An EOSC-hub Proposal for the EOSC Technical Architecture

Version:	1
Status:	Final
Dissemination Level:	Public
Document Link:	

Deliverable Abstract
This document describes the EOSC Technical Architecture proposed by EOSC-Hub. It is based on the concepts of service interoperability and end-to-end composition of services and foresees the definition of a reference architecture where all the EOSC main functions, interfaces, APIs and standards are identified.

EOSC-HUB RECEIVES FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT NO. 777536.

COPYRIGHT NOTICE

DELIVERY SLIP

Date	***Name***
From:	Diego Scardaci (EGI.eu) and Giacinto Donvito (INFN)

TERMINOLOGY

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

Contents

1 Introduction 6
2 The EOSC Portfolios and the EOSC Federating Core 7
3 Service composability 9
4 EOSC Technical Architecture 10
4.1 Main blocks of the technical architecture 10
4.2 EOSC Access Enabling and Federation services 12
4.3 EOSC Common services 16
4.4 EOSC Thematic Services 18
5 Relationship with the EOSC Architecture Working Group objectives 20
Appendix I. AAI Technical Specification 21
Appendix II. Cloud IaaS VM Management Technical Specification 25

Executive summary

This document describes the EOSC Technical Architecture proposed by EOSC-Hub. It is based on the concepts of service interoperability and end-to-end composition of services and foresees the definition of a reference architecture where all the EOSC main functions, interfaces, APIs and standards are identified. This work has taken into account the surrounding landscape and has followed the recommendations on the EOSC architecture from the European Commission and the EOSC governance.

As a basis to describe the proposed architecture, service categories were introduced, describing their functions, relationships and organisation in EOSC Service Portfolios. The concept of the end-to-end composition of the services was presented highlighting the most common integration scenarios and how services belonging to different categories can cooperate to create added-value solutions for research. EOSC-hub effort to foster service interoperability and the impact of the service composability on federating thematic services into the EOSC were also depicted.

Leveraging on the service categories and on the concepts of service interoperability and composition, a reference EOSC Technical Architecture was defined identifying main blocks and their interactions. The proposed architecture foresees the further specification of each block following a common approach that consists of (1) the identification of the main **macro-features/functions** offered per block and, for each macro-feature, (2) the definition of an **EOSC technical specification** that includes an high-level architecture, suggested EOSC standards and APIs and interoperability guidelines.

According to the proposed model, services offering a given macro-feature and compliant with the related EOSC technical specification will be able to interoperate, examples are the AAI services compliant with the AARC blueprint architecture and guidelines or monitoring and/or accounting systems able to exchange/share information and provide integrated views to the EOSC customers and service providers. We believe that this approach fits well with a varied environment like EOSC, where many solutions for a given technical need already exist. Furthermore, having well defined EOSC standards and APIs and related interoperability guidelines for each of the identified macro-features will foster the end-to-end composition of services. Indeed, other macro-features or thematic services can use the EOSC interfaces, described in the technical specifications, to exploit such macro-features.

In the proposed architecture, identifying macro-features and the related technical specifications for all the blocks could be a complex and long work, then we decided to follow an iterative approach starting from the functions that are more requested by the EOSC use cases. Also the technical specifications, initially prepared by the technical experts within the EOSC-hub project, will be iteratively improved collecting feedback by external people with expertise in the area and involving them in the maintenance and evolution of such specifications.

EOSC-hub already identified a considerable amount of macro-features per block and completed the technical specifications of the most relevant. However, we consider fundamental involving other relevant stakeholders in this work to have a real impact on the research world. For example, we think that including other technical experts to refine technical specifications and find consensus around them is fundamental. For this reason, we started a process to share our approach and collecting feedback. The first step was a webinar where we presented this work, then a formal feedback collection will start in the next few weeks and we are planning to organise a workshop by the end of this year involving the largest expected EOSC user groups.

Finally, EOSC-hub is intended to propose the EOSC technical architecture described in this document and the related approach to define the EOSC technical specification for macro-features to the EOSC Architecture WG for its adoption in the wider EOSC environment. EOSC-hub would also like to

collaborate with the WG on further refining the proposed architecture taking into account requirements and suggestions from the largest possible set of service providers and user communities.

1 Introduction

This document describes the EOSC Technical Architecture proposed by EOSC-Hub. It is based on the concepts of service interoperability and end-to-end composition of services and foresees the definition of a reference architecture where all the EOSC main functions, interfaces, APIs and standards are identified.

2 The EOSC Portfolios and the EOSC Federating Core

The proposed EOSC Technical Architecture is based on the different classes of EOSC services and on their interactions. Then, an introduction on such service categories and on their functions and relationships is necessary before describing the architecture.

As depicted in Figure 1, EOSC services are organised in two services portfolios:

- **EOSC Service Portfolio:** the external services which EOSC-hub either provides from its partners or onboards from the community to contribute to the larger portfolio of researcher-benefitting services within EOSC. The EOSC Service Portfolio contain:
  - **Thematic services:** community-specific capabilities including research core data, data products, scientific software, and pipelines. Examples of thematic services are: data resources and software tools to access, study and compare the data; data brokering services tailored to the needs of specific scientific communities;
  - **Common services:** they provide generic capabilities usable by any science discipline, each supporting aspects of the data lifecycle from creation to processing, analysis, preservation, access and reuse. Examples of services belonging to this category are multi-disciplinary services for data discovery, processing, workflow management and orchestration, data management, etc.
- **Hub Portfolio:** the internal services contributing to the federating core of EOSC, both for internal operation of the EOSC Hub and to offer as components to be integrated into the services of the EOSC Service Portfolio. They enable the other EOSC elements to deliver (greater) value to researchers across Europe. They can be further split in
  - **Access-enabling services:** deliver features allowing customers to easily exploit EOSC resources such as discovery, ordering and workflow enabling services;
  - **Federation services**: needed to operate the EOSC e.g. a common helpdesk, accounting information gathering, monitoring.

