EOSC Architecture and Interoperability Guidelines

Giacinto Donvito/INFN, EOSC-hub Technical Coordinator and WP10 Leader

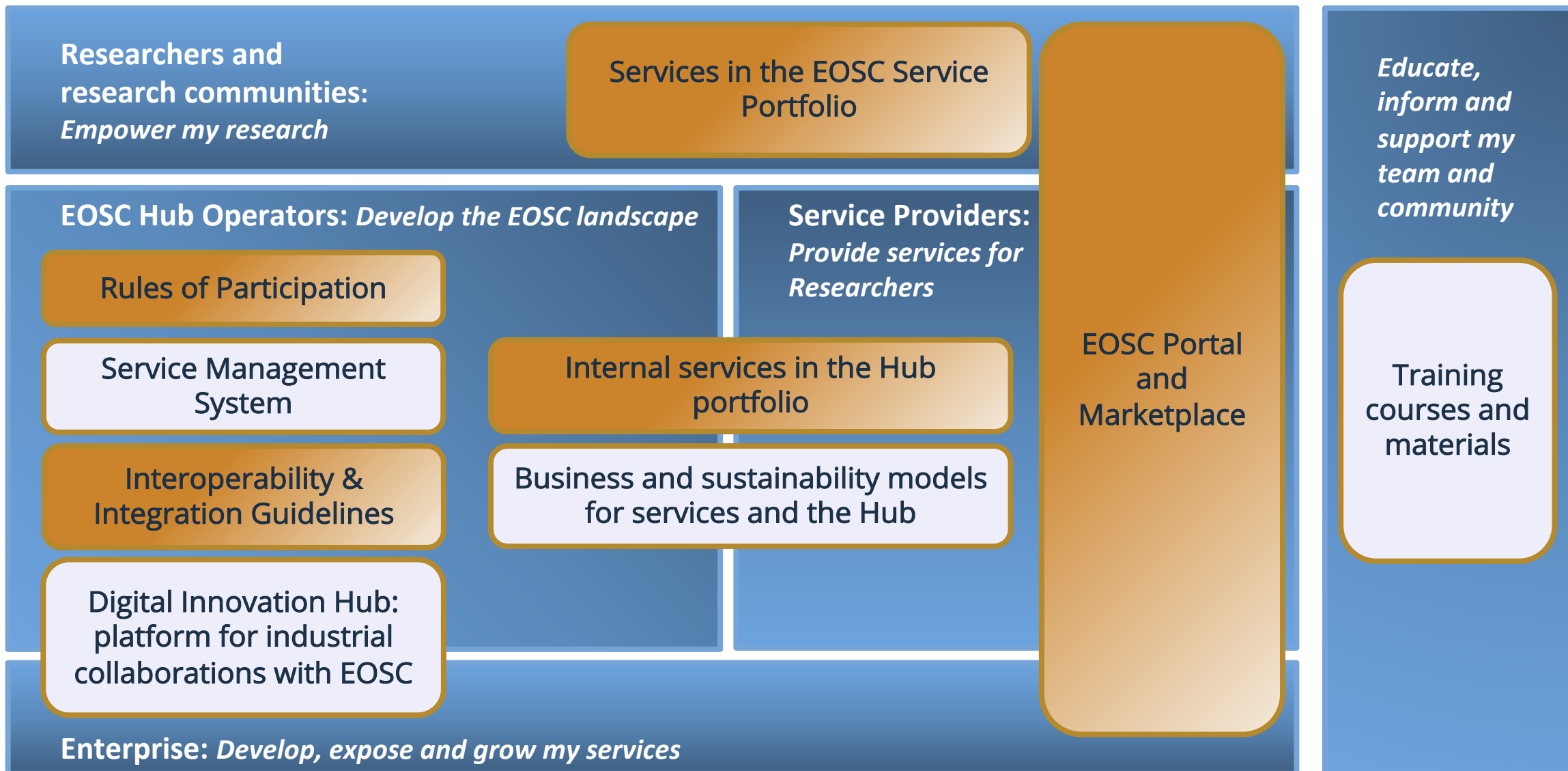**Luxembourg, 8-9 October 2019**

🔗   eosc-hub.eu

🐦   @EOSC_eu

Key Exploitable Results WP10

EOSC-hub

Researchers and research communities: *Empower my research*

Services in the EOSC Service Portfolio

Educate, inform and support my team and community

EOSC Hub Operators: *Develop the EOSC landscape*

Rules of Participation

Service Management System

Interoperability & Integration Guidelines

Digital Innovation Hub: platform for industrial collaborations with EOSC

Service Providers: *Provide services for Researchers*

Internal services in the Hub portfolio

Business and sustainability models for services and the Hub

EOSC Portal and Marketplace

Training courses and materials

Enterprise: *Develop, expose and grow my services*

Landscape and Service Composability

Overarching concepts

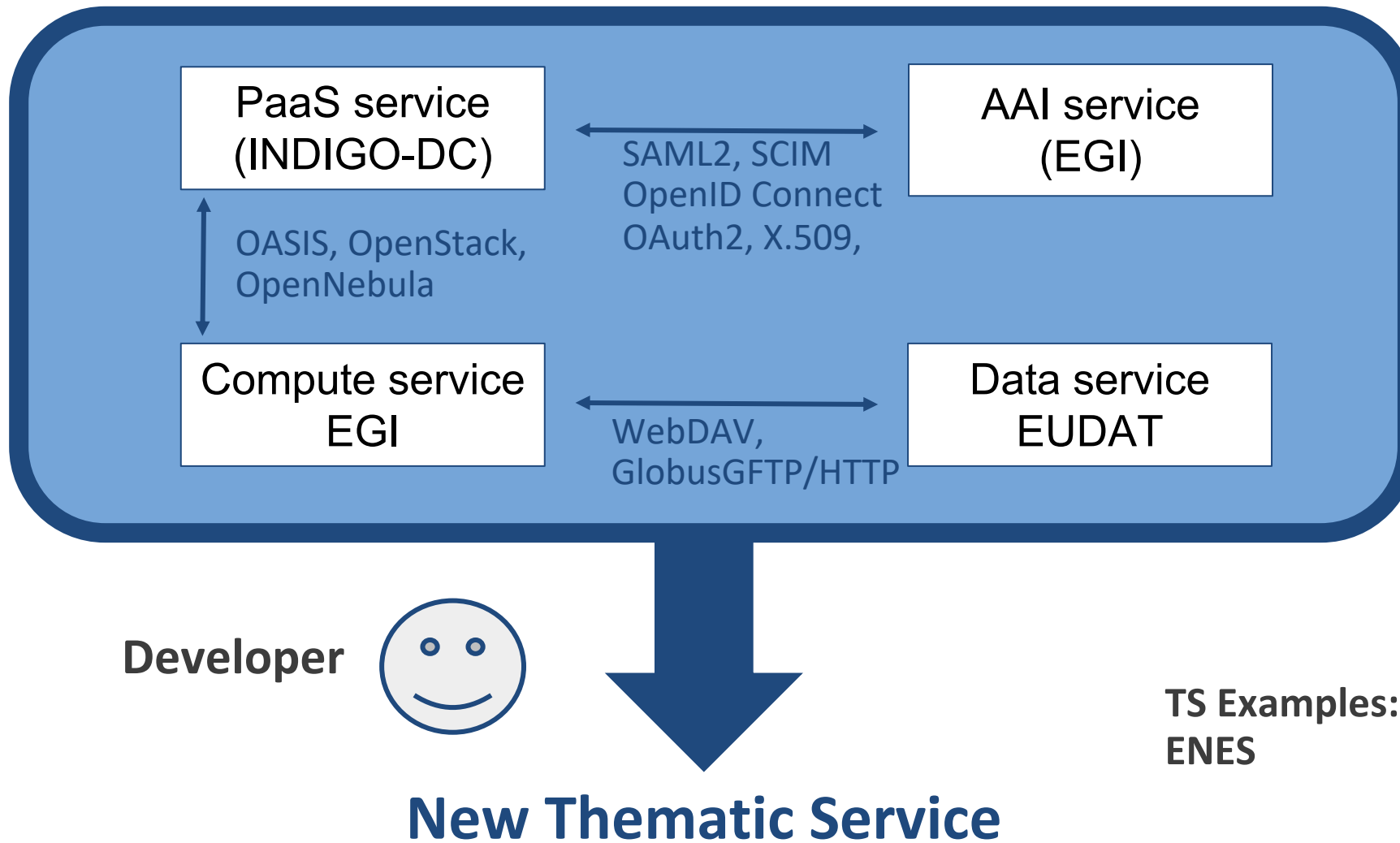Landscape & EOSC Architecture Working Group
End-to-end service composition
EOSC Interoperability Guidelines & Interoperability Framework

