoperability Guideline by completing the Interoperability Guideline Profile. Automated content validation is used to the greatest extent possible to ensure that all required information is included and that the information is of the correct type, size, etc. In case any difficulties arise during the application, the PR may communicate issues (1) and, depending on the nature of the issue, the EPOT will provide support. If the form is completed and passes all automatic validation, the Provider can submit it. This form shall be parsed by: • The Provider Dashboard in order to present an editable record of the Guideline's related metadata to Provider Admins, and • The EOSC Resource Catalogue in order to present a read-only view of the Guideline's related metadata to Researchers and to other Providers that may wish to utilise the Guideline. When the form is submitted the PR is prompted to confirm whether additional Guidelines will be onboarded. If yes, then Step 1a is repeated otherwise the process moves to Step 2. (1) onboarding@eosc-portal.eu
1b	The PR applies to onboard a Guideline via the API		The PR may apply for the onboarding of Interoperability Guidelines by using the Portal Open API. In brief, the Provider needs to follow the process to retrieve a new API token. Then, the Provider prepares the Guideline Profile by mapping the attributes to its own systems, according to the Interoperability Guideline Profile (3) and by calling the API's POST/Resource/validate method. Upon successful validation, the Provider calls the POST/Resource method to add the new Guidelines into the catalogue. Upon success, the Provider receives a new set of Resource IDs and the new Resources are onboarded to the Portal. A detailed description of the Portal Open API (2) is available. In case any difficulties arise during the employment of the API, the Provider may communicate issues (4) to the EPOT, who will support them. (2) https://providers.eosc-portal.eu/openapi (3) description of the Guideline according to the EOSC Interoperability Guideline data model. (4) onboarding@eosc-portal.eu

Step	Provider actions	EPOT actions	Action
2		The EPOT receives and triages the newly submitted Guideline profile	All Interoperability Guidelines submitted are to be triaged by the EOSC Portal Onboarding Team (EPOT) to the EOSC Interoperability Area Chairs (EIAC) in Jira. **Triggers** One of the following: • The EIAC Chair (member of EPOT) checks the EPOT's EOSCOB Kanban board or is notified by email of a new ticket that has automatically been created in the EPOT JIRA project at the point that a Provider submits a new Guideline Profile to be onboarded. • The EIAC Chair (member of EPOT) finds a new email in the onboarding@eosc-portal.eu inbox that provides notifications on new resource registrations • The EIAC Chair (member of EPOT) finds a new Interoperability Guideline with the state **Pending/InActive** in the Provider Portal. • Identified during regular (weekly) manual keyword searches and filters based on the 'Interoperability Guideline Catalog' dropdown filter in the Catalogue or during audit procedures.
3		The EIAC Chair oversees the review of the newly onboarded Guidelines	The EIAC onboarding representative examines the quality of the Guideline documentation and related Interoperability Guideline Profile referring to the Interoperability Guidelines Inclusion Criteria and the general recommendations and guidance that relates to the Interoperability Guideline Profile. This may include spelling, accuracy, composition and URLs, but the EIAC will not perform interoperability tests as a matter of routine. If the Guideline description does not comply with the inclusion criteria and the classification of the Interoperability Guideline's Profile, the PR may be asked to take an action (e.g. amend the description and to resubmit, etc.) or to join an information/training session. In this, often 1-to-1 call, the PR will have the chance to ask questions and get personalised consultation on the best way to onboard the Guidelines. To facilitate any required amendment, the EIAC provides recommendations taking into account potential Provider's specific contexts as understood by the team. In the case where a PR has submitted a thematic or horizontal EOSC-Exchange Interoperability Guideline for inclusion in the EOSC-IF Registry, the EIAC may request further information relating to evidence of prior validation and testing having been successful, and/or evidence of demonstrable impact, uptake and maturity of the Guideline. The Guideline may be rejected by the EIAC if: • the Guideline documentation or its Profile is not appropriate for the EOSC, • the Provider fails to respond to or act upon the request to amend the Guideline's Profile or documentation, which may also ultimately result in the Provider's suspension.

Step	Provider actions	EPOT actions	Action
			Otherwise, the EIAC notifies the PR of the approval of the Guideline's inclusion in the EOSC-IF Registry and the Guideline is published at the Portal. If the Provider description was also submitted alongside the Guideline proposal, the Provider description will be reviewed by the EPOT separately to the EIAC and publication of the first Guideline is dependent on the Provider Profile having been approved first. The EPOT or EIAC may also perform small corrections on the Guideline's Profile but will ask for the consent of the PR before publishing.
			• Action 1 - The EIAC representative assesses the Guideline Profile – The EIAC representative performs its review (as described above). – The EIAC representative member applies the inclusion criteria at https://www.eosc-portal.eu/providers-documentation/eosc-provider-portal-inclusion-criteria in particular: ○ Assess whether the submission meets the inclusion criteria ○ Check that the links provided point to the required content ○ Check that all fields are correctly filled in
			• Action 2 - If all is correct, the EIAC representative will consult with the EOSC Interoperability Area Chairs group and the EOSC Interoperability Advisory Board (EIAB) to determine whether the Guideline can be accepted. This is an offline process carried out without the assistance of EPOT, Jira or the Provider Portal.
			• Action 3 - If EIAC/EIAB approve the Guideline, the EIAC representative will move the Jira ticket to status "Record Approved" and click 'Approve' in the Provider Portal. – The PR will automatically be notified of the approval. – A notification is sent to onboarding@eosc-portal.eu
			• Action 4 - If the Guideline/Profile is not appropriate and/or the EIAC/EIAB judges it cannot be made so, the EIAC representative will click "Rejected" in the Provider Portal and record it as such in Jira. – Reject application in the provider portal – Send an email to the PR (and resource if different) email contact (with cc: onboarding@eosc-portal.eu) explaining why the submission is rejected – <To Add: TEXT FOR REJECTION> – Add note on reasons for rejection to Jira ticket comments

Step	Provider actions	EPOT actions	Action
			• Action 5 - If the metadata is not correct but the EIAC representative judges it can be corrected, move to Jira status "Application Requires Review" – Request that the PR makes the required adjustments in the Provider Portal by sending an email to the resource email contact – CC onboarding@eosc-portal.eu – Add a comment to the Jira ticket with the reason the application change was requested
			• Action 6 - Follow up – Monitor the progress of the ticket by referring to replies to email or through automatic notifications from the portal when the Guideline Profile is updated. – Respond to enquiries and updates from the PR. – For any change in status or new communication, add details to Jira ticket as a comment – Depending on the circumstances, move the ticket to 'Record Approved' or 'Rejected' status and proceed as appropriate. – If the EIAC representative has tickets assigned to them that they cannot follow up (due to e.g. a holiday), they must update the EPOT at the weekly onboarding meeting or onboarding-team@eosc-portal.eu email address to find another EIAC team member that can take on the activities, and when they agree, assign the tickets to them.

11.2 Guidance Note for Onboarding an Interoperability Guideline to EOSC

The following Guidance Note for Onboarding an Interoperability Guideline to EOSC[5] are published for reference by users and provides additional context beyond the EPOT Onboarding Procedure for Interoperability Guidelines[6].

11.2.1 Introduction

The EOSC-IF will be built upon a wide range of components, such as standards, APIs, policy frameworks, etc. Starting with human-readable interoperability guidelines that aim to promote and facilitate interoperability across services, resources, infrastructures and domains, the Framework will begin to incorporate more executable, machine-readable artefacts that would support composable or executable workflows at a later stage. Guidance notes for writing an EOSC Interoperability Guideline for inclusion in the EOSC Interoperability Framework are provided[7].

It is assumed that Interoperability Guidelines will be already established and published by an Infrastructure or Scientific Domain/Community, and therefore an author will not intend to write a new Interoperability Guideline for submission to the EOSC Interoperability Framework. However, it is conceivable that an existing Interoperability Guideline or collection of Guidelines would benefit from being structured into an EOSC Interoperability Guideline. In this case, a Word template has been provided[8] that is tailored towards Interoperability Guidelines produced under the EOSC Future project, but that can also be utilised (ensuring that the relevant copyright and project funding statements are updated as required), and the following guidance notes can be utilised in producing that document.