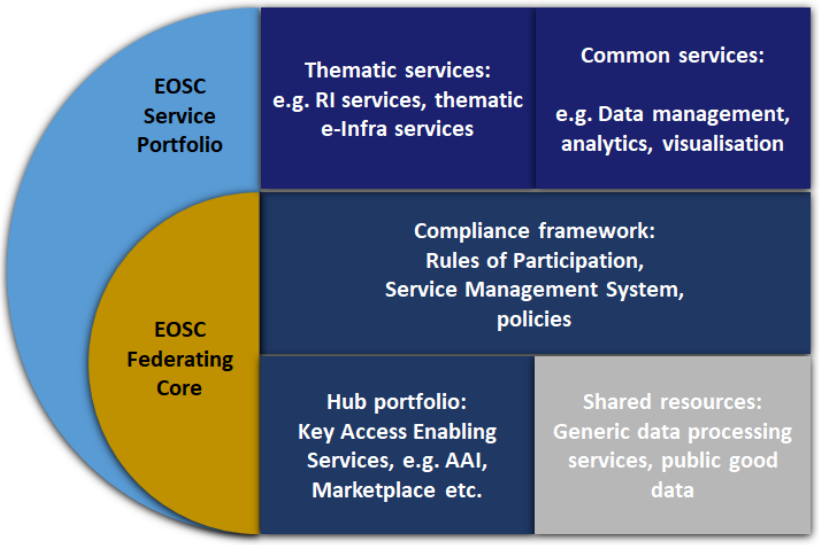

Figure 1. EOSC Service Portfolios and EOSC Federating Core.

Thematic services can be integrated with the services in the Hub portfolio to facilitate the users' access (e.g. the EOSC Portal and Marketplace) or to avoid to re-implement basic features, like authentication and authorisation, accounting, monitoring, etc. They can also adopt common services that already address some of their technical needs. Common services can also leverage on services of the Hub portfolio to deliver some of their functions. The integration of thematic, common, access enabling and federation services can be fostered through a large adoption of open and standard interfaces.

The Hub Portfolio is one of the key elements of the **EOSC Federating Core** together with the Compliance framework, made of Rules of Participation, EOSC Service Management System and other policies, and the Shared resources, a set of generic data processing services, commodity services, compute and storage resources and public good data managed and offered centrally by EOSC. More information about the EOSC Federating Core and the EOSC service portfolios are available in the EOSC-hub briefing paper "EOSC Federating Core Governance and Sustainability"[1] and in the EOSC-hub D2.6 "First Service roadmap, service portfolio and service catalogue"[2].

[1] https://documents.egi.eu/document/3479
[2] https://documents.egi.eu/document/3470

3 Service composability

The end-to-end composition of the services can be considered one of the most important added values provided by EOSC. Indeed, the service composability would allow EOSC service providers and users to select various services offered by EOSC and compose them according to their needs to create added-value solutions for research.

Typical service combinations are:

- A thematic service adopts some EOSC federation services to implement basic features (AAI, monitoring, accounting).
- A thematic service adopts common services that provide features to better exploit compute, storage and data resources.
- An EOSC user creates new scientific workflows integrating, for example, a data repository and some analytics services together.

The adoption of standard interfaces makes some of the services of the EOSC service catalogues already interoperable, these sub-classes of composable services need to be identified and made accessible through the EOSC Portal. Furthermore, according to the emerging needs from communities, other services can be made interoperable through integration activities. EOSC should provide technical guidelines (in terms of suggests EOSC standards and APIs) and technical support to both integrate services and facilitate the combined usage of (already) interoperable services.

Enabling the service composability would allow developers of the thematic services to easily reuse common, federation or access enabling services to implement basic features (AAI, accounting, monitoring, etc.) and exploiting in the best way compute, storage and data resources. Indeed, they, from one side, can focus on working on increasing the scientific added value of their services, and from the other side, rely on well-established and EOSC-compliant services for implementing the basic features. A large part of these reusable services will come from the experiences of the main European e-infrastructures and other relevant initiatives (such as those involved in the EOSC-hub project, EGI, EUDAT and INDIGO-DataCloud).

4 EOSC Technical Architecture

The EOSC Technical Architecture presented in this document is a reference architecture. In the field of software architecture or enterprise architecture, a reference architecture provides a template solution for an architecture for a particular domain. It provides a common vocabulary with which to discuss implementations, often with the aim to stress commonalities. A reference architecture often consists of a list of functions and some indication of their interfaces (or APIs) and interactions with each other and with functions located outside of the scope of the reference architecture[3].

Reference architectures can be defined at different levels of abstraction, in the context of EOSC, EOSC-hub decided to work at infrastructure/technical level. The architecture presented later in this document includes functions, interfaces, APIs and standard as technical concepts, with the final aim to foster interoperability and, then, the service composability. As part of this work, we are also defining a common vocabulary usable to define not only already available services, but also the ones will be joining EOSC catalogue in the future.

4.1 Main blocks of the technical architecture

Figure 2 describes the main blocks of the proposed EOSC technical architecture and their interactions. Each block contains all the functions offered by the related class of services (thematic, common, access enabling and federation services, see next sections for details).

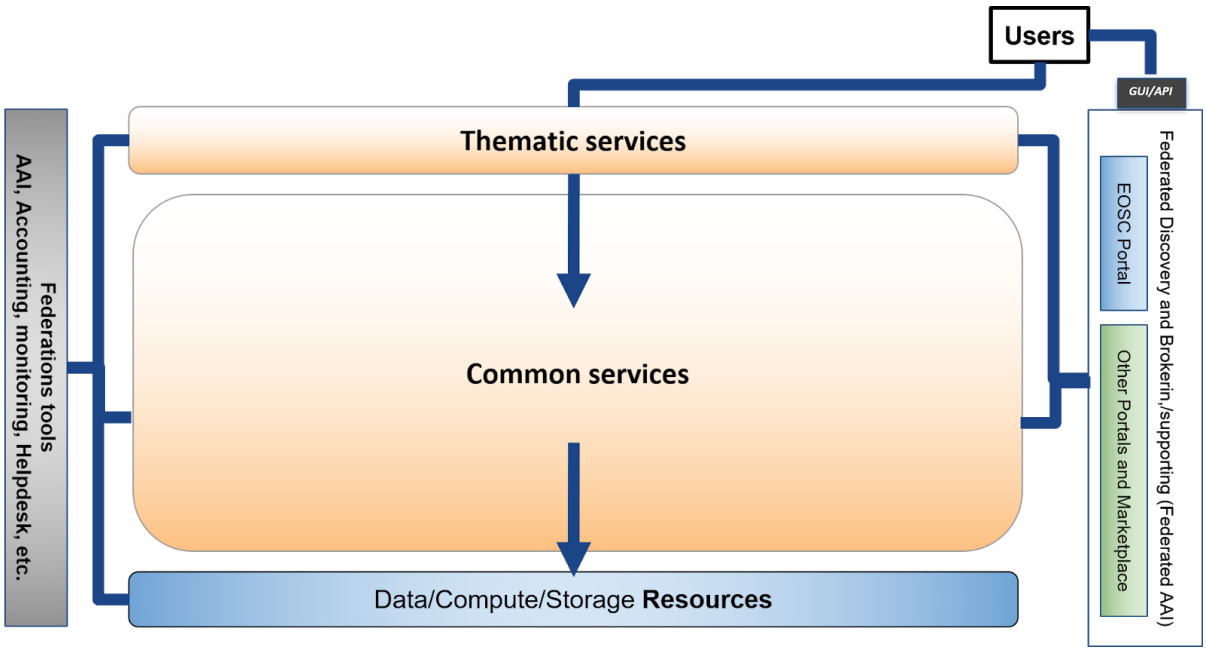

Figure 2. Main Service Blocks and their interactions.