Landscape

- Analysis of EOSC Pilot recommendations on Service Architecture
- Commission Staff Working Document on the implementation roadmap for the European Open Science Cloud
  - 'Architecture' Action Line
- EOSC Architecture WG
  - EOSC-hub representative contributing to the WG
  - The WG will propose the technical framework required to enable and sustain an evolving EOSC federation of systems
  - EOSC interoperability layer: standards, APIs and protocols that will facilitate interoperable services
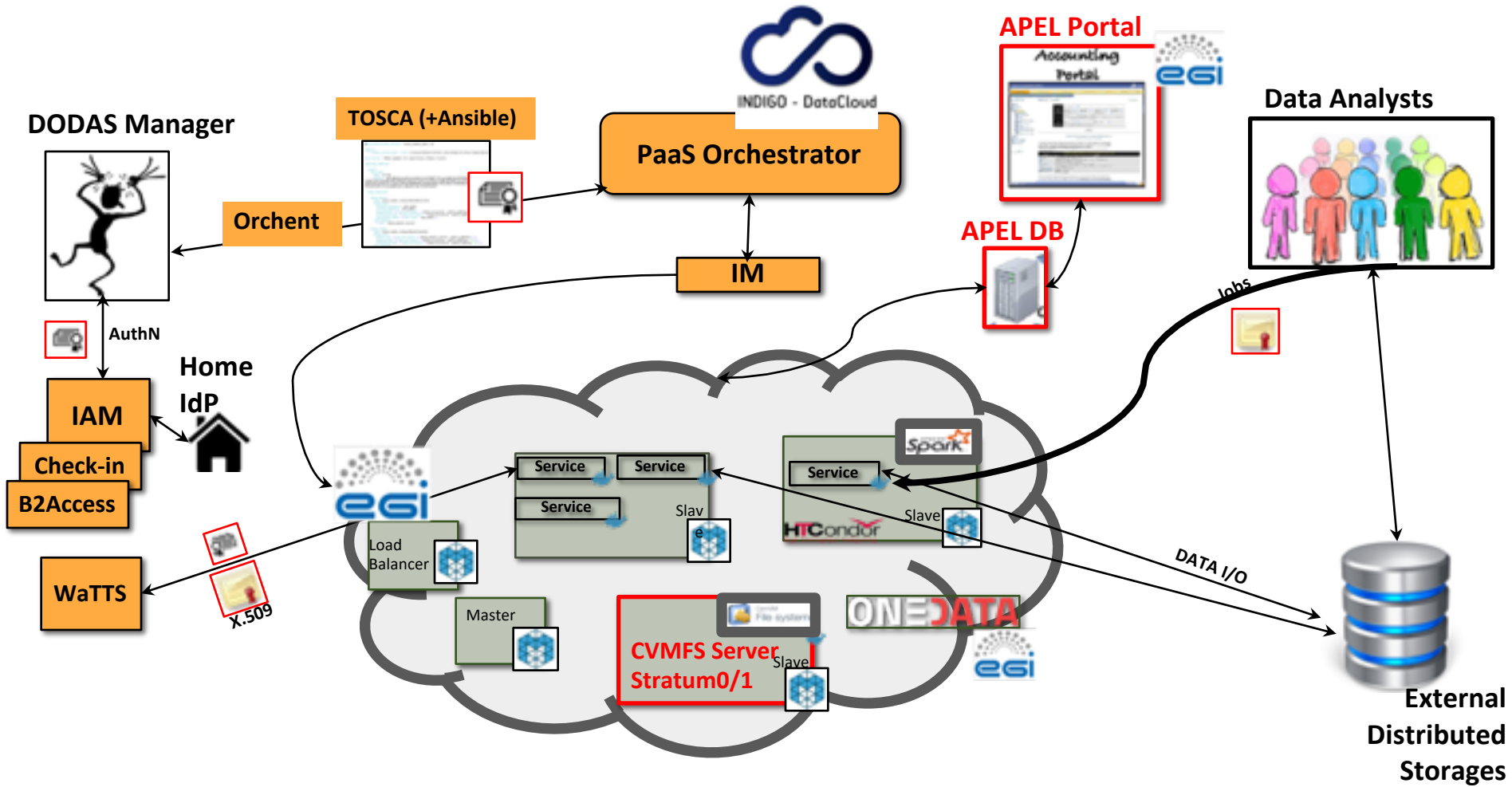
End-to-end Composition of Services

- Fostering service composability to create new added-value solutions for research
  - Promote the adoption of standards
  - Perform service integration activities

- Typical service combinations
  - Re-using federation services to implement basic features (e.g. AAI, monitoring, accounting)
  - Adopt common services to better exploit compute, storage and data resources (e.g. EGI Cloud Compute, EUDAT B2FIND, INDIGO Orchestrator)
  - Create new scientific workflows as combination of more services

EOSC Interoperability Guidelines

- Technical Interoperability Guidelines in terms of suggested EOSC standards and interfaces/APIs
  - Lower the barriers to allow services to interact and work together
- Benefits:
  - Scientists can rely on e-Infrastructures and EOSC-compliant services for implementing the basic features
  - Federated and common services can interoperate (e.g. AAI or accounting services from different initiatives)
- Interoperability guidelines as implementation of the EOSC interoperability layer (--> EOSC Architecture WG)
  - Based on existing community practices, well-known standards and interfaces
  - Adoption should not be forced but as a natural consequence of the advantages for a service
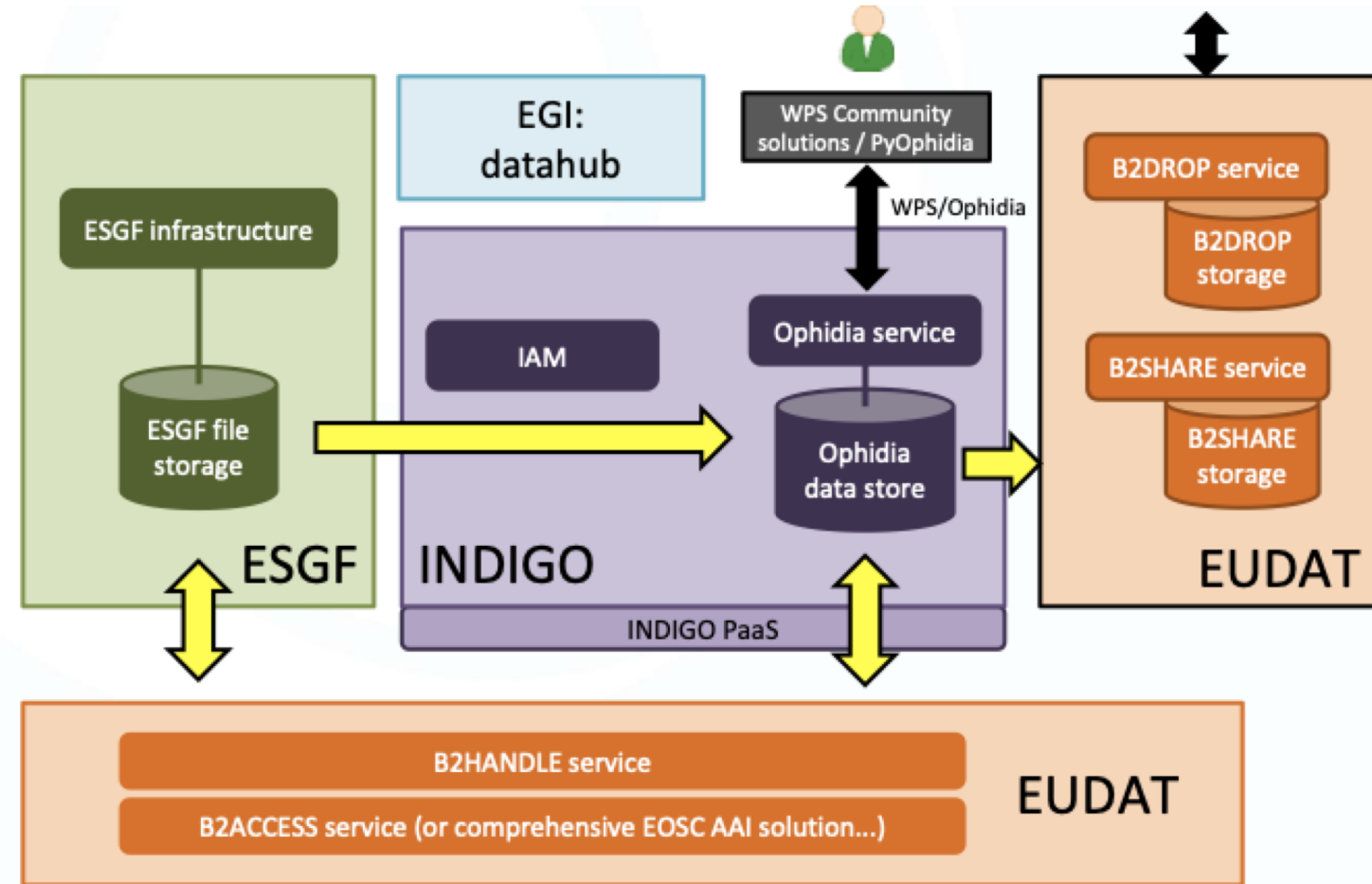
Common Service integration example

EOSC-hub

DODAS Manager

TOSCA (+Ansible)

INDIGO - DataCloud

PaaS Orchestrator

APEL Portal

Accounting Portal

Data Analysts

Orchent

IM

APEL DB

AuthN

Home IdP

IAM

Check-in

B2Access

WaTTS

X.509

Service Service

Service

Spark

Service

Slave

Load Balancer

Slave

HTCondor

Master

CVMFS Server Stratum0/1

ONEDATA

Slave

DATA I/O

Jobs

External Distributed Storages

Thematic Service

TS Examples: DODAS

Common Service integration example

Thematic Service

TS Examples:
ENES

EOSC Technical Architecture

The approach

Reference architecture
Define the EOSC Building Blocks
EOSC Technical Architecture
Access Enabling and Federation Tools Building Blocks
Common service Building Blocks
Thematic services Building Blocks
Technical Roadmap