Interoperability Guidelines that meet the Inclusion Criteria[9] can be onboarded to the EOSC Portal using the EOSC Providers Portal.

The EOSC Interoperability Registry allows Providers to onboard a description of the Interoperability Guideline that contains references to the relevant documentation, APIs, protocols, etc that the Guideline recommends. The description 'record' that results from the onboarding procedure is called an 'Interoperability Guideline Profile'.

11.2.2 Types of Interoperability Guideline

An **EOSC-Core Interoperability Guideline** describes specifications for the purposes of interoperating resources with EOSC-Core services. EOSC-Core guidelines provide context and description in order to provide technical instructions to scientific resource providers that would like to integrate their services and/or resources with (or be interoperable with) one or more EOSC-Core Services. These guidelines are created by operators of EOSC-Core Services to provide guidance to scientific service and resource providers who wish to benefit from the functionalities that the EOSC-Core services offer.

An **EOSC-Exchange Interoperability Guideline** (thematic/community) refers to where thematic or community services and resources interoperate with each other, usually at a domain level such as Cluster-specific. These guidelines are created by scientific community experts to provide guidance to their communities and highlight the importance of their interoperability efforts at the European-wide EOSC level. They are further intended to increase awareness of the existing thematic and community accomplishments to facilitate/increase interoperability relating to specific domains.

An **EOSC-Exchange Interoperability Guideline (horizontal)** refers to where services and resources interoperate with each other across communities and infrastructures. These are important at the EOSC Exchange level to help EOSC facilitate cross-linking of services and resources while staying domain-agnostic.

---

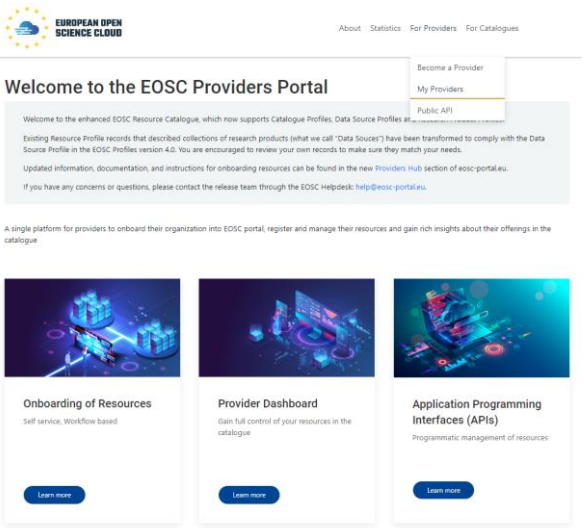
[5] https://zenodo.org/record/7929834
[6] https://wiki.eoscfuture.eu/display/EOSCOB/EPOT+Procedure-14%3A+Onboard+an+Interoperability+Guideline
[7] Guidance note: Writing a Guideline for the EOSC Interoperability Framework: 10.5281/zenodo.7929870
[8] https://eosc-portal.eu/eosc-interoperability-framework/eosc-if-templates-and-information
[9] https://wiki.eoscfuture.eu/display/EOSCSMS/EOSC-Exchange+inclusion+criteria#EOSCExchangeinclusioncriteria-3.6.Additionalinclusioncriteriatoonboardaninteroperabilityguideline

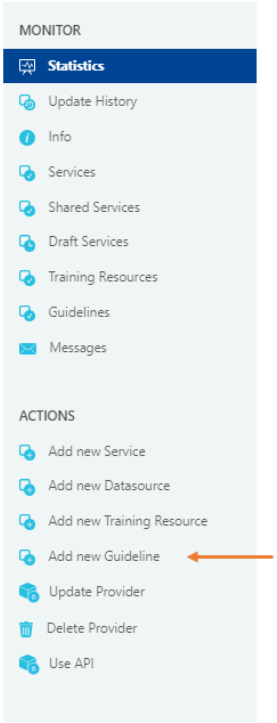
These are of wider applicability than the thematic/horizontal guidelines and have a much broader scope. These guidelines are created by scientific community experts to provide guidance to their own community and highlight their importance at the European-wide EOSC level. However, this, similarly, makes them more difficult to develop and requires cross-Cluster engagement and representation to ensure they are well-rounded and widely applicable to a diverse range of intended guideline users.

11.2.3 How to Onboard an Interoperability Guideline for inclusion in the EOSC Interoperability Framework

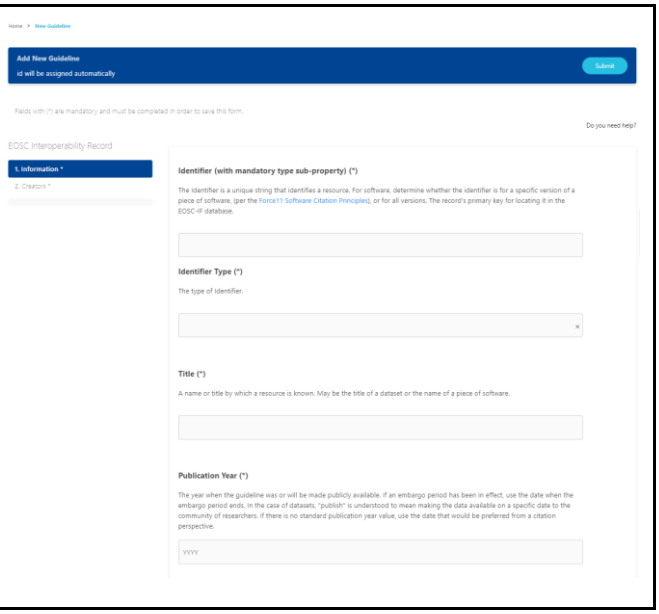
Please note, it is assumed that the owner of the Interoperability Guideline already has a Provider Profile within the EOSC Providers Portal and has administration rights for the related Provider Profile. If not, please refer to further instructions in the EOSC Provider Hub for creating such a Profile.

The Provider will be treated as the lead organisation for the Interoperability Guideline to be onboarded, but additional creators can be added to the Interoperability Guideline Profile.

Instruction	Detail
Navigate to the EOSC Providers Portal	providers.eosc-portal.eu/home
Select My Providers	

Instruction	Detail
The 'Add new Guideline' option can be selected from inside the Provider dashboard.	

11.2.4 Describe the Interoperability Guideline

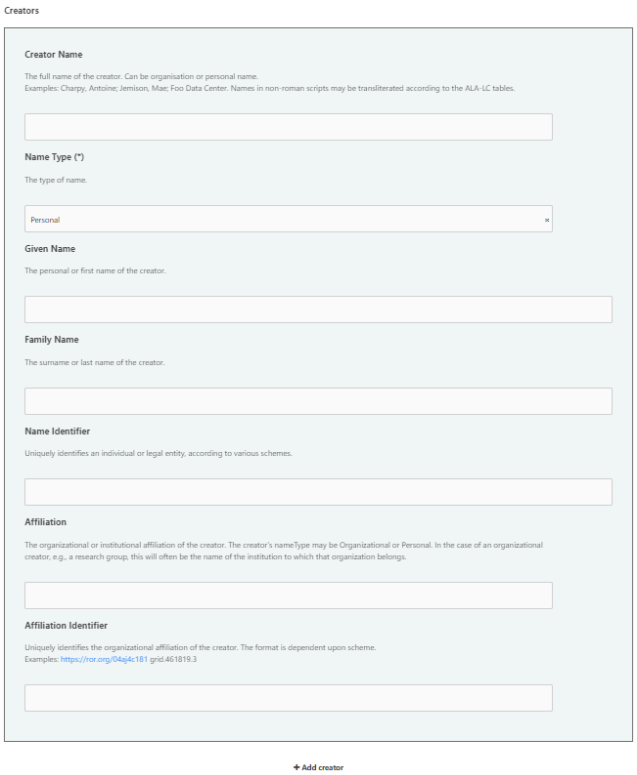
The Provider Representative is asked to complete a two-page form that is based on the EOSC Interoperability Guideline Profile[10]. The attributes are based on the Data Cite Metadata Scheme 4.4[11], and the guidance below should help to clarify how these attributes should be completed in the context of Interoperability Guidelines for the EOSC.	