EOSC users can exploit EOSC Thematic and Common services directly or through the GUI or API of an access enabling services like the EOSC Portal (or others portals and Marketplaces). Thematic services can leverage on Common services for added value features on top of data, compute and

[3] Definition from Wikipedia: https://en.wikipedia.org/wiki/Reference_architecture

storage resources. Federation tools support all these services providing basic features like authentication and authorisation, accounting, monitoring, etc. Pledged shared resources centrally managed by EOSC, including both commodity services and service capacity, are part of the Resources block and complement other EOSC resources directly managed by other service providers.

4.1.1 Approach to define the content of each block

EOSC-hub is working on defining the content of the main blocks of the architecture and specified a common approach to complete this task. It foresees the identification of the main **macro-features/functions** offered by each block and, for each of those, defining a **technical specification** that includes an **high-level architecture, suggested EOSC standards and APIs and interoperability guidelines**. This method would allow to have a set of services per macro-feature compliant with the related EOSC technical specification. Services belonging to this set would be able to interoperate thanks to the interoperability guidelines defined in the specification. Examples of these families of services can be AAI services compliant with the AARC blueprint architecture and guidelines or monitoring and/or accounting systems able to exchange/share information and provide integrated views to the EOSC customers and service providers. We believe that this approach fits well with a varied environment like EOSC, where many solutions for a given technical requirement already exist.

The definition of EOSC standards and APIs and related interoperability guidelines for each of the identified macro-features will foster the end-to-end composition of services. Indeed, from one side, being compliant with a specification for a given macro-feature, would allow a service to interoperate with other services offering the same macro-features and, from the other side, other macro-features/services can use the EOSC interfaces, described in the technical specification, to exploit such macro-features.

Identifying macro-features and the related technical specifications for all the blocks could be a complex and long work, then we agreed to follow an iterative approach starting from the functions that are more requested by the EOSC use cases. Also the technical specifications, initially prepared by the technical experts within the EOSC-hub project, could be iteratively improved collecting feedback by external people with expertise in the area and involving them in the maintenance and evolution of such specifications. The same is true also for the list of macro-features: they will evolve and change in the future adding/deleting features depending on the user requirements and on the projects/service providers that would join the EOSC in the future.

4.1.2 Technical Specification template

We have defined a template to collect information for each of the identified macro-features and define a technical specification, regardless of the block they belong to. It is structured as follows:

- Introduction: short description of the macro-feature highlighting its main functions.
- High-level Service Architecture: reference architecture of the macro-feature with highlighted the interfaces towards the other macro-features. It does not refer to any specific service.
- Adopted Standard: list with references of the main adopted standards and protocols/API.

- Interoperability guidelines: describe how similar services can be made interoperable with this macro-feature.
- Examples of solutions implementing this specification: list of already available Open Source services that are compliant with this specification.

In the following sections of the document, we describe the current status of the work to define the content of each main block of the infrastructure.

4.2 EOSC Access Enabling and Federation services

In the Access Enabling and Federation services blocks, **we intend as macro-feature any key access- and federation-enabling function needed to operate the EOSC**. Services offering these features according to the EOSC specification could be onboarded on the Hub service portfolio described in section 2.

We already identified an initial list of macro-features for this block leveraging on the experiences from some of the largest European e-infrastructures that are involved in the project. This list is detailed in the table below.

Table 1 – Access Enabling and Federation macro-features

Macro-features	***Short description***
EOSC Portal	The EOSC Portal provides a European-level delivery channel connecting the demand-side (the EOSC Customers) and the supply-side (the EOSC Providers) to allow researchers to conduct their work in a collaborative, open and cost-efficient way for the benefit of society and the public at large. In particular it delivers the following functions: • Enable different kinds of users, with different skills and interests, to discover, access, use and reuse a broad spectrum of EOSC Resources (services, datasets, software, support, training, consultancy, etc) for advanced data-driven research • Support interdisciplinary research and facilitate Resource discovery and access at the institutional and inter-institutional level • Allow researchers and institutions to focus on value creation through sharing and reuse as opposed to duplicating Resources and increase excellence of research and European competitiveness • Improve the provisioning of access to integrated and composable products and services from the EOSC Catalogue • Facilitate the composition of services and products to support multi-disciplinary science for example with high-level community-specific interfaces for running workflows involving EOSC services • Help Providers gain additional insight into potential Users outside their traditional constituencies • Give Providers the possibility to offer Resources under homogeneous terms of use, acceptable use policies, and in different configuration options, so that Users are guided in the choice.