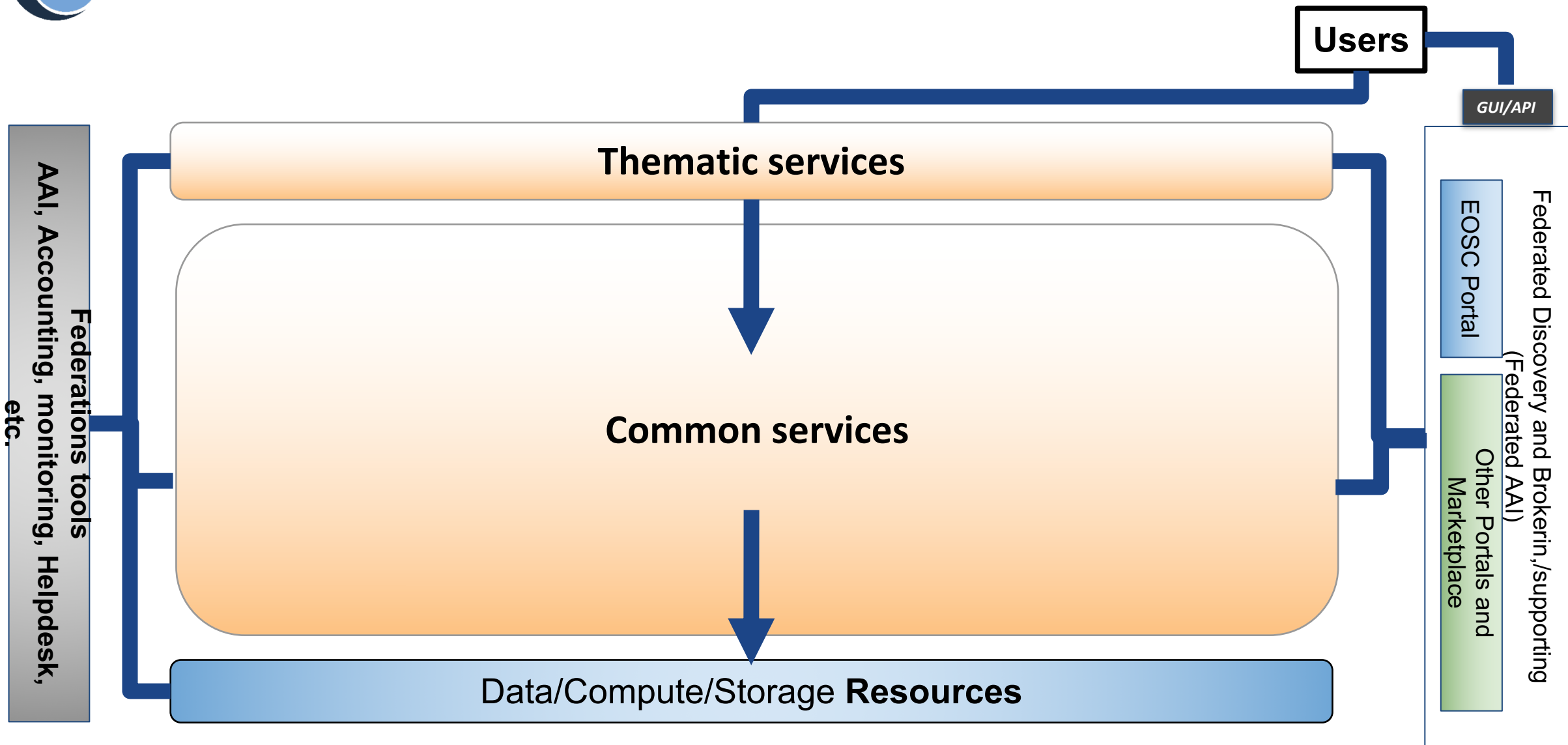# Define the Building Blocks of the EOSC architecture

- Identify and define **EOSC architecture building blocks** for Access Enabling and Federation, Common and Thematic services
  - Common approach for all service categories
  - Mapping the features to **Technical Specifications** including High-level architecture & Interoperability guidelines
- Iterative approach to identify building blocks and improve the technical specification
  - Start from the most relevant use cases
  - Define the technical specification and get feedback
  - Involve external people/initiatives with expertise in the are

Family of EOSC Compliant Services per Block

EOSC-hub
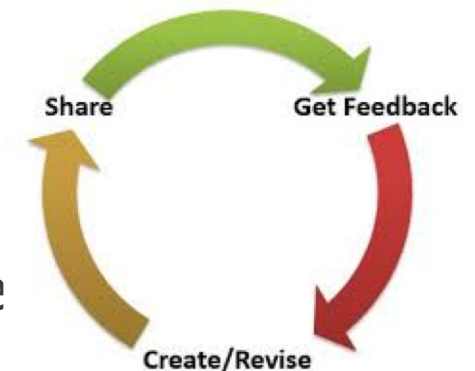
- **Services in a Building Block**
  - Follow the EOSC technical specification
  - Are interoperable
- **Examples:**
  - AAI services compliant with the AARC Blueprint
  - Monitoring/Accounting systems able to exchange/share information

Building Block

Technical specification template per Block

- **Introduction**
  - *Short description of the building block highlighting its main functions*
- **High-level Service Architecture**
  - *Reference architecture of the building block highlighting the interfaces towards the other building blocks*
  - *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**
  - *Services compliant with this specification*

EOSC Technical Architecture - Functional view

Users

GUI/API

Thematic services

Common services

Data/Compute/Storage **Resources**

AAI, Accounting, monitoring, Helpdesk, etc.

Federations tools

Federated Discovery and Brokerin,/supporting (Federated AAI)

EOSC Portal

Other Portals and Marketplace

Access Enabling and Federation Services

Access Enabling and Federation services

- Building block as any key acce... ... to operate the E... ... e.

  - Leveraged o... xperiences from large European e-infra... ucture to identifies the needed s... vices

  - Part of the E... Hub Portfolio

Technical specifications for each of these services

Hub Portfolio
AAI
Helpdesk
Accounting
CMDB
Monitoring
Order Management

| Hub Portfolio |
|---|
| EOSC Portal |
| Operations portal |
| Service Portfolio Management Tool |
| Collaboration software & platforms |
| Messaging |
| Security Monitoring |

An example: AAI

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 …	"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)	https://www.rfc-editor.org/info/rfc5280 https://www.rfc-editor.org/info/rfc3820
		ietf.org/html/rfc4

Technical specification ready for public comments
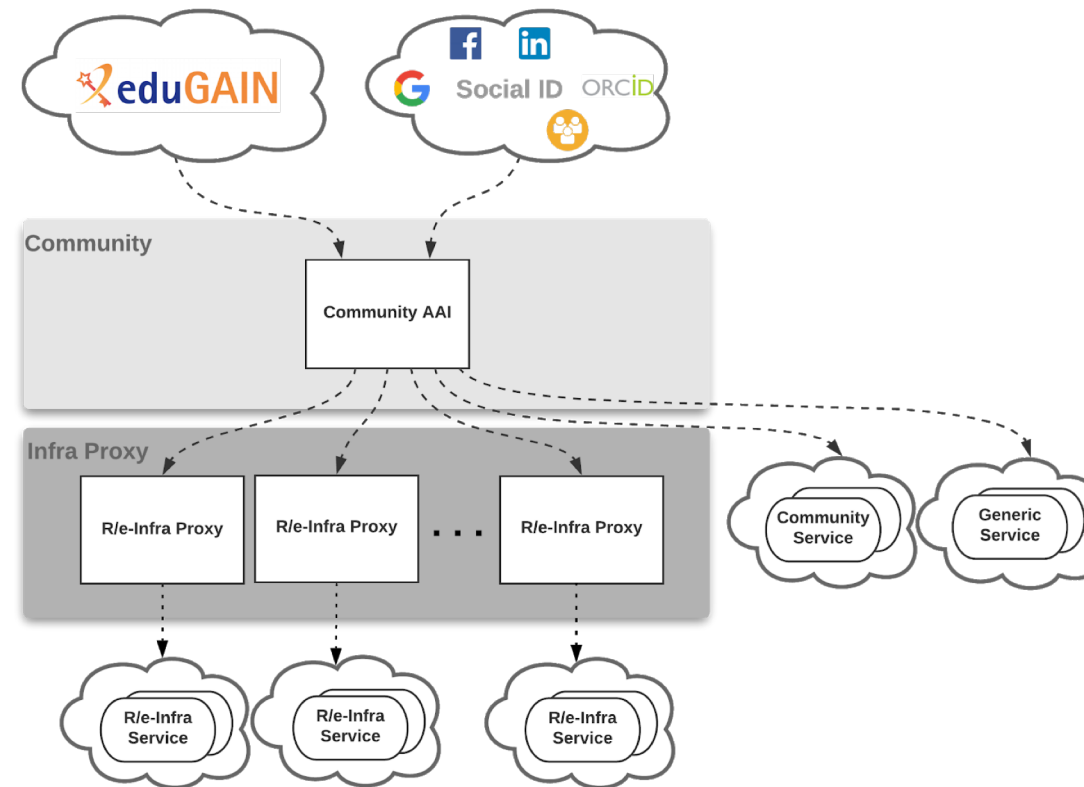
High-level Service Architecture