[10] https://wiki.eoscfuture.eu/display/PUBLIC/EOSC+Interoperability+Guideline+Profile+-+Data+Model
[11] DataCite Metadata Working Group. (2021). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.4. DataCite e.V. https://doi.org/10.14454/3w3z-sa82

Identifier	The Identifier is a unique string that identifies the Guideline. This is likely to be the DOI or URI where the document is published. DOI is preferred. Note that this must be unique, and the URL inserted here will be treated as the primary key.
Identifier Type	Select the Identifier Type from the drop down list.
Title	The name or title by which the Guideline is known. The title will influence the search results returned by the users that look for Interoperability Guidelines, so it should be as concise and specific as feasible. Maximum 250 characters allowed.
Publication Year	The year when the guideline was or will be made publicly available. If an embargo period has been in effect, use the date when the embargo period ends.
Resource Type	This should be a brief abstract description of the resource, and should include the nature of the Guideline, i.e., whether it is API, metadata schema, specification, policy, framework, human-readable guideline, human readable guideline, configuration template. Maximum 1000 characters allowed.
Resource Type General	This will default to 'Guideline' in order to signal an Interoperability Guideline in the EOSC Marketplace; additional values will be added to the Controlled Vocabulary for this field in future releases.
Related Standards	This attribute is designed to capture details regarding the main standards, protocols, APIs, etc., that are adopted by this Interoperability Guideline. This should point to related standards only when it is a prerequisite/dependency, and where it is likely to influence the manner in which a Provider would design towards interoperability based on the guideline. Multiple Related Standards can be added. Maximum 250 characters allowed. The Related Standard URI field should be filled with the related URI. A human-readable name is to be added in the 'Related Standard Identifier' field.
Rights	The 'Rights' field is mandatory, and should include any rights information related to the Guideline. The property may be repeated to record complex rights characteristics. The Provider must include a rights management statement for the resource or reference a service providing such information. Include embargo information if applicable. Use the complete title of a licence and include version information if applicable. May be used for software licences. Examples: Creative Commons Attribution, 3.0 Germany License, Apache License, Version 2.026.

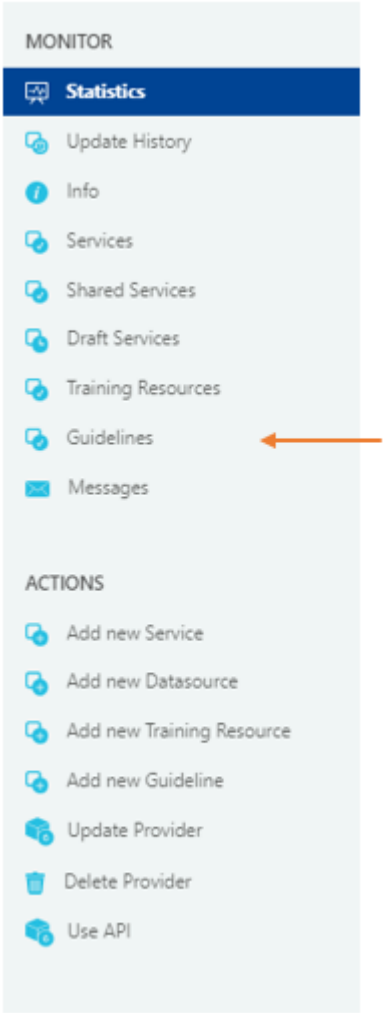
	Maximum 500 characters allowed.
Rights URI	This field should be completed with the URI relating to the rights statement.
Rights Identifier	This field should be completed with the short, standardised version of the licence name. Example: CC-BY-3.0.
Description	Complete the box on the left-hand side with a summary of the guideline. The description can include additional information that does not fit in any of the other categories and may be used for technical information. Consider including a concise description of the main features and benefits of interoperation in the manner described, and a brief description of the use case, problem statement, gap or need that this interoperability addresses. The box on the left-hand side allows formatting, and the box on the right allows you to see how the text will be represented in the UI. Maximum 500 characters allowed.
Status	Providers should initially select the 'Proposed' value from the drop-down menu. The EPOT team will select and/or advise the most appropriate value during the Onboarding procedure. This field relates to the status of the Guideline in relation to its inclusion in the EOSC Interoperability Framework, and not to the lifecycle status of the related APIs, standards, protocols, service, product, etc.
Domain	Please select the most appropriate top-level Domain to which the Guideline applies from the drop-down menu. Guidelines that can be utilised across all domains can be classified as 'Generic'.
EOSC Guideline Type	Please select the most appropriate Guideline Type from the drop-down menu. Please refer to section 2 above for more guidance.
EOSC Integration Options	This optional attribute is generally applicable to EOSC-Core Interoperability Guidelines, where a number of types of Integration can be achieved with a Core Service. This field may not be applicable to all types of Interoperability Guideline.

11.2.5 Describe the Creators of the Interoperability Guideline

The second page of the form relates to the Creators of the Guideline. The Creator field is optional, but it can be used when multiple authors, organisations, communities, projects or initiatives have been involved in the creation of the Guideline. Creators can be organisations as well as people, but please note that the Provider that onboards the Interoperability Guideline will be its lead organisation, rather than any Creator organisation captured in this form. Please create a Creator form for each Creator to be cited.	
Creator Name	This is the full name of the creator. Can be an organisation or a personal name. Examples: Charpy, Antoine; Jemison, Mae; Foo Data Center. Names in non-roman scripts may be transliterated according to the ALA-LC tables.
Name Type	This field allows the user onboarding the Interoperability Guideline to specify whether the Creator Name is personal or the name of an organisation.
Given Name	The personal or first name of the creator.
Family Name	The surname or last name of the creator.
Name Identifier	This optional field uniquely identifies an individual or legal entity, and can be used to add a Name Identifier according to the required scheme.
Affiliation	This optional field can be used to convey the organisational or institutional affiliation of the creator. The creator's nameType may be Organisational or Personal. In the case of an organisational creator, e.g., a research group, this will often be the name of the institution to which that organisation belongs.
Affiliation Identifier	This optional field can be used to uniquely identify the organisational affiliation of the creator. The format depends on the scheme. Examples: https://ror.org/04aj4c181 grid.461819.3

When the forms are complete, scroll to the top of the page and click 'Submit. At this point a new ID will be assigned in the Registry's database and the form will commence the EOSC Onboarding Procedure for Interoperability Guidelines, where a member of the EOSC-IF Governance team will help you through the process.	

11.2.6 Viewing and Maintaining Interoperability Guidelines

It is possible to view a summary of a Pending, Approved or Rejected Interoperability Guideline Profile in the Provider Dashboard, and to edit or delete the Profile itself.	

11.2.7 Related documents

- EPOT Procedure-14: Onboard an Interoperability Guideline (relating to EPOT team activities).[12]
- Guidance note: Writing a Guideline for the EOSC Interoperability Framework: 10.5281/zenodo.7929870.
- Guidance note: Reviewing a Guideline for the EOSC Interoperability Framework: 10.5281/zenodo.7929900.
- Guidance note: Assigning an Interoperability Guideline to an EOSC Service Profile: 10.5281/zenodo.7929856.
- EOSC Interoperability Guideline Profile Data Model.

11.2.8 References

(1) Eosc-portal.eu. 2021. EOSC Portal. [online] Available at: https://eosc-portal.eu/.

(2) EOSC Onboarding Inclusion Criteria, [online] Available at: https://wiki.eoscfuture.eu/display/EOSCSMS/EOSC-Exchange+inclusion+criteria.

(3) Data Cite Metadata Schema 4.4, 201 [online] Available at: DataCite Metadata Working Group. (2021). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.4. DataCite e.V. https://doi.org/10.14454/3w3z-sa82.

(4) EOSC Interoperability Guideline Profile Data Model [online] Available at: https://wiki.eoscfuture.eu/display/PUBLIC/EOSC+Interoperability+Guideline+Profile+-+Data+Model.