	Use case. The Portal is particularly relevant to support on-demand access to EOSC through Business-to-User (B2U) and Business-to-Business (B2B) transactions. • B2U is applicable for consumer-oriented Resources appealing to a large potential User pool. B2U transaction will address the digital needs of individual researchers and short- and medium-term research projects. Because of the large user base, B2U transactions will be possible for those Resources supporting automated or semi-automated provisioning, a short acquisition process, requiring a low-level of specialisation, and which can be easily compared and chosen without requiring expert support. • On the other hand, B2B applies for the acquisition of bespoke solutions and/or of large quantities of EOSC Resources involving potentially multiple Providers. B2B suits the needs of research performing organisations and research infrastructures which need to cater for the long-term needs of a large pool of end users. The EOSC Portal Concept 2.0 provides extensive information on potential use cases and a participatory model for resource providers, which are provided with the choice of selecting different EOSC participation levels (e.g. Entry, Standard and High).
AAI	The EOSC AAI enables seamless access to research data and services in EOSC in a secure and user-friendly way. It also provides authorisation management for access control. It is based on the AARC blueprint architecture. The EOSC AAI follows the architectural and policy recommendations defined in the AARC project [AARC-Community]. As such, it enables interoperability across different SP-IdP-Proxy services, each of which acts as a bridge between the community-managed proxies (termed Community AAIs) managing the researchers' identity and the generic services offered by Research Infrastructures and e-Infrastructures (termed R/e-Infrastructures or Infrastructures). This is the "community-first" approach to the AARC Blueprint Architecture [AARC-G045], which enables researchers to sign in with their community identity via their Community AAI. Community-specific services are connected to a single Community AAI, while Infrastructure Services are connected to a single Infrastructure Proxy. Lastly, generic services may be connected to more than one Community AAI. Each Community AAI in turn serves as a bridge between external identity providers and the proxies to the e-infrastructure services. Specifically, Community AAIs connect to eduGAIN as service providers but act as identity providers from the services point of view, thereby allowing users to use their credentials from their home organisations. Complementary to this, users without an account on a federated institutional Identity Provider are still able to use social media or other external authentication providers for accessing services. Research communities can leverage the EOSC AAI services for managing their users and their respective roles and other authorisation-related

	information. At the same time, the adoption of standards and open technologies, including SAML 2.0, OpenID Connect, OAuth 2.0 and X.509v3, facilitates interoperability and integration with the existing AAIs of other e-Infrastructures and research communities. **Use Cases.** Access to all EOSC shared resources and access enabling services (e.g. the Portal, the Helpdesk, EOSC data and compute and storage resource tier) will require federated authentication and authorisation. In addition, the Life Science Research Infrastructure cluster, as well as other research infrastructures from other scientific domains like Social Sciences and Humanities and Physics, have been piloting different solutions to get AAI as a managed service.
Helpdesk	The helpdesk is the tool that supports Incident and Service Request Management to restore normal/agreed service operation within the agreed time after the occurrence of an incident, and to respond to user service requests. The service works as a unified ticketing system, by connecting individual providers' helpdesks to the central helpdesk instance, offering a standalone service interface. **Use case.** The helpdesk tool is necessary to support Incident and Service Request Management of the resources provided by EOSC. The helpdesk can be implemented as a distributed platform linking together the helpdesks of the suppliers offering resources to EOSC. The linking of existing helpdesks allows streamlining of support processes involving multiple suppliers, and in particular facilitates the work of the support teams that, through linking, are able to use existing in-house helpdesk tools.
Monitoring	Monitoring provides the capability of checking the status of service end-point interfaces and aggregating such information for the production of service reports. In particular, it should provide a scalable framework for monitoring status, availability and reliability. It provides monitoring of services, visualisation of their status, dashboard interfacing, notification system and generation of availability and reliability reports. Third parties can gather monitoring data from the system through a complete API. **Use case.** Monitoring information supports Service Report Management, and is consumed to produce Service Reports, i.e. the documents that provide the details of the performance of a service against the service targets defined in service level agreements (SLAs) – often based on key performance indicators (KPIs). Typical users are the EOSC service suppliers.
Accounting	Accounting is about collecting, aggregating, storing and displaying EOSC resource usage data produced by the providers participating in EOSC, for example from the providers of Shared Resources. It gathers usage information from the individual resource providers and aggregates it centrally in a secure, GDPR-compliant manner. Accounting is necessary for providing control over resource consumption by the funders, and reduces the overhead of defining accounting information models, architecture and setup. Accounting is a key service of the EOSC federating core that will support its business models, and provides transparency on which resources are being used. The correlation of

	usage data to service identifiers, scientific product identifiers and user identifiers, supports the development of metrics that relate scientific impact to the extent a researcher and/or project has been embracing open science practices. **Use case.** Accounting of resource usage is required for any EOSC customers (e.g. platform operator and research infrastructure managers) to get aggregated information on usage of scientific products and services used from the EOSC portfolio, to scale up the in-house.
CMDB	The configuration database is an ITIL database used by an organisation to store information about hardware and software assets (commonly referred to as Configuration Items). This database acts as a data warehouse for the organisation and also stores information regarding the relationship between its assets. The CMDB provides a means of understanding the organisation's critical assets and their relationships. **Use case.** The availability of an EOSC CMDB is relevant to EOSC shared resource suppliers, and is requested by the IT configuration management process. It allows the management of the provision of services owned and managed by the EOSC governance. It is envisaged that the management of resources published in EOSC just for the purpose of improving their discoverability, will be delegated to the respective providers and will not be registered in an EOSC CMDB.
Order management	The Order management is a tool that allows to handle orders received through the EOSC Portal. It implements interfaces towards service providers order management tool in case of orders that should not be centrally processed in EOSC. **Use case.** Managing orders from the EOSC Portal.
Operations Portal	The Operations Portal refers to the set of control dashboards that support the work of EOSC infrastructure managers in charge of supervising the overall status, allocation and accessibility of the EOSC shared resources. It provides central operations management of federated resources. The Operations Portal offers a portfolio of management tools to support communications, customer relationship management, infrastructure oversight, and metrics gathering. **Use case.** The Operations Portal can support multiple service management activities like incident management and order management if used as a back-office tool of the EOSC Portal.
Service Portfolio Management	The Service Portfolio management tool allows lifecycle management of the resources provided by EOSC. In particular, it aims at facilitating service management in IT service provision, including federated scenarios. SPMT represents a complete list of the services managed by a service provider; some of these services are visible to the customers, while others are internal. The service management system is designed to be compatible with the FitSM standard. **Use case.** The tool is used by the team involved in the activities of the Service Portfolio Management process, which defines and maintains a service portfolio. A service portfolio is the entity that provides information such as the service value proposition, target customer base, service description, relevant technical specifications, cost and price, risks to the service provider, service level packages offered, etc.

Collaborations software & platforms	Issue management and documentation co-development and sharing. **Use case.** Collaborations between EOSC users and/or service providers.
Security monitoring	Provide features to secure monitoring the EOSC services and resources. **Use case.** Identify security threats in the EOSC.
Messaging	A real-time messaging service that allows to send and receive messages between independent applications. **Use case.** Enabling asynchronous communication between EOSC services.
Software quality assurance	A tool that allows to deliver quality software for the EOSC consumption. The software is compiled, validated and distributed following the Software Provisioning Process (SWPP), where the Quality Criteria (QC) definition sets the minimum quality requirements for acceptance. The growing number of software components currently existing to support EOSC infrastructure favours the adoption of automated solutions towards the manual-based validation mechanisms. **Use case.** Automated validation of software quality.