Interoperability Guidelines

Common Services

Building Blocks in Common Services (1/2)

- In Common Services, a building block is a technical function that
  - offer added value on top of the EOSC resources (computing, storage, etc.)
  - can be adopted by multiple thematic services
- Examples of building blocks
  - IaaS VM/Container management, Cloud Orchestration, metadata management, making scientific artefacts FAIR, etc.
  - A macro-feature can be offered by one or more generic services
- Number of relevant building blocks in Common Services can be huge
  - Different technical areas
  - Need to prioritise
    - Starting from the most relevant for users -> use cases requirements

Building Blocks in Common Services (2/2)

- Split the work leveraging on technical areas
  - EOSC-hub technical areas

Technical Areas
HTC/HPC Compute
Cloud Compute (inc Containerisation and orchestration)
PaaS Solutions
Workflow management and user interfaces and Data analytics

Technical Areas
Data Platforms for Processing
Data Publishing and Open Data
Data Preservation/Curation/Provenance
Metadata Management and Data Discovery

  - Other tech areas suggested by other initiatives
    - E.g. OpenAIRE suggested to add Scholarly Communication

Building Blocks per technical area

Technical Area	Building Blocks
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

Building Blocks per technical area

Technical Area	Building Blocks
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

Building Blocks per technical area

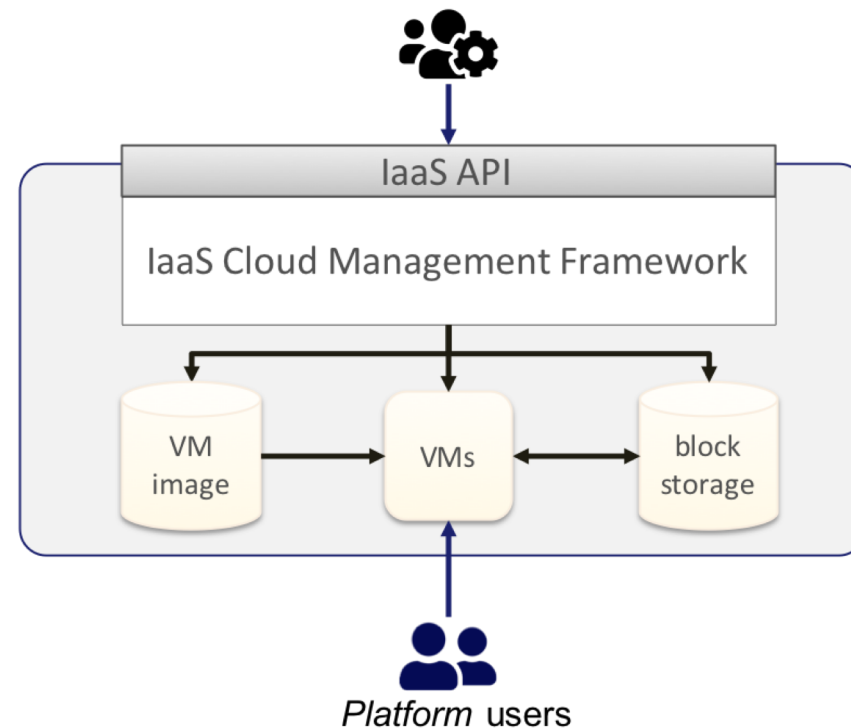| Technical Area	Building Blocks
Workflow management and user interfaces and Data analytics	• Portals • Big data analytics • ML/DL analytics services • Cloud based IoT Platforms interoperability
Scholarly Communication	• Data Management Plans • Digital Preservation • Overlay platforms: Peer-review • Anonymization • Aggregator • Broker • Entity Registry • Metadata validation • Annotation • Usage stats • VRE: RI Services for experiments

An example: Cloud IAAS VM Management

Standards & API

Standard	Short description	References
Open Virtualization Format (OVF)	Packaging format for software solutions based on virtual systems (VM image format)	OVF 2.1.1
OpenStack	OpenStack is an Open Source cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter…	OpenStack API
Amazon EC2/EBS/VPS & 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		Specification

Technical specification ready for public comments

IaaS VM Management users

High-level Service Architecture

Interoperability Guidelines

EOSC Technical Architecture

EOSC-hub

Users

GUI/API

Thematic services

Federations tools
AAI, Accounting, monitoring, Helpdesk, etc.

| APIs and std interfaces | APIs and std interfaces | APIs and std interfaces | APIs and std interfaces |

Workflow Management | **Big Data Analityics** | **Data Annotation** | **Making science products FAIR**

| APIs and std interfaces | APIs and std interfaces | APIs and std interfaces | APIs and std interfaces |

Common services

| APIs and std interfaces | APIs and std interfaces | APIs and std interfaces | APIs and std interfaces |

Cloud IaaS | **Cloud Orchestrator** | **Data Discovery** | **Metadata management**

| APIs and std interfaces | APIs and std interfaces | APIs and std interfaces | APIs and std interfaces |

Data/Compute/Storage **Resources**

EOSC Portal

Other Portals and Marketplace

Federated Discovery and Brokerin./supporting (Federated AAI)

EOSC Technical Architecture

Users

GUI/API

Thematic services

Federations tools
AAI, Accounting, monitoring, Helpdesk, etc.

Federated Discovery and Brokerin,/supporting (Federated AAI)

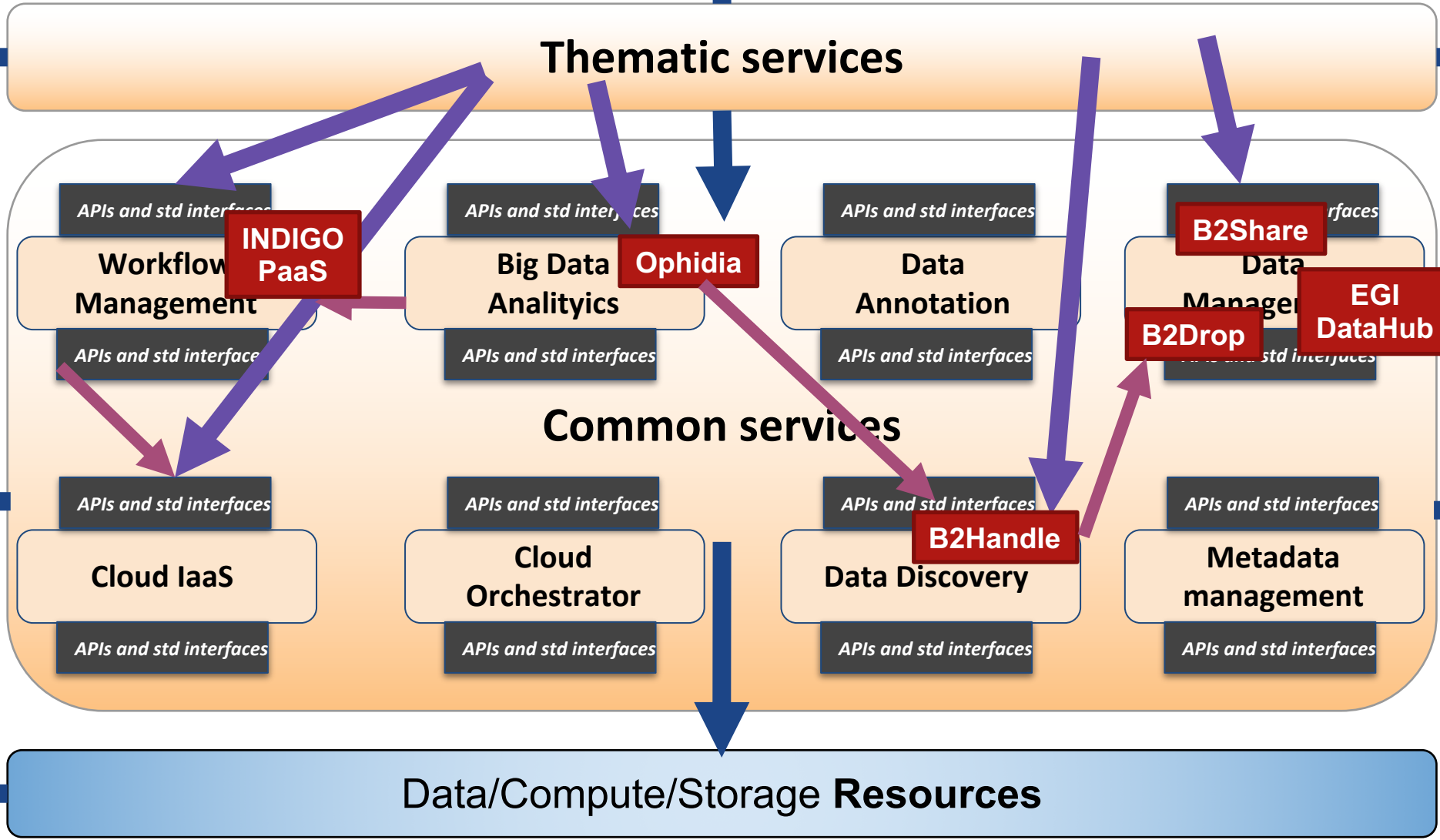EOSC Portal

Other Portals and Marketplace

| APIs and std interfaces | APIs and std interfaces | APIs and std interfaces | APIs and std interfaces |