[12] https://wiki.eoscfuture.eu/display/EOSCOB/EPOT+Procedure-14%3A+Onboard+an+Interoperability+Guideline

12. Appendix E - Interoperability Guideline Profile

The following data model was inspired initially by the Data Cite Metadata Schema 4.4[13] to requirements articulated by participants of EOSC Future Work Package 3 and incorporated into the EOSC Portal Profiles.[14]

12.1 Basic Information

Code	Attribute Name	Human Readable Attribute Name	Description	Type	Multiplicity	Attribute Grouping	Required	Public
EIG.BAI.01	**EOSC Interoperability ID**	**EOSC Interoperability ID**	(auto-assigned)	Integer Auto-assigned by database	1	N/A	Mandatory	No
EIG.BAI.02	**Identifier**	**Identifier**	The Identifier is a unique string that identifies a resource. For software, determine whether the identifier is for a specific version of a piece of software, (per the Force11 Software Citation Principles), or for all versions. The record's primary key for locating it in the EOSC-IF database.	String DOI (Digital Object Identifier) registered by a DataCite member. Format should be "10.1234/foo"	1	N/A	Mandatory	Yes: PP & MP

[13] DataCite Metadata Working Group. (2021). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.4. DataCite e.V. https://doi.org/10.14454/3w3z-sa82

[14] https://wiki.eoscfuture.eu/display/PUBLIC/EOSC+Portal+Profiles

EIG.BAI.03	**identifierType**	**Identifier Type**	The type of Identifier.	Controlled List DOI URI	1	N/A	Mandatory	Yes: PP & MP

12.2 Provider Information

Code	Attribute Name	Human Readable Atribute Name	Description	Type	Multiplicity	Attribute Grouping	Required	Public
EIG.BAI.04	**ProviderID**	**Provider**	The Provider of the Guideline. Individual Creators may be added to the record if required to cite contributors and/or authors of the the guideline.	Auto-assigned by database (Provider is logged in when onboarding a new Guideline; ID defaults to ProviderID of Provider that the user has logged into)	1	N/A	Mandatory	Yes: PP & MP
EIG.BAI.27	**CatalogueID**	**Catalogue**	The ID of the Catalogue this Resource is originally registered at. (Catalogue Provider is logged in when onboarding a new Guideline)	Auto-assigned by database (catalogueID)	1	N/A	Optional	Yes: PP & MP

EIG.BAI.05	**creatorName**	**Creator Name**	The full name of the creator, if necessary to include individual contributors or creators. May be a corporate/institutional or personal name.	String (100) Examples: Charpy, Antoine; Jemison, Mae; Foo Data Center Note that the personal name, format should be: family, given. Names in non-roman scripts may be transliterated according to the ALA-LC tables.	Multiple	Creators	Optional	Yes: PP & MP
EIG.BAI.06	**nameType**	**Name Type**	The type of name	Controlled List: Organizational Personal	1		Mandatory	Yes: PP & MP
EIG.BAI.07	**givenName**	**Given Name**	The personal or first name of the creator Example: Antoine; Mae	String (100)	1		Optional	Yes: PP only
EIG.BAI.08	**familyName**	**Family Name**	The surname or last name of the creator Example: Charpy; Jemison	String (100)	1		Optional	Yes: PP only

EIG.BAI.09	**nameIdentifier**	**Name Identifier**	Uniquely identifies an individual or legal entity, according to various schemes. The format is dependent upon scheme utilised.	String (100) Examples: 0000-0001-nnnn-nnnn (ORCID), or ISNI (16 digits, the last character being a check character. The check character may be either a decimal digit or the character "X".	1		Optional	Yes: PP & MP
EIG.BAI.10	**affiliation**	**Affiliation**	The organizational or institutional affiliation of the creator The creator's nameType may be Organizational or Personal. In the case of an organizational creator, e.g., a research group, this will often be the name of the institution to which that organization belongs.	String (250)	1		Optional	Yes: PP & MP
EIG.BAI.11	**affiliationIdentifier**	**Affiliation Identifier**	Uniquely identifies the organizational affiliation of the creator. The format is dependent upon scheme. Examples : https://ror.org/04aj4c181 grid.461819.3	String (250)	1		Optional	Yes: PP & MP

12.3 Detailed information

Code	Attribute Name	Human Readable Atribute Name	Description	Type	Multiplicity	Attribute Grouping	Required	Public

EIG.BAI.12	**Title**	**Title**	A name or title by which a resource is known. May be the title of a dataset or the name of a piece of software.	String (250)	1	N/A	Mandatory	Yes: PP & MP
EIG.BAI.13	**PublicationYear**	**Year of Publication**	The year when the guideline was or will be made publicly available. If an embargo period has been in effect, use the date when the embargo period ends. In the case of datasets, "publish" is understood to mean making the data available on a specific date to the community of researchers. If there is no standard publication year value, use the date that would be preferred from a citation perspective.	Integer YYYY	1	N/A	Mandatory	Yes: PP & MP
EIG.BAI.14	**ResourceType**	**Resource Type Description**	A description of the resource, for example API, metadata schema, specification, policy, framework, human-readable guideline, human readable guideline, configuration template	String (1000)	Multiple	Resource Types Info	Mandatory	Yes: PP & MP
EIG.BAI.15	**resourceTypeGeneral**	**Resource Type Category**	The general type of a resource	Controlled List: Guideline	Multiple* (*additional values will be introduced in future releases)		Mandatory	Yes: PP & MP

EIG.BAI.16	**Created**	**Created**	Time/date the record was created. Auto-assigned by database	ISO8601 DateTime (2022-07-27)	1	N/A	Mandatory	Yes: PP only
EIG.BAI.17	**Updated**	**Updated**	Time/date the record was last saved, with or without modifications. Auto-assigned by database	ISO8601 DateTime (2022-07-27)	1	N/A	Mandatory	Yes: PP only
EIG.BAI.18	**RelatedStandard**	**Related Standard URI**	Standards related to the guideline This should point out to related standards only when it is a prerequisite/dependency, and likely to influence a Provider's design towards interoperability based on the guideline.	URL	Multiple	Related Standards	Optional	Yes: PP & MP
EIG.BAI.28	**RelatedStandard Identifier**	**Related Standard Identifier**	The name of the related standard.	String (250)	Multiple		Optional	Yes: PP & MP
EIG.BAI.19	**Rights**	**Rights**	Any rights information for this resource. Provide a rights management statement for the resource or reference a service providing such information. Include embargo information if applicable. Use the complete title of a license and include version information if applicable. May be used for software licenses.	String (500)	Multiple	Rights	Mandatory	Yes: PP & MP

			Examples: Creative Commons Attribution 3.0 Germany License Apache License, Version 2.026					
EIG.BAI.20	**rightsURI**	**Rights URI**	The URI of the license Example: https://creativecommons.org/licenses/by/3.0/de/	URL	Multiple		Mandatory	Yes: PP & MP
EIG.BAI.21	**rightsIdentifier**	**Rights Identifier**	A short, standardised version of the licence name Example: CC-BY-3.0 A list of identifiers for commonly-used licenses may be found here: (https://spdx.org/licenses/).	String (250)	Multiple		Mandatory	Yes: PP & MP
EIG.BAI.22	**Description**	**Description**	Provide an abstract or summary of the guideline. Description can include additional information that does not fit in any of the other categories. May be used for technical information.	String (1000)	1	N/A	Mandatory	Yes: PP & MP

EIG.BAI.23	**Status**	**Status**	Status of the resource.	Controlled List: Candidate Proposed Consultation On Hold Update Pending Accepted Operating Deprecated Abandoned Withdrawn Rejected	1	N/A	Mandatory	Yes: PP & MP
EIG.BAI.24	**Domain**	**Domain**	Intended Audience for the Guideline	Controlled List: EOSC Profile Domains	1	N/A	Optional	Yes: PP & MP
EIG.BAI.25	**EOSC Guideline Type**	**EOSC Guideline Type**	The type of Interoperability Guideline according to the definitions at https://eosc-portal.eu/eosc-interoperability-framework	Controlled List: EOSC-Core Interoperability Guideline EOSC-Exchange Interoperability Guideline (Thematic) EOSC-Exchange Interoperability Guideline (Horizontal)	1	N/A	Mandatory	Yes: PP & MP

EIG.B AI.26	**EOSC Integration Options**	**EOSC Integration Options**	A short summary of any options to integrate this guideline (if applicable).	String (100)	Multiple	N/A	Optional	Yes: PP & MP