Technical specifications for all these macro-features are under preparation and will be published for feedback as soon as they are ready. An example of an already mature specification is the AAI technical specification that is described later in the document.

The EOSC Portal is as special case within this block. It is currently being further enhanced and developed by a large collaboration that includes EOSC-hub, OpenAIRE Advance and key partners of the eInfraCentral project. More information is available in the EOSC Portal concept paper[4]. The outcomes of this collaboration will be adopted by this work to technically specify the EOSC Portal.

4.3 EOSC Common services

In the Common services block, we intend as macro-feature a technical function that offers added value on top of the EOSC resources (computing, storage, etc) and that can be adopted by multiple thematic services. Examples of macro-features for this block are IaaS VM/Container management, Cloud Orchestration, metadata management, making scientific artefacts FAIR, etc. A macro-feature in the Common Service block can be implemented and, then, offered by one or more common services.

In this block, the number of relevant macro-features can be huge, this required to split the work on sub-areas. We used the technical areas we are working on in EOSC-hub to start the process to identify the macro-features:

- HTC/HPC Compute
- Cloud Compute (inc Containerisation and orchestration)
- PaaS Solutions
- Data Platforms for Processing
- Data Publishing and Open Data
- Data Preservation/Curation/Provenance

[4] https://wiki.eosc-hub.eu/display/EOSC/EOSC+Portal

- Metadata Management and Data Discovery
- Workflow management and user interfaces and Data analytics

Other technical areas could be added by other initiatives according to their expertise. For example, OpenAIRE suggested to add the Scholarly Communication technical area and proposed macro-features for this area during the last EOSC-hub technical workshop in Amsterdam[5].

Also for the Common Services, we agreed to prioritise the preparation of the technical specifications for the macro-features that are more relevant for users according to the use cases analysis. The current list of identified macro-features, organised per technical area, is listed in the following table.

Table 2. Common services macro-features per technical area

Macro-features	***Short description***
HTC/HPC Compute	• Multitenant job submission • Multitenant container based job submission • HTC / HPC clusters on demand
Cloud Compute (inc Containerisation and orchestration)	• IaaS: VM Management • IaaS: Orchestration • IaaS: Containers
PaaS Solutions	• PaaS Solution for Cloud service automation and federation of hybrid Cloud resources
Data Platforms for Processing	• Transparent data processing using POSIX in distributed and hybrid cloud environment including Dockers and Kubernetes and Jupiter • Data Ingesting and movement for processing in hybrid cloud environment • Metadata Management in processing workflows • QoS based data access optimization and tight integration with preservation services • Authorization based on attributes from IdP • Results sharing and experiment repeatability • Distribution of software for the processing tasks
Data Publishing and Open Data	• Data Repository
Data Preservation/Curation/Provenance	• Data Preservation • Tracking of provenance metadata • Data Curation
Metadata Management and Data Discovery	• Data Discovery and Access • Metadata cataloguing and indexing • Annotation service
Workflow management and user interfaces and Data analytics	• Portals • Big data analytics • ML/DL analytics services • Cloud based IoT Platforms interoperability
Scholarly Communication	• Data Management Plans • Digital Preservation

[5] https://indico.egi.eu/indico/event/4675/overview

	• Overlay platforms: Peer-review • Anonymization • Aggregator • Broker • Entity Registry • Metadata validation • Annotation • Usage stats • VRE: RI Services for experiments

An example of an already completed technical specification for the Common Services block is presented in Annex. It details the Cloud IaaS VM Management macro-feature.

The following picture shows how the EOSC technical architecture will appear when the first set of macro-features and the related technical specifications of this block will be well defined. Thematic services could easily exploit macro-features offered in the common services block through the EOSC standard interfaces (purple arrows in the figure). Also services offering such macro-features in the Common Service block can be made interoperable in an easier way thanks to the EOSC standard interface (red arrows in the figure) offering a combined usage to the thematic services. In this scenario, the service composability would be easier to obtain and the cost of integration works will be reduced with respect to the current one.

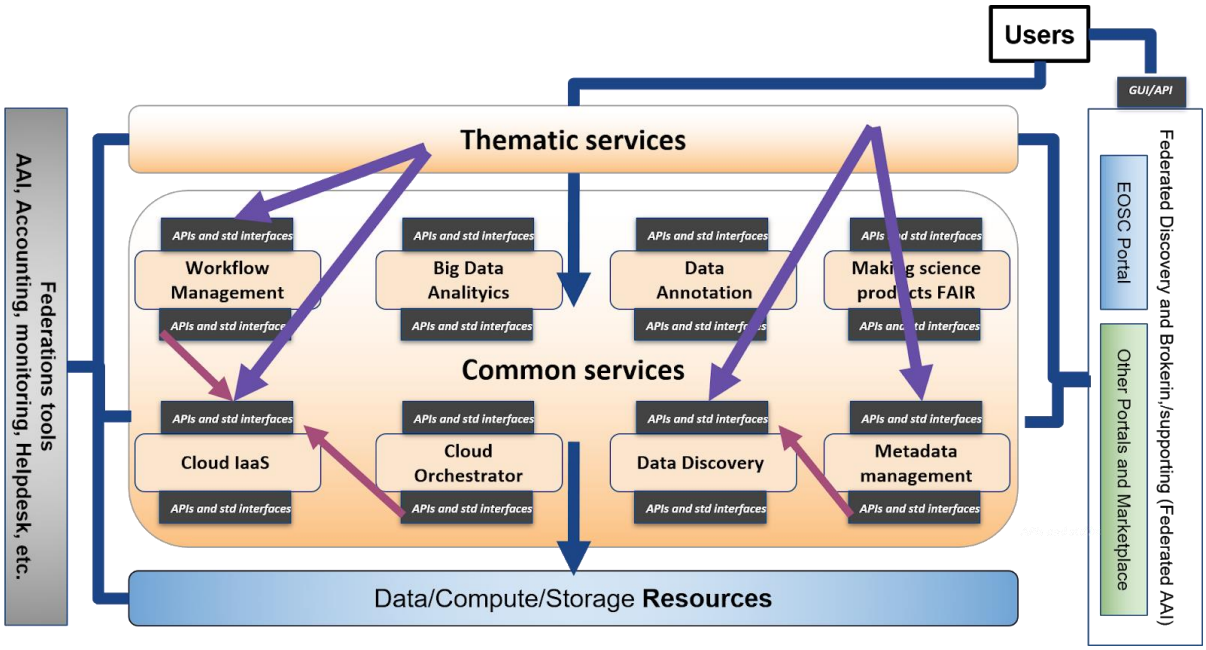

Figure 3. EOSC Technical Architecture. Interactions between thematic and common services.