Workflow Management | **Big Data Analityics** | **Data Annotation** | **Making science products FAIR**

| APIs and std interfaces | APIs and std interfaces | APIs and std interfaces | APIs and std interfaces |

Common services

| APIs and std interfaces | APIs and std interfaces | APIs and std interfaces | APIs and std interfaces |

Cloud IaaS | **Cloud Orchestrator** | **Data Discovery** | **Metadata management**

| APIs and std interfaces | APIs and std interfaces | APIs and std interfaces | APIs and std interfaces |

Data/Compute/Storage Resources
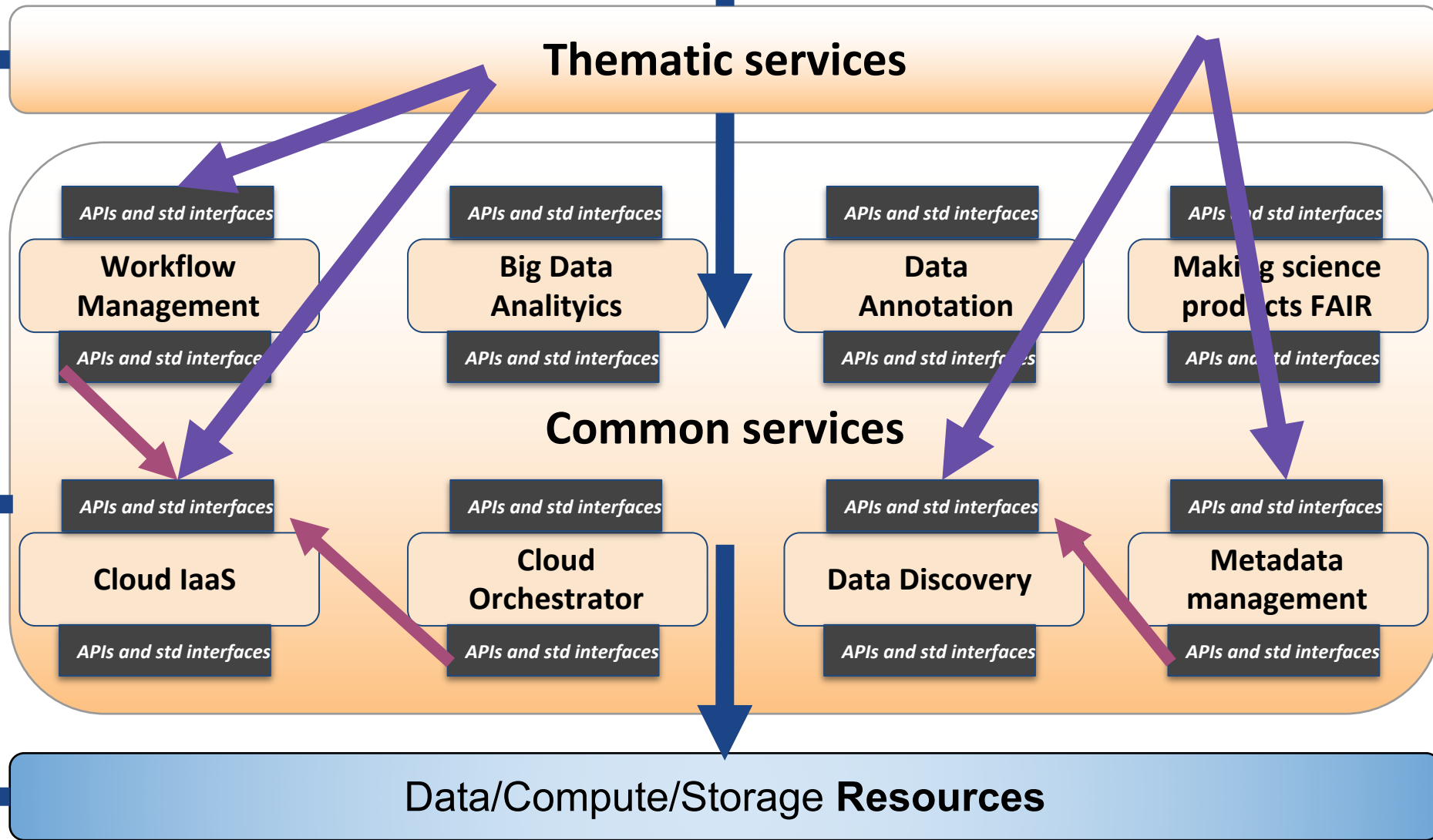
EOSC Technical Architecture: ENES TS

Users

GUI/API

Thematic services

AAI, Accounting, monitoring, Helpdesk, etc.

Federations tools

APIs and std interfaces

INDIGO PaaS

Workflow Management

APIs and std interfaces

Ophidia

Big Data Analityics

APIs and std interfaces

Data Annotation

APIs and std interfaces

B2Share

Data Management

B2Drop

EGI DataHub

APIs and std interfaces

Common services

APIs and std interfaces

Cloud IaaS

APIs and std interfaces

APIs and std interfaces

Cloud Orchestrator

APIs and std interfaces

APIs and std interfaces

B2Handle

Data Discovery

APIs and std interfaces

APIs and std interfaces

Metadata management

APIs and std interfaces

Data/Compute/Storage Resources

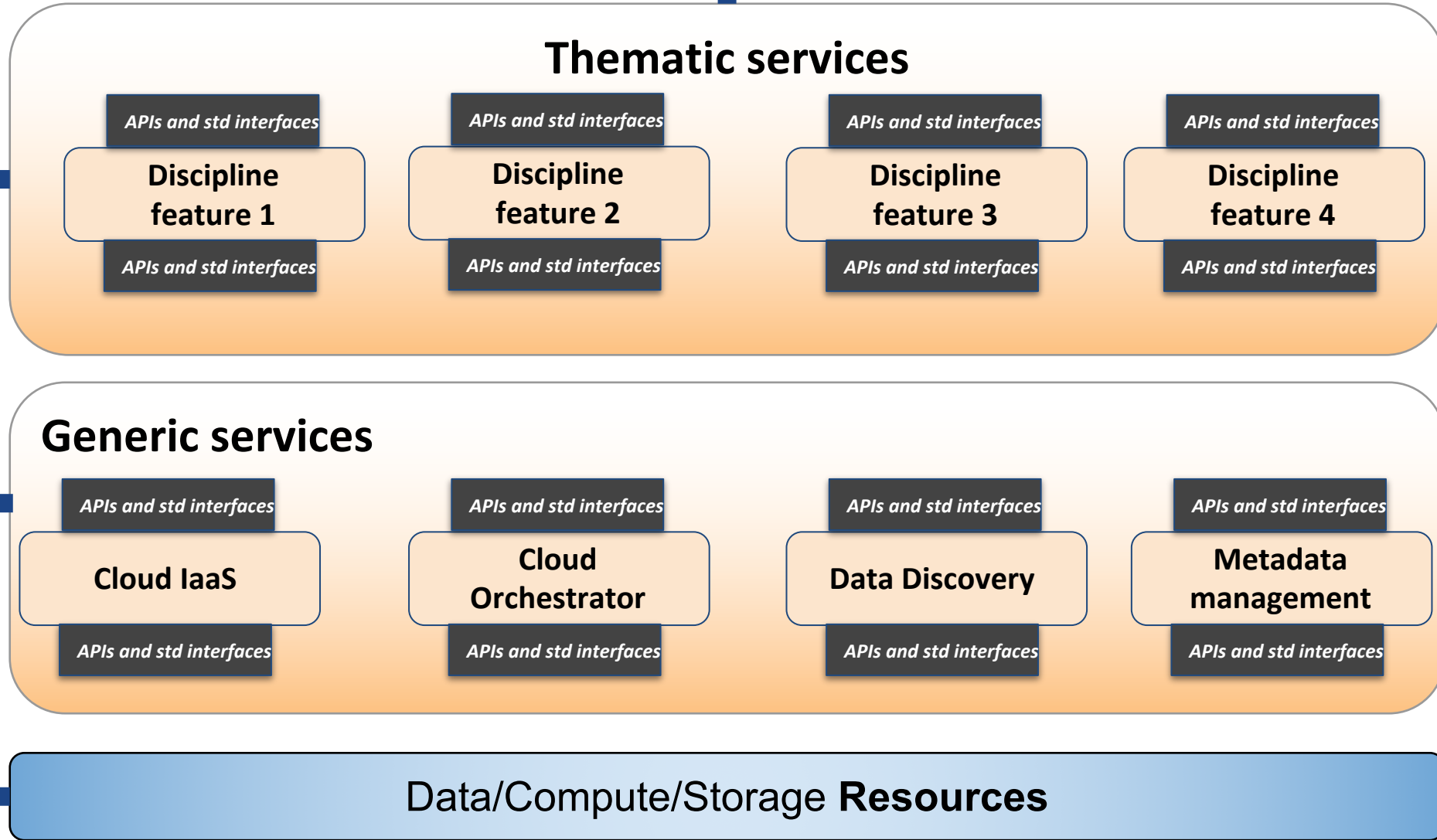Federated Discovery and Brokerin,/supporting (Federated AAI)

EOSC Portal

Other Portals and Marketplace

Thematic Services

Thematic Services

- A building block in the Thematic Services is a technical function that
  - is discipline oriented
  - can be re-used in multiple thematic services within one thematic domain

- Discipline oriented building blocks need to be identified and specified by experts of disciplines
- EOSC-hub will start this effort with communities participating to the project
  - Community oriented projects need to be involved (e.g. Cluster projects)

EOSC Technical Architecture

Users

GUI/API

Thematic services

APIs and std interfaces	APIs and std interfaces	APIs and std interfaces	APIs and std interfaces
Discipline feature 1	Discipline feature 2	Discipline feature 3	Discipline feature 4
APIs and std interfaces	APIs and std interfaces	APIs and std interfaces	APIs and std interfaces

Generic services

APIs and std interfaces	APIs and std interfaces	APIs and std interfaces	APIs and std interfaces
Cloud IaaS	Cloud Orchestrator	Data Discovery	Metadata management
APIs and std interfaces	APIs and std interfaces	APIs and std interfaces	APIs and std interfaces

Data/Compute/Storage Resources

Federations tools AAI, Accounting, monitoring, Helpdesk, etc.