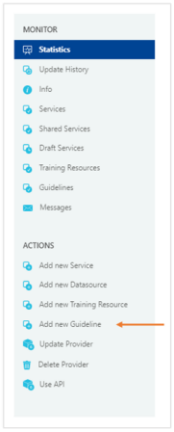
12.4Controlled Values

Status	Description	Next state/s
Candidate	Identified and registered, but no proposal has been submitted.	Proposed, on hold, abandoned, withdrawn, rejected
Proposed	Proposed but consultation has not commenced	Consultation, on hold, abandoned, withdrawn, rejected
Consultation	This would operated as 'Pending Approval' status (like 'patent pending')	Accepted, rejected, on hold, withdrawn, abandoned
On Hold	To be used in the event that further information is required before consultation can conclude	Consultation, abandoned, withdrawn, rejected
Update Pending	To be used either: 1. During the proposal process to indicate that a new version is to be considered before the proposed version is accepted. OR 2. During the operating phase to indicate that a new version will be published shortly.	1. Proposed, consultation, on hold, accepted 2. Deprecated, abandoned, withdrawn, rejected

Status	Description	Next state/s
Accepted	To identify Artefacts that have been accepted by the EOSC Interoperability Advisory Board but have not yet been announced as operational.	Operating, deprecated, withdrawn
Operating	To identify Artefacts that have been accepted by the EOSC Interoperability Advisory Board and have been announced as operational.	Deprecated, withdrawn
Deprecated	Artefacts that have been superseded or replaced by new versions but are to be retained in the Registry.	n/a
Abandoned	An artefact that was proposed but its Guardian and/or Proposer is unresponsive to requests from the [role]	n/a
Withdrawn	An artefact that has been withdrawn from the proposal process by its Guardian or Proposer	n/a
Rejected	An artefact that was rejected at the proposal stage by the EOSC Interoperability Advisory Board	n/a

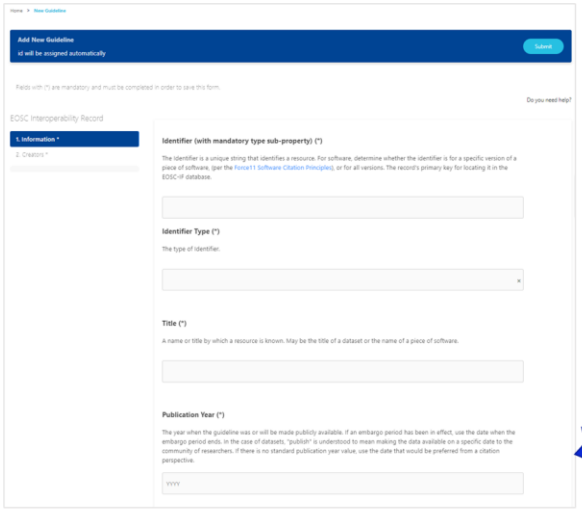
13. Appendix F – Updated EOSC Interoperability Guideline Features

13.1 New features: Interoperability Guideline Onboarding capabilities



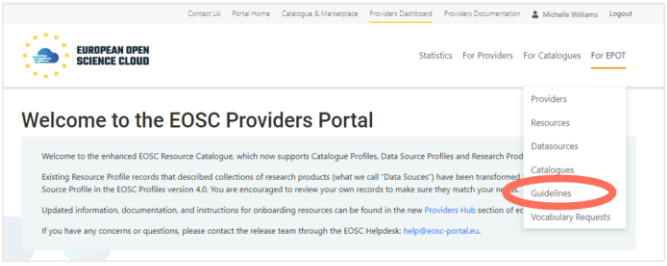
M25

- 'Add new Guideline' option can be selected from inside the Provider dashboard.



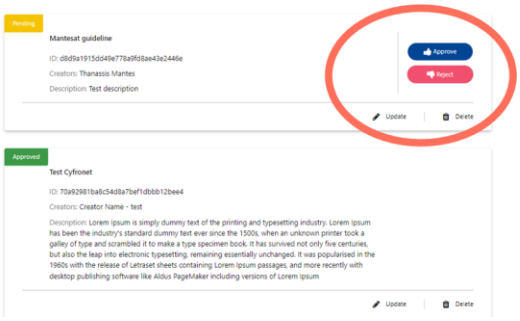
- Provider can add Guideline metadata to the EOSC-IF Registry
- Presents a form to by completed by the user that maps to the EOSC Interoperability Guideline Profile attributes
- IGs point out to the existing documentation in formal repositories, e.g. Zenodo
- Initially deployed in EPOT Admin area but M25 moves it to the Provider Portal

M22 M25



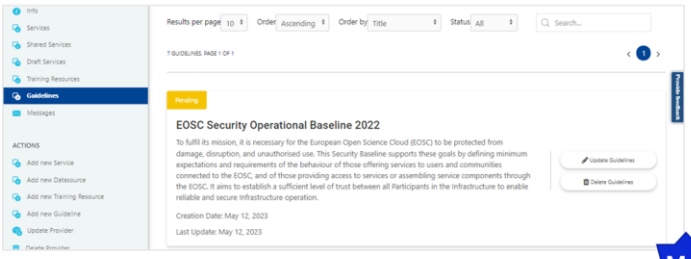
M22

- EOSC-IF Registry deployed initially inside the EPOT administration area
- Allowed registration of guidelines from within the Project as an early trial of the feature



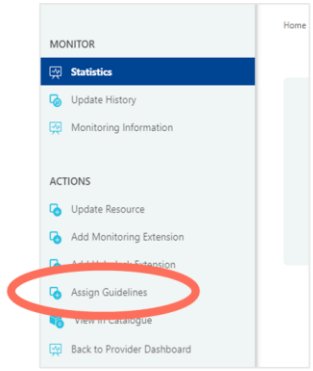
M25

- 'For EPOT' provides EPOT admin tools
- EPOT analyse the onboarded guideline referring to EIAC/EIAB and decide whether to approve or reject



M25

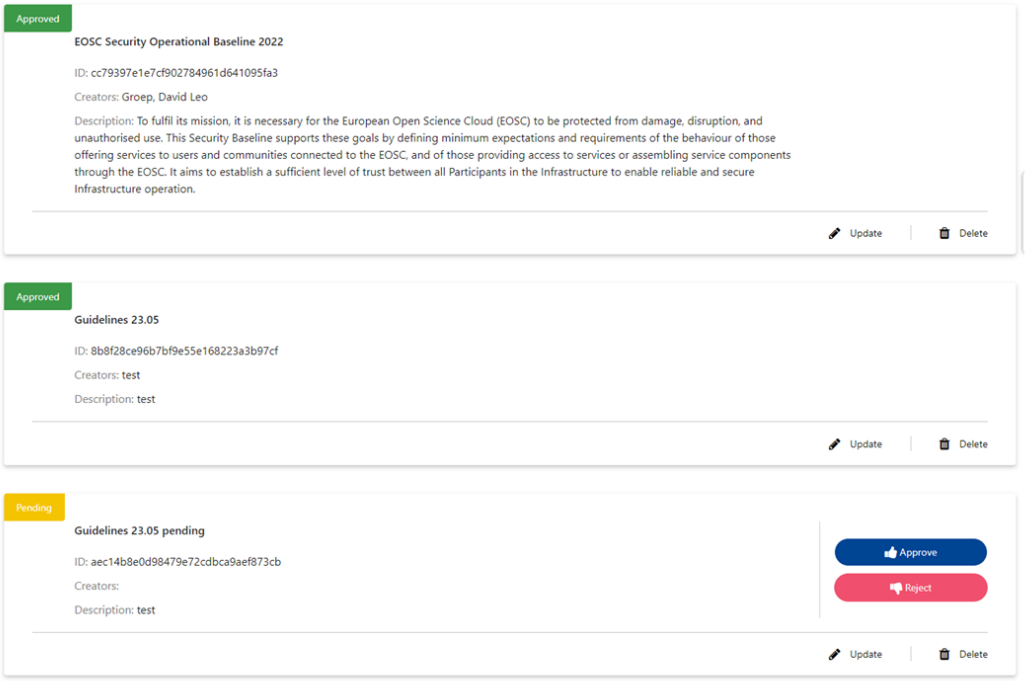
- Provider summary view of Pending, Approved or Rejected Interoperability Guideline Profiles in the Provider Dashboard, and to edit or delete the Profile itself.