4.4 EOSC Thematic Services

As we previously wrote, we want to apply the same process to identify and technical specify macro-features also to the Thematic services. In this block, **we intend as macro feature a technical function that is discipline oriented and that can be reused in multiple thematic services**.

Discipline oriented macro features need to be identified and specified by experts of the related disciplines. Then, EOSC-hub will start the work to detail this block with the communities participating in the project. However, community oriented projects need to be involved to further enhance this activity.

The following picture shows how the EOSC technical architecture will appear when the first set of macro-features (and the related technical specifications) for the Thematic Services block will be identified.

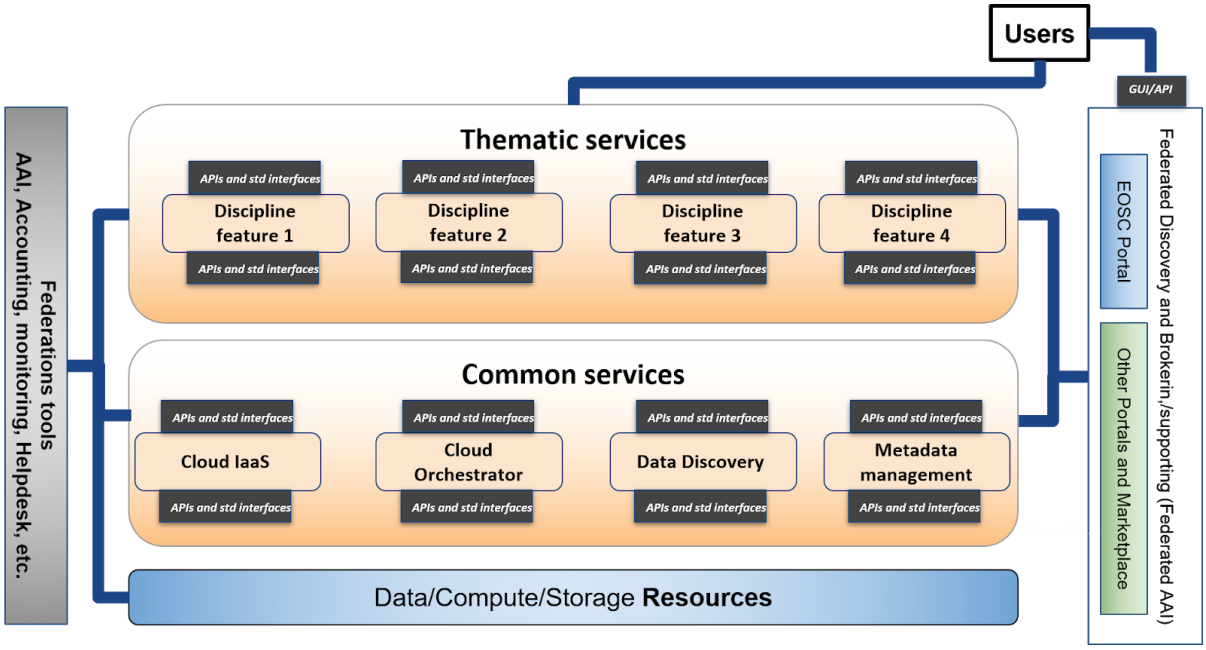

Figure 4. EOSC Technical Architecture. Macro-features in the thematic services block.

5 Relationship with the EOSC Architecture Working Group objectives

One of the working groups that has been setup by the EOSC Governance is about the EOSC Architecture. Its main objective is to *review of the current offering and the required evolution of the EOSC technical architecture, its standards and best practices*. It has been translated into five sub-objectives listed in the table below.

The output of EOSC-hub activity on the EOSC Tech Architecture is intended to become an important input for the activity of this working group and it has been shaped to try to provide valuable results for each of the five sub-objectives. The following table shows the mapping between the Architecture WG sub-objectives and what we expect to deliver within this activity.

Table 3. Mapping between EOSC Architecture WG objectives and the outcome of the EOSC-hub activity on the EOSC technical architecture

EOSC Architecture WG objectives - The WG will describe and/or define:	***Outcome of the EOSC-hub activity on the EOSC technical architecture***
EOSC core services and their interfaces	• Definition of the EOSC Access Enabling and Federation services and interfaces
EOSC open source APIs for reuse by thematic services	• Interoperability guidelines for Common services (EOSC APIs and standards) • Interoperability guidelines for Thematic services (EOSC APIs and standards)
EOSC portal components and federated catalogues of service offerings	• Outcomes of the collaboration between EOSC-hub, OpenAIRE and key partners from eInfraCentral on the EOSC Portal design and development
The EOSC data description standards	• To be described in the technical specification of the metadata management macro-feature
Standards and best practices necessary to ensure the evolution of EOSC and the widening of its user base to the industry and the public sectors	• Interoperability guidelines for Common services (EOSC APIs and standards) • Interoperability guidelines for Thematic services (EOSC APIs and standards)

Appendix I. AAI Technical Specification

Introduction

The EOSC AAI enables seamless access to research data and services in EOSC in a secure and user-friendly way.

High-level Service Architecture

The EOSC AAI follows the architectural and policy recommendations defined in the AARC project [AARC-Community]. As such, it enables interoperability across different SP-IdP-Proxy services, each of which acts as a bridge between the community-managed proxies (termed Community AAIs) managing the researchers' identity and the generic services offered by Research Infrastructure and e-Infrastructures (termed R/e-Infrastructures or Infrastructures). This is the “community-first” approach to the AARC Blueprint Architecture [AARC-G045], which enables researchers to sign in with their community identity via their Community AAI. A high-level view of the EOSC AAI is provided in the figure below.