EOSC Portal

Other Portals and Marketplace

Federated Discovery and Brokerin,/supporting (Federated AAI)

Technical Roadmap

- Built to facilitate the service composition
  - Adoption of standards and participation to standardisation bodies (RDA, OGF, etc.)
  - Develop new solutions as result of integration activities
  - Fostering the integration of thematic services with federation and common services
- Enhance the EOSC-Hub access channels to improve
  - Discoverability, access and combined usage of services
  - See EOSC Portal in particular
- Establish and further enhance collaborations with other EOSC projects
  - OpenAIRE Advance, GNT4-3, ESFRI Clusters, regional and national projects, implementation projects, etc.

EOSC-hub examples of service integration work
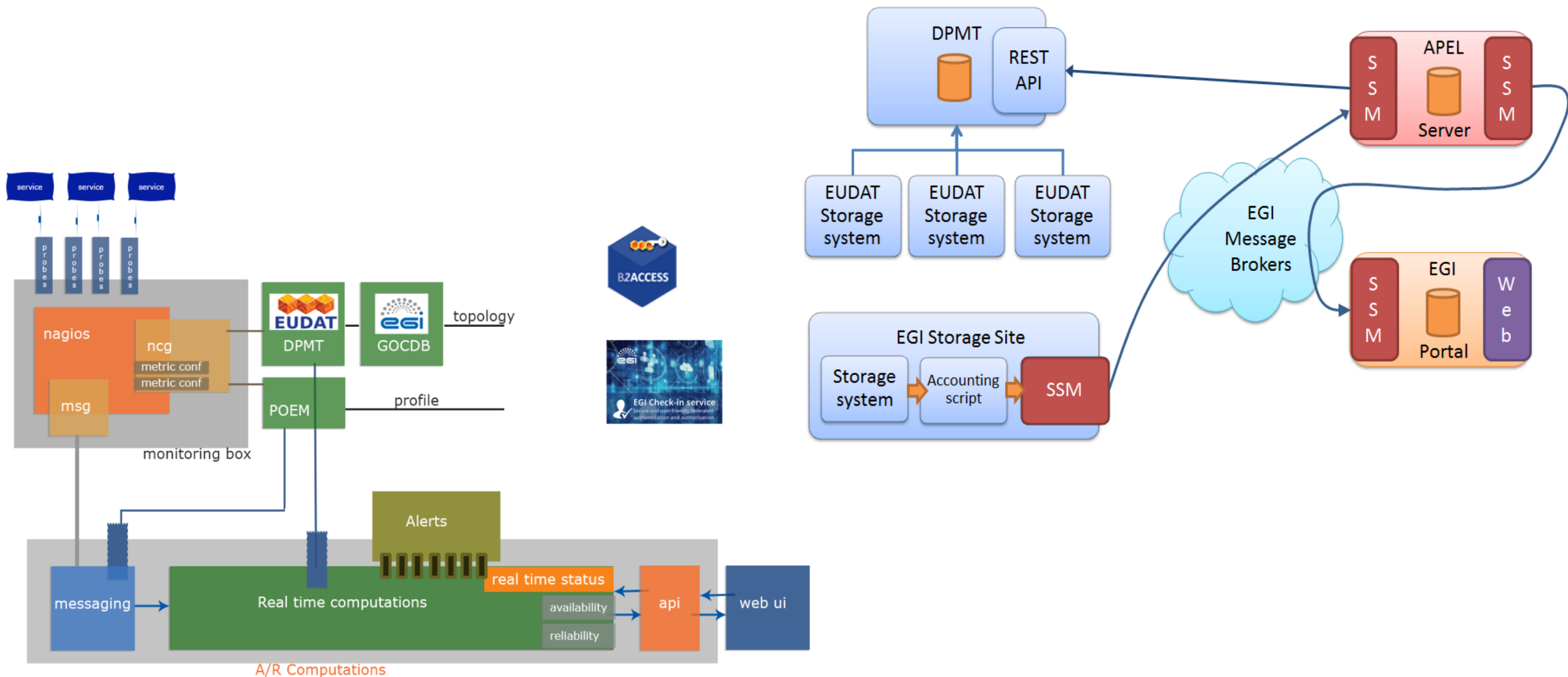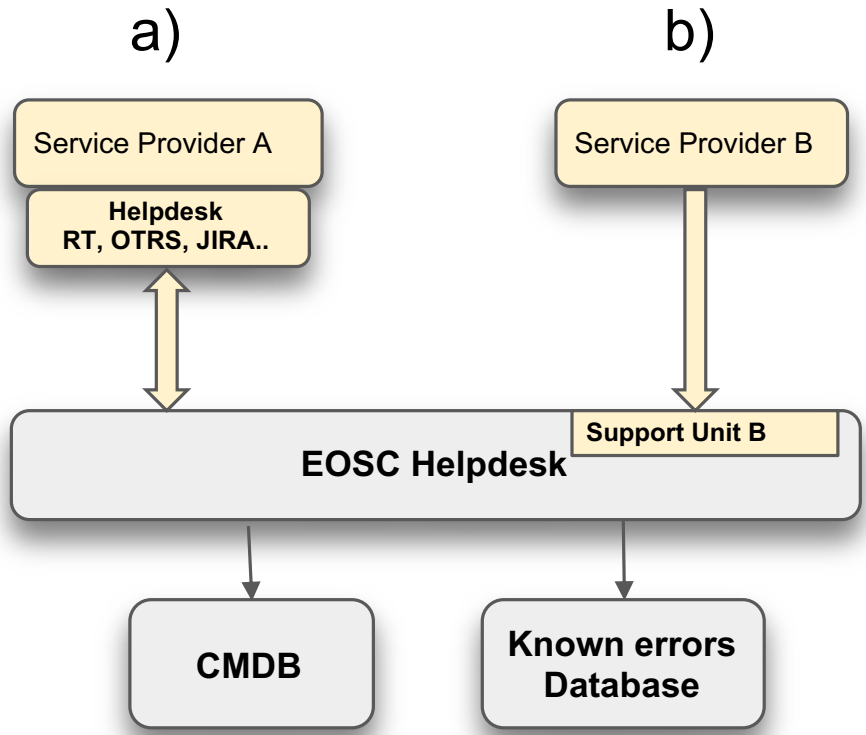
EOSC Federation Services (functional roles)

T5.1	T5.2	T5.3,4,5	T5.6
Access	**Service Discovery and Ordering**	**Operations Support**	**Software and App Repositories**
Check-in	EOSC Portal / Market place	Monitoring Accounting	AppDB
B2ACCESS	Service Portfolio Tool	Config. Database	GitLab
eduTEAMs IAM	SOMBO (Service Order Mgmt. System)	Helpdesk	EGI Software Repo