M22

Service Providers can Assign Guidelines to Services

13.2 New features: EPOT Admin

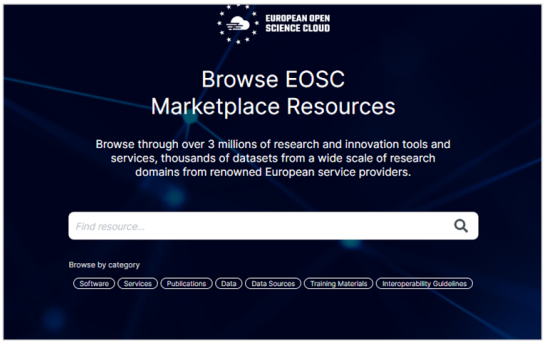


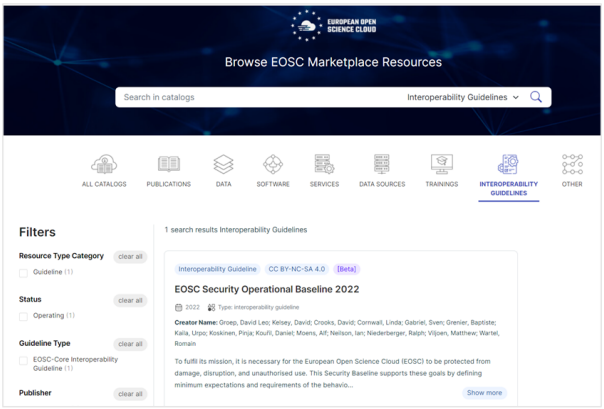
- EPOT Admin users can view, update, delete, approve or reject Interoperability Guidelines

M25

13.3 New features: Interoperability Guidelines in the Marketplace

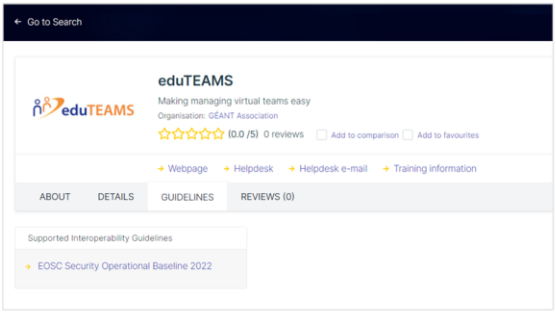
(as illustrated at M25 release; improvements has been requested for the UI for the next release)

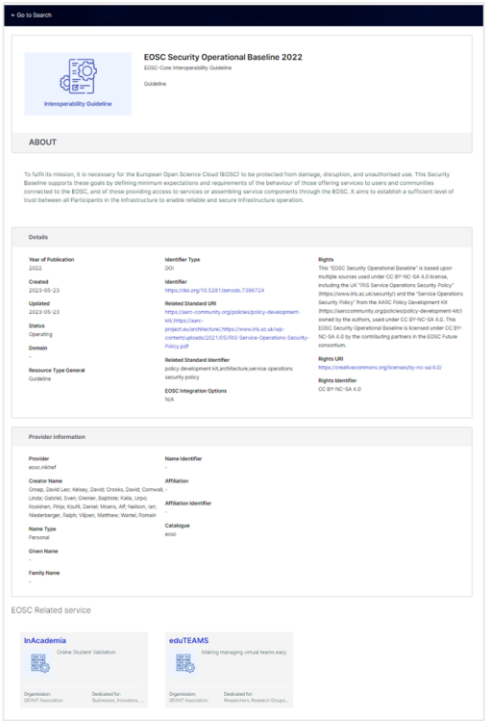




Users can search and browse for Interoperability Guidelines

M25





Users can view:

- the Interoperability Guideline description, and
- the EOSC Services that assert that they support the Guideline

M25

14. Appendix G – Summary of META-COVID Test Science Project (EOSC Future WP6.3)

COVID-19 metadata findability and interoperability in EOSC (META-COVID)

EOSC-Future Del. 3.2: Input from TSP "META-COVID"

Authors: C. Ohmann, S. Canham, M. Panagiotoupolou (ECRIN)

Holub P., Majcen K. (BBMRI)
Saunders G., Fratelli M., Tang J. (EATRIS)
Gribbon P., Reagon K. (EU-OPENSCREEN)
Kleemola M., Moilanen K. (CESSDA)
Broeder D., Daelemans W., Fivez P. (CLARIN)

Date: 22.05.2023

Version: Draft

The TSP was performed within the Horizon Europe funded project EOSC-Future and here within the test science project "META-COVID". This document is a summary of the project report together with an outline for a guideline to be developed for the EOSC Interoperability Framework.

Background

There are many metadata schemas that can describe the concrete outputs of research, e.g., papers and datasets, but relatively little work has been done on finding a metadata schema for the research process itself. This issue is handled differently between the metadata schemas. The problem is, however, that different disciplines have vastly different ways of organising research activities, for instance because of differences in funding models and mechanisms, or in requirements for approval, and thus differences in how and when research is split into discrete activities and labelled. In addition, research efforts take place at a variety of scales, have varying requirements for pre-published protocols, use a huge range of different methodologies and workflows, and may even draw upon different underlying assumptions. Research design, approach, strategy, and method are heavily influenced by the researchers' epistemology and research philosophy. The type of the research (e.g., hypothesis testing versus hypothesis generating), the methodology chosen (e.g., experimental, survey, cohort, case study) and the research methods applied (e.g., type of sampling) are of major importance in understanding the data generated, and thus in supporting any secondary use of that data. Therefore, metadata should go beyond a description of the data itself to include the basic elements of the research process ("contextual metadata").

Objectives

The objective of this TSP is to develop a framework for a contextual metadata, especially within research in the domains represented by the participating RIs. The intention is to initiate and apply an iterative process of consensus building amongst experts drawn from different disciplines in the LS and SSH.

Methods

As a first step in this process semi-structured interviews of representatives of the participating RIs were performed (BBMRI, ECRIN, EATRIS, EU-Openscreen from the LS; CESSDA, CLARIN from the SSH). The analysis was restricted to the following services: MDR for ECRIN, Directory for BBMRI, MICHA for EATRIS, ECBD for EU-OPENSCREEN, Data Catalogue for CESSDA and META-SHARE for CLARIN. For the majority of the RIs (BBMRI, EU-OPENSCREEN, CESSDA, CLARIN) the primary goal is to describe resource collections. In ECRIN the goal is to describe a clinical study and in EATRIS a chemosensitivity assay. Except for EATRIS and EU-OPENSCREEN, the primary purpose is to provide metadata that can be searched to find resources. EU-OPENSCREEN is primarily a data hub and EATRIS defines standards for annotation of drug screening protocols.