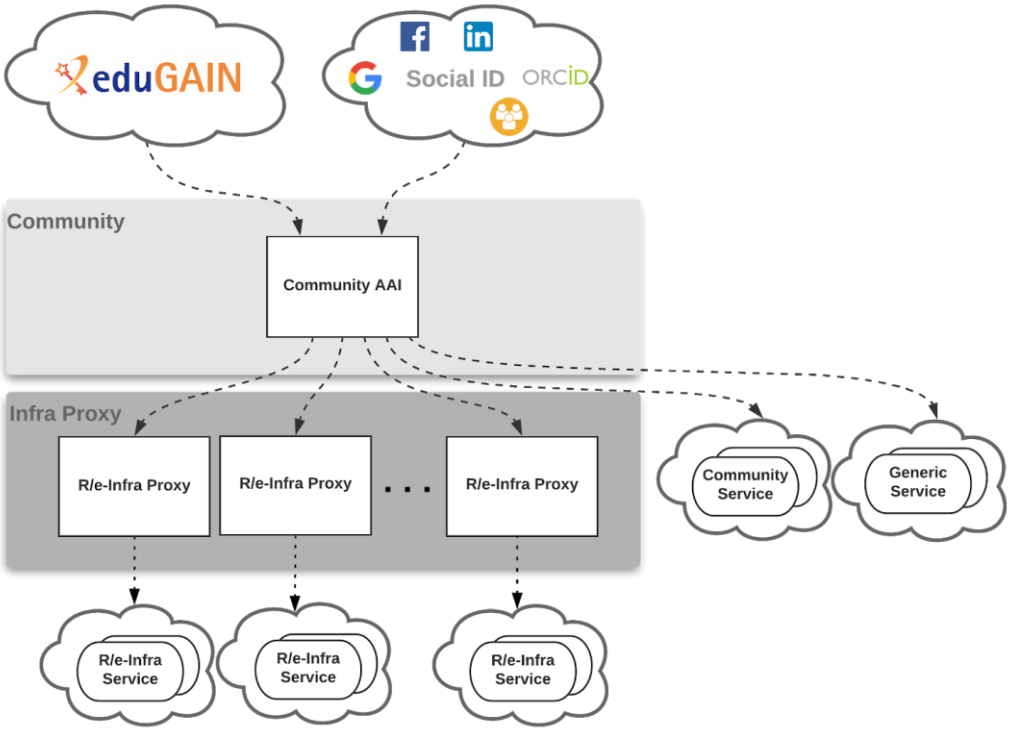

Figure 5. AARC Blueprint Architecture.

Community-specific services are connected to a single Community AAI, while Infrastructure Services are connected to a single Infrastructure Proxy. Lastly, generic services may be connected to more than one Community AAI. Each Community AAI in turn serves as a bridge between external identity providers and the proxies to the e-infrastructure services. Specifically, Community AAIs connect to eduGAIN as service providers but act as identity providers from the services point of view, thereby allowing users to use their credentials from their home organisations. Complementary to this, users

without an account on a federated institutional Identity Provider are still able to use social media or other external authentication providers for accessing services.
Research communities can leverage the EOSC AAI services for managing their users and their respective roles and other authorisation-related information. At the same time, the adoption of standards and open technologies, including SAML 2.0, OpenID Connect, OAuth 2.0 and X.509v3, facilitates interoperability and integration with the existing AAIs of other e-Infrastructures and research communities.

Adopted standards

Standard	Short description	References
Security Assertion Markup Language (SAML) 2.0	OASIS standard for exchanging authentication and authorisation data between parties.	https://www.oasis-open.org/standards#samlv2.0
OAuth 2.0	Standard for authorisation that enables delegated access to server resources on behalf of a resource owner	"The OAuth 2.0 Authorization Framework", RFC 6749, https://www.rfc-editor.org/info/rfc6749
OpenID Connect 1.0	Identity layer on top of the OAuth 2.0 protocol. It enables Clients to verify the identity of the End-User based on the authentication performed by an Authorization Server, as well as to obtain basic profile information about the End-User in an interoperable and REST-like manner	“OpenID Connect Core 1.0”, https://openid.net/specs/openid-connect-core-1_0.html
X.509	ITU-T standard for a public key infrastructure (PKI), also known as PKIX (PKI X509)	"Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 5280, https://www.rfc-editor.org/info/rfc5280 "Internet X.509 Public Key Infrastructure (PKI) Proxy Certificate Profile", RFC 3820, https://www.rfc-editor.org/info/rfc3820
Lightweight Directory Access Protocol (LDAP)	Provides access to distributed directory services that act in accordance with X.500 data and service models.	https://tools.ietf.org/html/rfc4511

Protocol/API	Short description	References
OAuth 2.0 Token Introspection	Protocol that allows authorised protected resources to query the authorisation server for determining the set of metadata for a given OAuth2 token, including its current validity.	https://tools.ietf.org/html/rfc7662
OAuth 2.0 Token Exchange	Protocol for requesting and obtaining security tokens from OAuth 2.0 authorization servers, including security tokens employing impersonation and delegation.	https://tools.ietf.org/id/draft-ietf-oauth-token-exchange-14.html
OAuth 2.0 Device Authorization Grant	Enables OAuth 2.0 clients on input-constrained devices to obtain user authorisation for accessing protected resources without using an on-device user-agent.	https://tools.ietf.org/html/draft-ietf-oauth-device-flow-15
System for Cross-domain Identity Management (SCIM) 2.0	Open API for managing identities	SCIM: Core Schema , RFC7643, https://tools.ietf.org/html/rfc7643 SCIM: Protocol, RFC7644, https://tools.ietf.org/html/rfc7644 SCIM: Definitions, Overview, Concepts, and Requirements, RFC7642, https://tools.ietf.org/html/rfc7642

Interoperability guidelines

Technical interoperability guidelines

- The attributes used to express user information should follow the REFEDS R&S attribute bundle, as defined in [REFEDS-R&S]
- VO/group membership and role information, which is typically used by relying parties for authorisation purposes, should be expressed according to [AARC-G002]
- Capabilities, which define the resources or child-resources a user is allowed to access, should be expressed according to [AARC-G027]
- Affiliation information, including (i) the user's affiliation within their Home Organisation, such as a university, research institution or private company, and (ii) affiliation within the Community, such as cross-organisation collaborations, should be expressed according to [AARC-G025]
- Assurance information used to express how much relying parties can trust the attribute assertions about the authenticating user should follow:
    - REFEDS Assurance framework (RAF) [RAF-version-1.0]

  - Guideline on the exchange of specific assurance information [AARC-G021]
  - Guideline for evaluating the combined assurance of linked identities [AARC-G031]
  - Guideline Expression of REFEDS RAF assurance components for identities derived from social media accounts [AARC-GO41]
  - Guidelines for expressing the freshness of affiliation information, as defined in [AARC-G025]
- OAuth2 Authorisation servers should be able to validate tokens issued by other trusted Authorisaton servers. Extending existing flows, such as the OAuth2 Token Exchange flow [OAuth2-Token-Exchange-draft], will need to be considered for enabling the validation of such externally issued tokens.