Authentication & Authorization Infrastructure
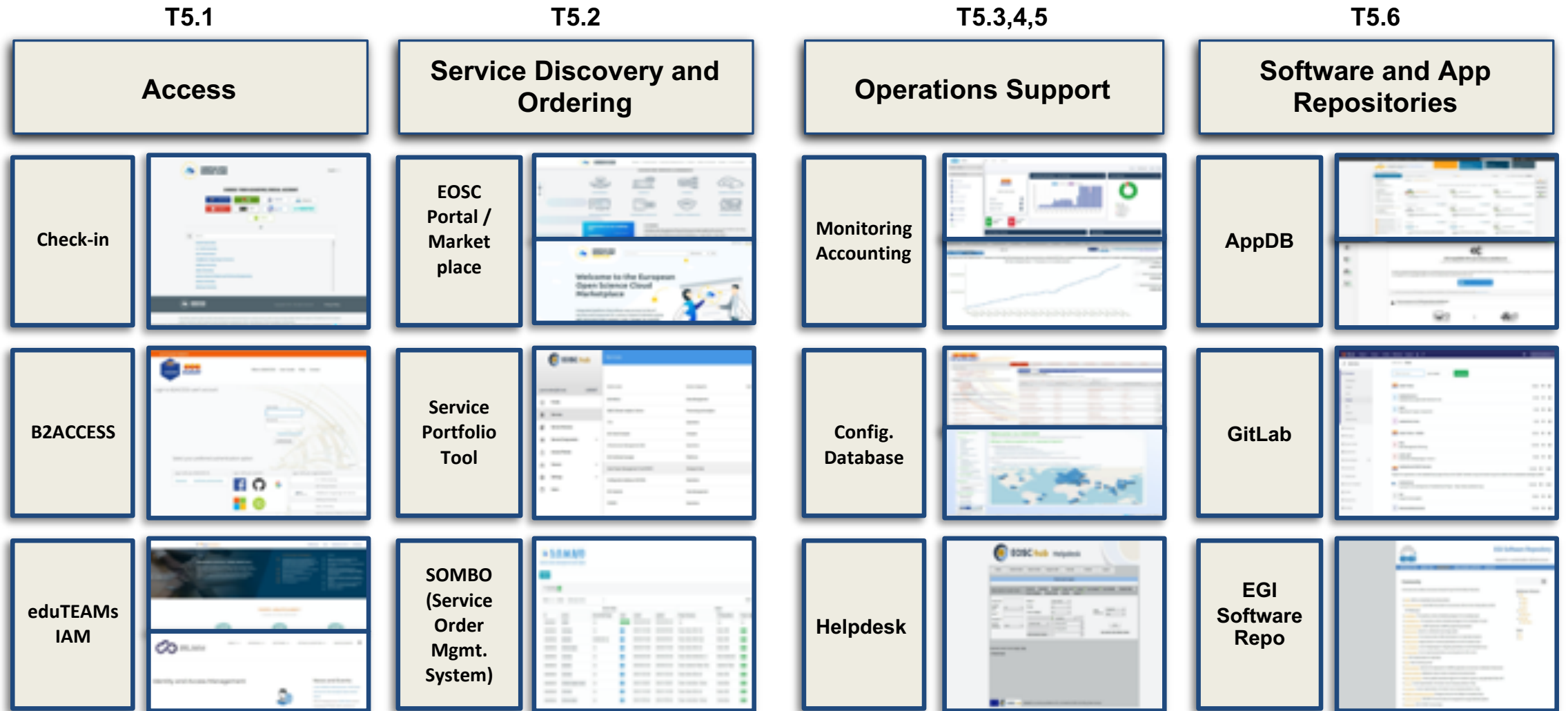
- **EOSC-hub AAI high-level architecture** based on the "community-first" approach of the AARC Blueprint Architecture (BPA)

- **Community:** Enables the use and management of community identities for access to EOSC resources

- **Infra Proxy:** Enables access to resources offered by Service/Resource Providers connected to the R/e-Infrastructures
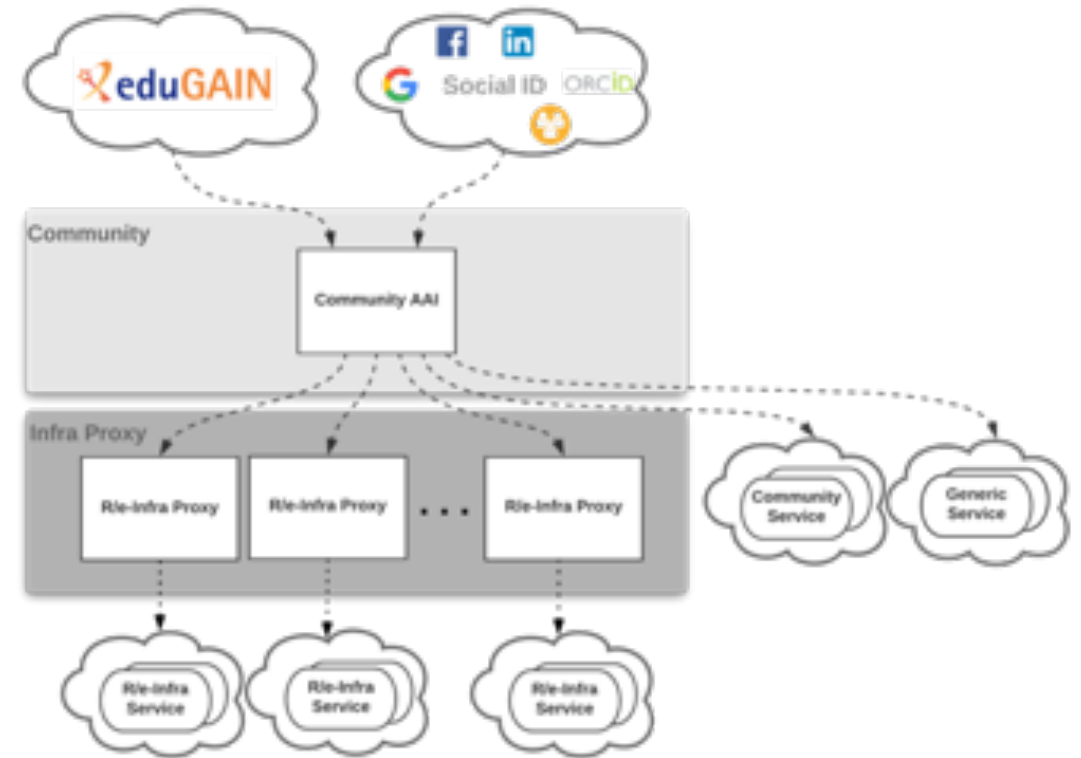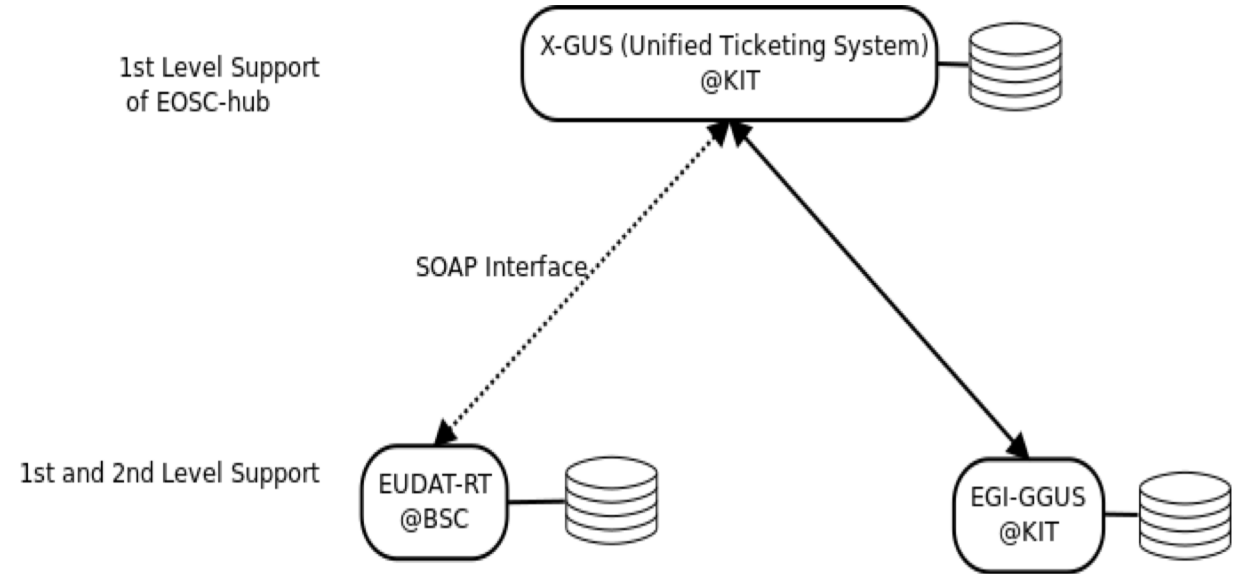

High-level AAI architecture for access to EOSC resources

Activity	B2ACCESS	Check-in	eduTEAMS	INDIGO-IAM
Alignment of privacy statements	✓	M21	✓	✓
Alignment of operational security and incident response policies	✓	✓	✓	✓
Alignment of Acceptable Use Policies (AUPs)	M24	M24	✓	✓

Helpdesk interoperability

a)

b)

```
Service Provider A

   Helpdesk
   RT, OTRS, JIRA..

        EOSC Helpdesk    Support Unit B

   CMDB           Known errors
                  Database
```

Service Provider B

High level EOSC Helpdesk Architecture

1st Level Support
of EOSC-hub

X-GUS (Unified Ticketing System)
@KIT

SOAP Interface

1st and 2nd Level Support

EUDAT-RT
@BSC

EGI-GGUS
@KIT

Integration of EOSC-hub helpdesk with EUDAT-RT and EGI-GGUS

Goal: Common data discovery and access layer for EOSC-hub

Major Achievements:

- Integration of B2FIND with EOSC data services
  - Harvest from EGI-Datahub
  - Harvest from B2SAFE (harbadrop)