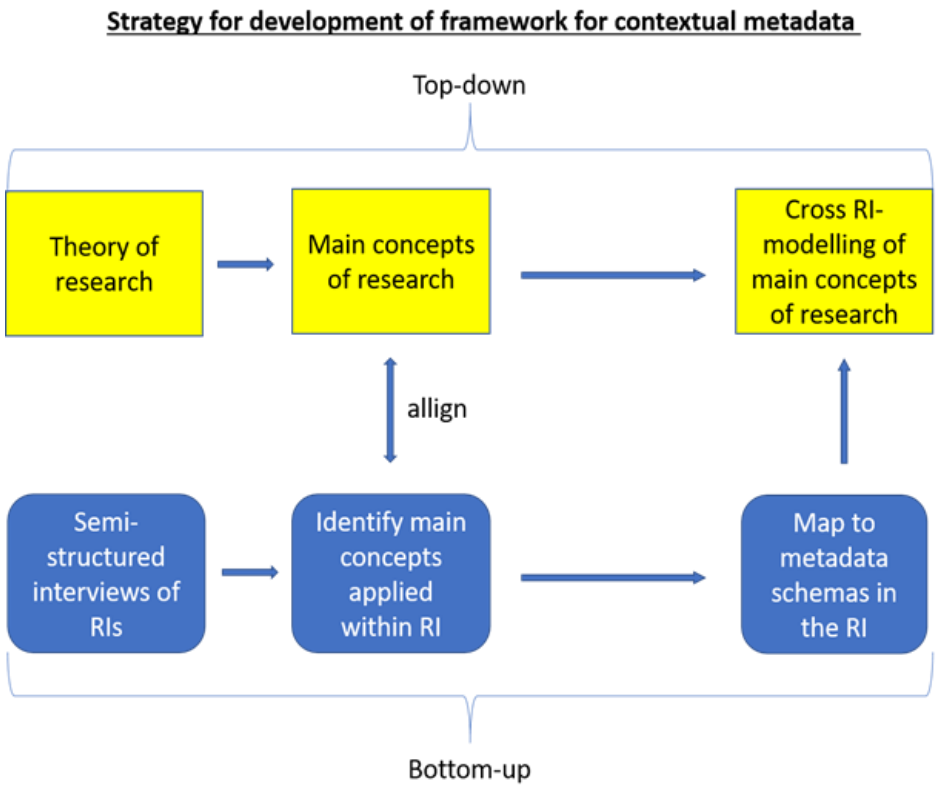
During the interviews, objective aspects of the use of contextual metadata in the RIs domain were assessed as well as opinion-based and subjective views of the RIs about use and potential value of contextual metadata in their domain. During the interviews, the main research concepts applied within a RI were identified and a mapping to metadata schemas in the RI was performed.

Within the semi-structured interviews, the following questions were discussed with the RI:

- What does "contextual metadata" mean to your RI?
- What services, protocols, standards, APIs are implemented in your RI to support harvesting of contextual metadata from outside (e.g., public, or non-public API)?
- Are the contextual metadata in your RI already linked to a research process graph or is it planned to do so?
- Do you believe that a greater generation and use of contextual metadata would be valuable enough to justify the additional effort that would likely be involved?
- From your viewpoint how could interoperability for contextual metadata between RIs be improved?
- What could be the best organisational framework for moving this work forward within EOSC?

The semi-structured interviews were performed between 9/2022 and 2/2023.

In the second part of the report, main entities of research in general were specified based on selected literature and mapped to the individual RIs. Finally, the analysis was synthesised cross-RI to provide a basis for a common framework for contextual metadata. The methodologic approach for this work is summarised in the figure:



The main entities of contextual metadata identified and of relevance for this TSP are:

- Research project
- Research activity
- Research programme
- Research process

Definitions for these basic entities are provided in the report and relations between entities specified. The key entity "research activity" is linked to the following attributes:

Linked attributes to "research activity"
- has a goal
- has a location
- has actor
- has (researchable) subject
- has procedure
- has methods
- uses tools/services
- produces research output

In the summary table below, an overview is given on the identification of attributes of a "research project" respectively "research activity" in the individual RIs. The metadata elements are classified according to "explicit available", "implicit available" and "not available" (Detailed information available in the full report).

"Research activity"/ "research project"	ECRIN MDR	BBMRI Directory (data collection)	EATRIS MICHA	EU-OPENSCREEN EBCD	CESSDA Data Catalogue	CLARIN CCR
Type	*Explicitly available*	*Implicitly available*	*Explicitly available*	*Explicitly available*	*Implicitly available*	*Implicitly available*
- has a goal	*Implicitly available*	*Implicitly available*	*Implicitly available*	*Implicitly available*	*Implicitly available*	*Implicitly available*
- has a location	*Partly explicitly/ implicitly available*	*Explicitly available*	*Not available*	*Not available*	*Explicitly available*	*Explicitly availabe*
- has actor	*Implicitly available*	*Explicitly available*	*Not available*	*Not available*	*Explicitly available*	*Explicitly available*
- has (researchable) subject	*Partly explicitly/ implicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*
- has procedure	*Partly explicitly/ implicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*

- has methods	*Partly explicitly/ implicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*	*Explicitly available*
- uses tools/services	*Implicitly available*	*Implicitly available*	*Explicitly available*	*Explicitly available*	*Not available*	*Explicitly available*
- produces research output	*Explicitly available*	*Not available*	*Not available*	*Not available*	*Explicitly available*	Explicitly available

A first full report of the semi-structured interviews and the data synthesis was provided and distributed to the participating RIs (April 2023). Written feedback was collected and an internal workshop with all participants was performed on 25 April 2023, where this draft was discussed and suggestions for improvement made. The document was revised, and a second version was provided for final written feedback (May 2023). From this feedback currently a final report will be provided. Final version of the report is foreseen for June 2023.

Discussion

The results of the study have shown that contextual metadata are dealt with quite differently between the RIs and their services. For three RIs, "research project" or "research activity" are the main entity (ECRIN, EATRIS, EU-OPENSCREEN). For these RIs information on most of the attributes of a "research project or/"research activity" is available, however, not always explicitly named but included somewhere in the metadata. In the other three RIs, primarily resources are described that have been derived from "projects", partly "research projects" and partly "non-research projects". In BBMRI the entity "study" is already foreseen but not yet implemented and a model has already been constructed linking "study" with the other entities "biobank" and "sample collection". For CESSDA and CLARIN, information about a "project" (respectively "study") underlying the sample or resource collection, is often available but distributed over different parameters of the metadata documentation. Here, primarily datasets that are outputs of projects are described. These datasets can be reused in different ways and for answering different research questions. Nevertheless, a considerable amount of information is already covered as demonstrated in the summarizing table.

Knowledge about a sample or data resource could be considerably increased when a "project" (whether a "research project" or "non-research project") could be explicitly identified as the source of a resource or data collection. PIDs for projects would be a useful concept, grant numbers may be available but not sufficient. If, in addition, the attributes of such a "project" could be explicitly linked to the "project", a basic corpus of contextual metadata would be available with a common structure cross-RIs but with RI-specific content for the attributes.

Whether a framework for contextual metadata could be transformed into a machine-actionable metadata model, depends primarily on the machine-actionability of the metadata schemas of the individual RIs. If this is not the case, the first step would be to explore if and how 'machine-actionability' could be increased. The situation is complex and difficult, however, because so much depends on the source systems in use within a domain and the willingness of metadata creators (ultimately, researchers) to agree on, learn and then use a highly structured descriptive system. This TSP represents a good starting point but much more is needed to progress. Making metadata 'machine actionable' needs, primarily, less ambiguous and more structured metadata, with entities, categories, and ontologies all unambiguously labelled and clearly defined. Here additional and extensive input from the individual RIs would be required. If this work can be completed, innovative methodological strategies for investigating and documenting metadata schemas could be applied as a basis for possible comparisons and convergence.

There are different options to support implementation of contextual metadata. Examples are knowledge graphs and business process modelling (BPMN). It may be that a graph-based data structure, which would echo the actors – processes – objects structure more closely, would be a better approach to storing data about the different aspects of "research activity". Here it is explicitly stated that research graphs can connect entities. Most widely used and of major relevance for the EOSC is the OpenAIRE research graph data model. The OpenAIRE Graph includes metadata and links between scientific products (e.g. literature, datasets, software,

and "other research outputs"), organizations, funders, funding streams, projects, communities, and (provenance) data sources. As such, the OpenAIRE graph already includes some of the basic entities to model contextual metadata (e.g., funder, project, organisation). Unfortunately, the research process, covering "research projects" and "research activities" as described in this report, is not modelled explicitly. Within the FAIRCORE4EOSC project, an EOSC Research Discovery Graph (RDGraph) will be developed that will become a flexible and federated EOSC search service across EOSC repositories that extends EOSC Research Catalogue making it compatible with the specifications provided by the RDA's "Open Scientific Graph for FAIR Data" working group and incorporating additional entities like the Research Activity Identifiers (RAiDs)[1]. In its core, it is based on the OpenAIRE Research Graph, and it will become exposed through its APIs and data dumps.