Policy interoperability guidelines

- For the EOSC AAI, compliance with the GÉANT Data Protection Code of Conduct version 1 (DPCoCo-v1) [DPCoCo-v1] is implicit, since it reflects the Data Protection Directive and means compliance with applicable European rules (see [AARC-G040]). To explicitly declare compliance with DPCoCo-v1, the privacy notice of each EOSC AAI service should include a reference to DPCoCo-v1.
- The entities of the EOSC AAI registered with eduGAIN should meet the Sirtfi [Sirtfi-v1.0] requirements and express Sirtfi compliance in their metadata in order to facilitate coordinated response to security incidents across organisational boundaries.
- To reduce the burden on the users and increase the likelihood that they will read the AUP as they access resources from multiple service and resource providers, the EOSC AAI services should adopt the WISE Baseline AUP model [WISE-AUP].

Examples of solutions implementing this specification

AAI services:

- B2ACCESS
- Check-in
- eduTEAMS
- INDIGO-IAM

Identity and Access Management:

- Perun
- Comanage
- HEXAA

Token Translation Services:

- WaTTS
- MasterPortal
- RCauth.eu

Procedure to integrate a service with the EOSC Hub AAI

- B2ACCESS
- Check-in
- eduTEAMS
- INDIGO-IAM
- Perun
- WaTTS
- MasterPortal
- RCauth.eu

Appendix II. Cloud IaaS VM Management Technical Specification

Introduction

Services of Cloud IaaS VM Management provide on-demand API-based access to computing resources as Virtual Machines that can run user-defined arbitrary software (including operating systems and applications). Services in this category also allow management of block storage that can be associated to the VMs and network management to provide connectivity between VMs and external networks.

High-level Service Architecture

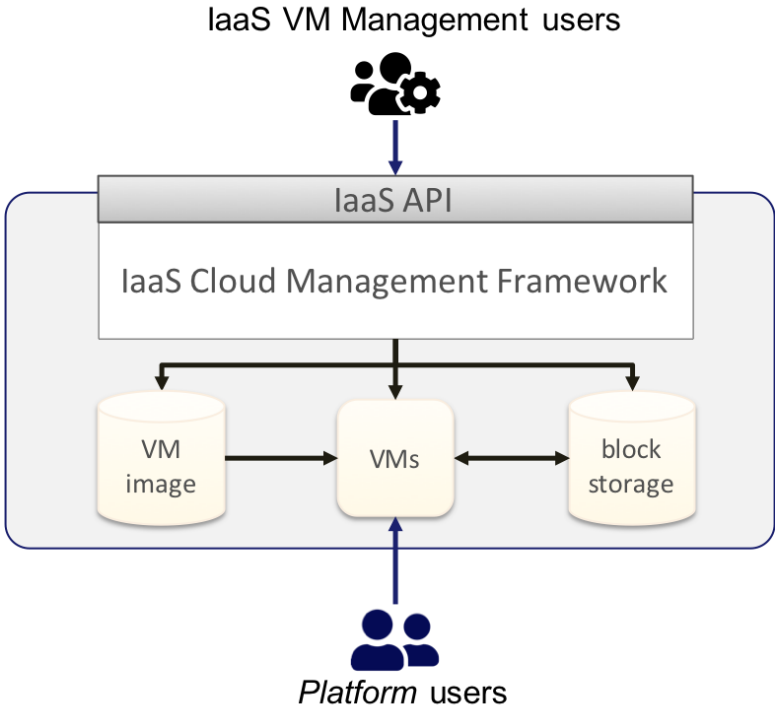

Figure 6. Cloud IaaS VM Management. High level architecture.

IaaS VM Management services allow users to manage VMs that are instantiated from VM images and can be associated with permanent block storage. The VMs can execute any kind of workload, including new services or platforms that are accessed by platform users, which may be different from the IaaS VM Management users that manage the IaaS resources.

Adopted standards

Standard	Short description	References
Open Virtualization Format (OVF)	Packaging format for software solutions based on virtual systems (VM image format)	OVF 2.1.1

Protocol/API	Short description	References
OpenStack	OpenStack is an Open Source cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter, all managed and provisioned through APIs with common authentication mechanisms.	OpenStack API
Amazon EC2/EBS/VPS & AWS VPN	Amazon Elastic Compute Cloud (EC2), Elastic Block Storage (EBS), Virtual Private Cloud (VPS) and AWS Virtual Private Network (AWS VPN) provide management of Virtual Machines and associated block storage and network features	AWS EC2 API
Azure Virtual Machines/Disks/VNet	IaaS VM management services from Microsoft Azure	Azure Virtual Machines API
Google Cloud Compute Engine	IaaS VM management service from Google Cloud Platform	Google Cloud Compute Engine API
OGF OCCI	Open community-lead specifications delivered through the Open Grid Forum. OCCI is a Protocol and API for all kinds of Management tasks, focused on IaaS	OCCI Specification
DMTF CIMI	The CIMI specification describes the model and protocol for management interactions between a cloud Infrastructure as a Service (IaaS) provider and the consumers of an IaaS service.	CIMI 2.0.0

Interoperability guidelines

Interoperable service in this category must:

- **Provide API access** for on-demand management of VMs and associated resources. Open and/or Standard APIs are preferred. Services that provide the capability to manage VMs through graphical dashboards but limit API access to users cannot be considered interoperable. See table above for a non-comprehensive list of APIs that may be supported by the service.

AAI interoperability

- Services should provide access to users authenticated with one of the EOSC-hub AAI federated identity protocols (OpenID Connect and/or SAML).

Orchestration interoperability

- Services should expose APIs that are supported by the IaaS Orchestrator services of EOSC-hub.

Federation interoperability:

- Services in this category that need to be federated into a cloud federation should provide API-based access to:
    - Management of VM images, i.e. allow to create (upload) and delete VM images from which VMs can be instantiated.
    - Access usage information of individual VMs and block storage so accounting records can be generated for integration into the EOSC-hub central services.

Examples of solutions implementing this specification

EOSC-hub services:

- EGI Cloud Compute

OpenSource implementations:

- OpenStack
- OpenNebula