The work performed in this TSP should be closely linked to the Metadata Schema & Crosswalk Registry (MSCR), currently under development in FAIRCORE4EOSC. The MSCR allows registered users and communities to create, register and version schemas and crosswalks with PIDs. The published content can be searched, browsed and downloaded without restrictions. The MSCR also provides an API to facilitate the transformation of data from one schema to another via registered crosswalks. The framework for "contextual metadata" and crosswalks to other metadata schemas can be shared with the community for reuse and extension, thus improving step by step the possibility to explicitly characterise "research projects" and "research activities", to provide links between these entities and finally to improve interoperability cross-RIs and domains.

Conclusions

The work in our TSP has shown that already a substantial corpus of metadata is available. The amount and quality of these contextual metadata is highly dependent on the domain, the structure and goals of the RIs (project-centric, service-centric, resource-centric) and the metadata schemas applied in the RIs. A major problem is that quite often the contextual metadata are not explicitly identifiable in dedicated fields of metadata schemas but are distributed over several fields or implicitly included in the text of other fields. In addition, often essential contextual metadata are (at least partly) missing. The situation may get more difficult when different resources, which are related to the same "research project", are considered with respect to contextual metadata. Here, for example, contextual metadata identified through catalogues or repositories and contextual metadata implicitly contained in "research outputs" (e.g., publications) may not be consistent or even contradictory. The situation would considerably improve, if the contextual metadata specified in this report could be identified, the information cleaned and summarised in a separate entity "research project", to which other information objects related to a "research project" could refer. For legacy data, this can be done retrospectively but, in the future, a prospective approach would be preferable.

Guideline on contextual metadata for the EOSC Interoperability Framework

The guideline has take-up and acceptance from already a broad set of communities that have been included in the project (6 RIs from LS and SSH). Based on an assessment of the use of contextual metadata in the RIs and a converging process towards a metadata framework for contextual metadata with several rounds of feedback and updated versions, there is already evidence for the required communication, necessary for accepting a guideline in the scientific community.

The planned guideline will be based on the output of the TSP and will cover the following points:

- Raising awareness for the importance of contextual metadata

In summary, there was agreement between the majority of RIs participating in the TSP that a greater generation and use of contextual metadata would be valuable enough to justify the additional effort that would be likely involved. Availability and adequate use of contextual metadata is seen as necessary to improve research replicability and reliability. What is strongly needed is therefore to raise awareness for the importance of contextual metadata. Half of the RIs in the TSP suggested to explore whether input into the EOSC Interoperability Framework could be useful. Input into the EOSC Association (e.g., TF Semantic Interoperability) is also advocated by the majority of the RIs. In addition, a collaboration of EOSC with FAIRsharing and other resources is suggested.

The guideline will provide arguments and examples, demonstrating the benefit of using contextual metadata.

- Putting contextual metadata into perspective with other types of metadata (e.g., provenance)

For many researchers the definition of "contextual metadata" is still unclear, especially in relation to the concept "provenance". Contextual metadata will be defined more clearly in the proposed guideline than this was done in the past, by splitting it into 'direct output / data provenance' (e.g., how the data was generated and by whom, when, where etc., which is already covered to some extent), and 'the contextual research activity' (what was the study and what was its aim, approach, goal, programme context etc.), which is more difficult to deal with.

- Provide definitions for essential elements of contextual metadata

The essential elements of contextual metadata, covering "research project", "research activity", "research programme", "research process" will be defined in the guideline.

- Provide relations between contextual metadata elements

The relationships between the essential contextual metadata elements will be defined and explained in the guideline.

- Provide attributes for contextual metadata elements

Necessary attributes, characterising the basic entity "research activity" (respectively "research project") will be specified in the guideline.

- Practical implementation of the metadata framework in the different domains and RIs

The implementation of the framework for contextual metadata will be based upon existing and used metadata schemas in the different domains and RIs. No adaption or extension is needed to start the process. It will be specified how contextual metadata (explicitly or implicitly available) can be assessed and used for supporting of interoperability. The technical prerequisites for using the contextual metadata will be described (e.g., APIs, knowledge graphs, NLP). Use cases, demonstrating the benefit of taking contextual metadata into consideration, will be listed. Guidance will be given, how the documentation and management of contextual metadata can be improved in the future if existing metadata schemas used in the RIs are adapted or updated.

Procedure

The initial version of the guideline will be provided by a writing group, consisting of representatives of the different RIs participating in the TSP, completed by representatives from selected other domains and RIs. The structure of the guideline will follow the EOSC IF Guideline template. The draft guideline will be communicated to an even broader community of domains and RIs for feedback and improvement. Alignment between the different stakeholders will be sought and a first version will be produced. The guideline will be updated regularly. The impact of the guideline on interoperability between EOSC resources and services will be investigated when the guideline is 2 years in use.

15. References

[1] D3.2a *EOSC Architecture and Interoperability Framework* https://eoscfuture.eu/wp-content/uploads/2022/12/EOSC-Future-WP3-GEANT-D3.2a-EOSC-Architecture-and-Interoperability-Framework-2021-12-22.pdf

[2] D3.3a *Architecture and Interoperability Guidelines for Operational Services of the EOSC-Core* is not publicly available yet

[3] *EOSC Interoperability Framework Report* https://dx.doi.org/10.2777/620649

[4] https://eosc-portal.eu/eosc-interoperability-framework

[5] https://wiki.eoscfuture.eu/display/PUBLIC/EOSC+Portal+Profiles Note that this reference will be updated as we transition to a formal Github repository of profile specifications

[6] https://c4model.com/

[7] D10.1 *EOSC Future Stakeholder Engagement and Communication Strategy and Plan* https://eoscfuture.eu/wp-content/uploads/2022/12/EOSC-Future-WP10-CSC-D10.1-EOSC-Future-Stakeholder-Eng.-Com.-Strategy-Plan-2021-09-08.pdf

[8] https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/infraeosc-05-2018-2019

[9] https://eosc-portal.eu/eosc-regional-projects

[10] https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/infraeosc-04-2018

[11] https://eosc-portal.eu/esfri-thematic-cluster-projects

[12] https://c4model.com/

[13] https://icepanel.io/

[14] https://www.eosc.eu/advisory-groups/aai-architecture: The final documentation of the EOSC Association Architecture Task Force "EOSC AAI Architecture v2022" is due to be published soon

[15] *EOSC AAI Architecture* report https://data.europa.eu/doi/10.2777/8702

[16] *AARC Blueprint Architecture and Guidelines* https://aarc-project.eu/guidelines/

[17] Deliverable 7.3 *EOSC Federated Authorisation and Authentication Activities EOSC-Future-WP7-GEANT-D7.3-EOSC Federated Authorisation and Authentication Activities-2023-02-03.pdf (submitted to the EC, not yet publicly available)*

[18] *AARC Blueprint Architecture* https://doi.org/10.5281/zenodo.3672784

[19] Section 3.6 of https://wiki.eoscfuture.eu/display/EOSCSMS/EOSC-Exchange+inclusion+criteria

[20] *Charter for the EOSC Interoperability Advisory Board and EOSC Interoperability Area Chairs* https://eosc-portal.eu/eosc-interoperability-framework/eiab-and-eiac-charter

[21] EOSC Interoperability Framework Consultation https://eosc-portal.eu/news-and-events/news/eosc-interoperability-framework-consultation

[22] *Guidance Note for Reviewing EOSC Interoperability Guidelines* https://zenodo.org/record/7929900

[23] *EOSC Interoperability Guideline Profile* https://wiki.eoscfuture.eu/display/PUBLIC/EOSC+Interoperability+Guideline+Profile+-+Data+Model

[24] *EOSC Security Operational Baseline* https://search.marketplace.eosc-portal.eu/guidelines/eosc.df8f717f77566b7adde87ede8f55b31d?pv=search%2Fguideline&q=*

[25] EOSC IF Templates https://eosc-portal.eu/eosc-interoperability-framework/eosc-if-templates-and-information

[26] *Guidance Note for Writing an EOSC Interoperability Guideline* https://zenodo.org/record/7937507

[27] https://www.sshopencloud.eu/news/how-sshoc-responding-covid-19

[28] https://www.desy.de/coronavirus/index_eng.html

[29] https://by-covid.org/

[30] D4.1b *Back-Office design, functional and technical specifications*, is not publicly available yet.

[31] D5.1b *Front-Office Design, Functional and Technical Specifications*, is not publicly available yet.