- Integration B2DROP-B2SHARE

- Enable transparent cross infrastructure data transfers
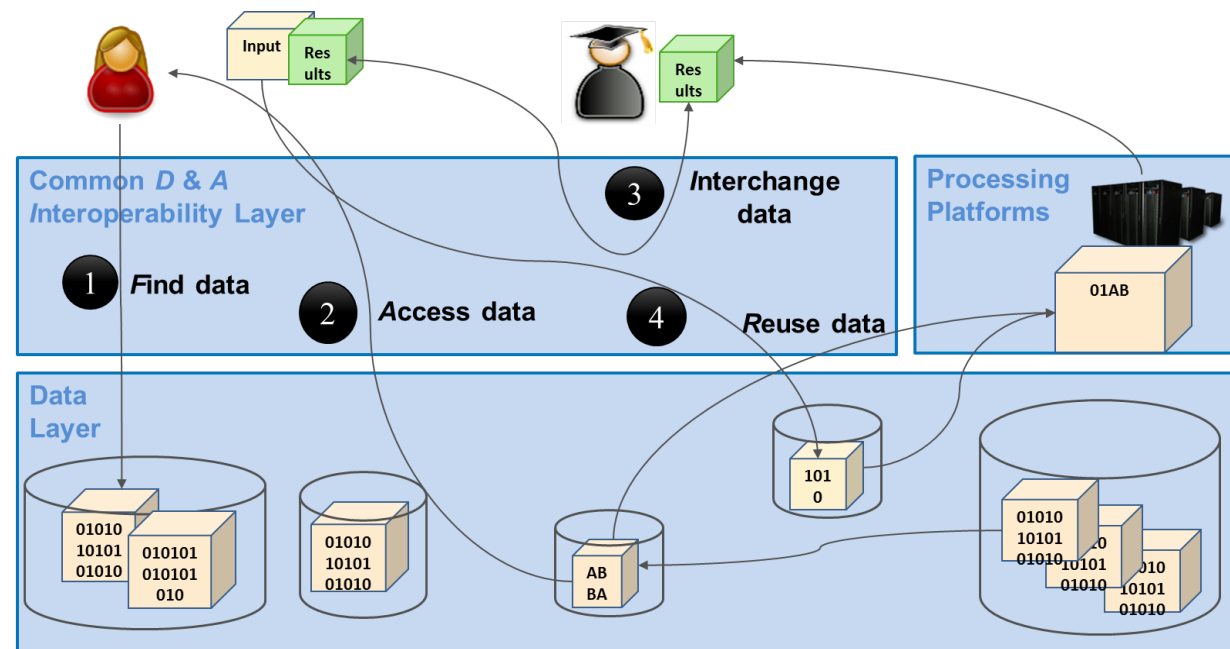
1. [F]ind distributed data in EOSC-hub and beyond
2. [A]ccess distributed data resources wherever they are located
3. [I]nteroperate by sharing and publishing research output
4. [R]reuse, exchange and stagedata (depending on the use case)

Services: B2FIND, EGI Datahub, INDIGO IAM, B2STAGE, B2SAFE, B2DROP.

Data Preservation – Curation - Provenance

**Goal:** Integration of a Certified Trusted Digital Repository into EOSC-hub

Major Achievements:

- Two ingest points (milestone 6.7)
  - Production ready CINES TDR
  - DANS-KNAW proof of concept
- CINES TDR being added to EOSC marketplace
- Two different technical solutions:
  - B2SAFE and B2SHARE as ingest points
  - Same quality constraints (CTS certification)
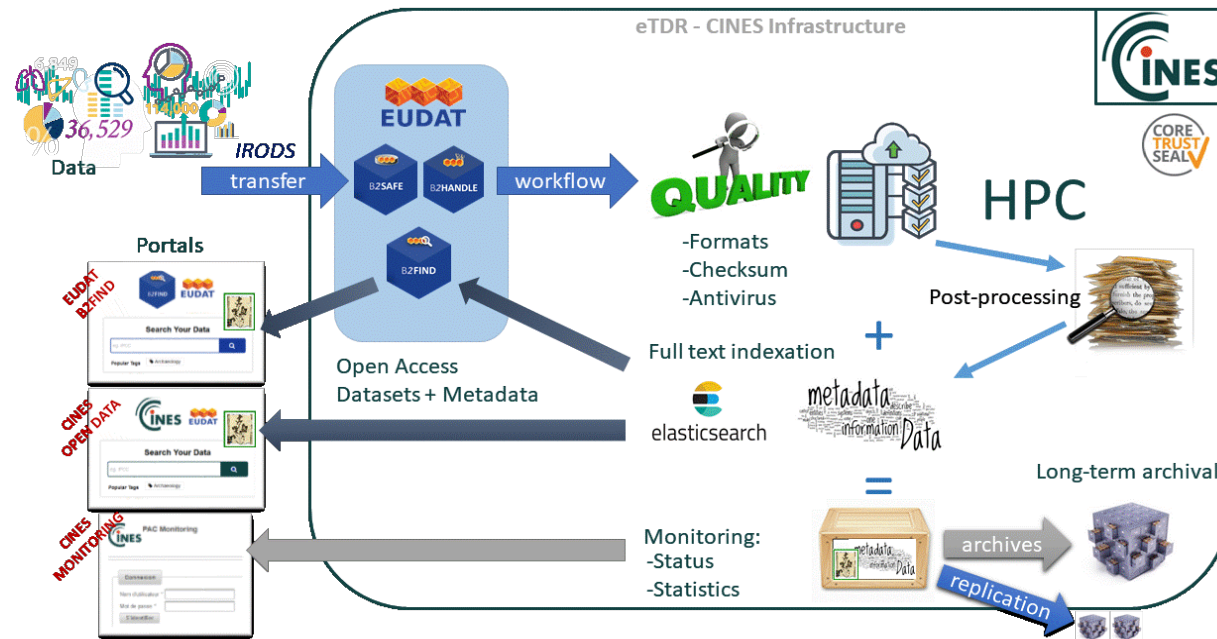
CINES eTDR:
- DSA Certified TDR (CTS in progress)
- B2SAFE as ingest
- Findable data (B2FIND)
- Being added to EOSC marketplace
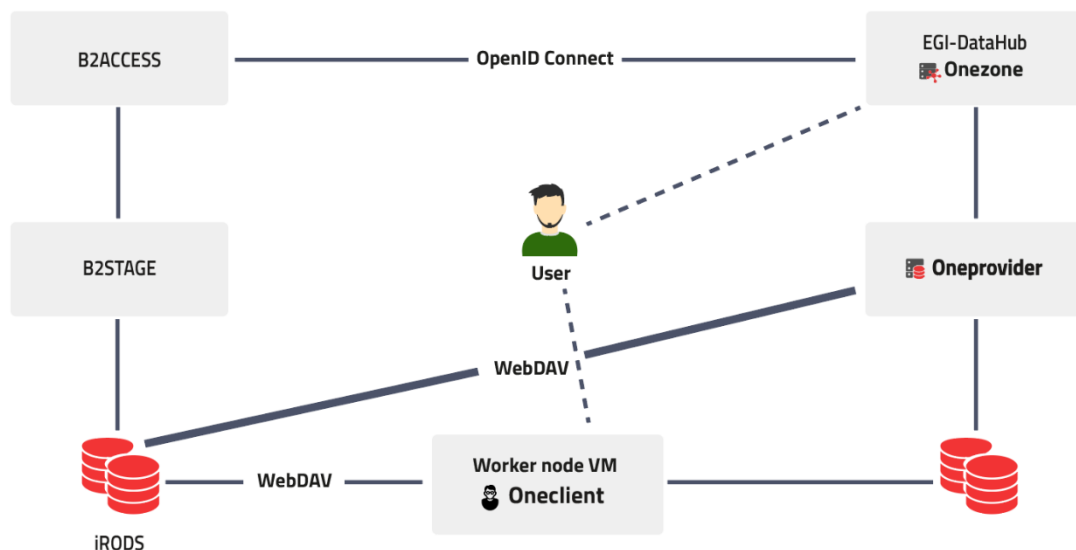
The integration of the EGI-DataHub and B2STAGE/B2SAFE services for the purpose of transparent data transfer between these infrastructures included development of the following features:

Goal: Transparent data exchange between e-infrastructures (EGI<->EUDAT)

- Real end-user driven activity
- Core WP6 integration task
  - Milestone 6.6
- Solution demonstrated
- Dissemination material provided to WP11

EOSC-hub Liaison with external bodies

Collaboration with RDA
Other bodies

Collaboration with RDA

- **FAIR Data Maturity Model WG**
  - EOSC-hub regularly contributing to meetings providing expertise on the assessment criteria for FAIR data that the WG is developing, and their validation against known use cases.
- **Software Source Code Identification WG**
  - Software code sharing through collaborative platforms like the Application Database
- **Array Database Assessment WG**
  - Supported by Task T7.3 (ECAS, climate research)
- **Health & privacy IG/WGs**
  - Contributions from WP2.4
- Meeting and workshops
  - RDA11th Plenary (at Berlin), EOSC-Hub, FREYA, and OpenAIRE-Advance projects together organized a **"EOSC-related European Projects getting Global: Engaging with the RDA - RDA 11th Plenary BoF meeting"** to:
  - share interests in RDA activities for each of the three projects,
  - identify new activities relevant to the RDA context,
  - discuss RDA recommendations and outputs of interest,
  - liaise with interested RDA members from all regions.

Collaboration with RDA (cont) and other bodies

- "**Research Data Repository Interoperability**" WG

- "**Metadata**" IGs

- "**Sensitive Data in the Open Science**" WG

- "**Data Usage Metrics**" WG

- "**Data Management Plans** (DMP) Common Standards"

- Together with OpenAIRE: RDA WG "Make Data Count"

- Contribution to **TOSCA standard**: for the orchestration of Cloud services

- **OGC standard:** in the field of the Earth Observation cases, and the GLUE standard

- Strict correlation with **AARCH** for the AAI implementation

EOSC-hub Technical Support

Users' requirements gathering

Workflow for users' requirements gathering (T10.3/T10.4 to be completed)

Technical support is provided to all potential EOSC-hub users (internal and external) by T10.3 and T10.4 in collaboration with other technical work-packages (WP5, WP6, WP7)

Technical Support
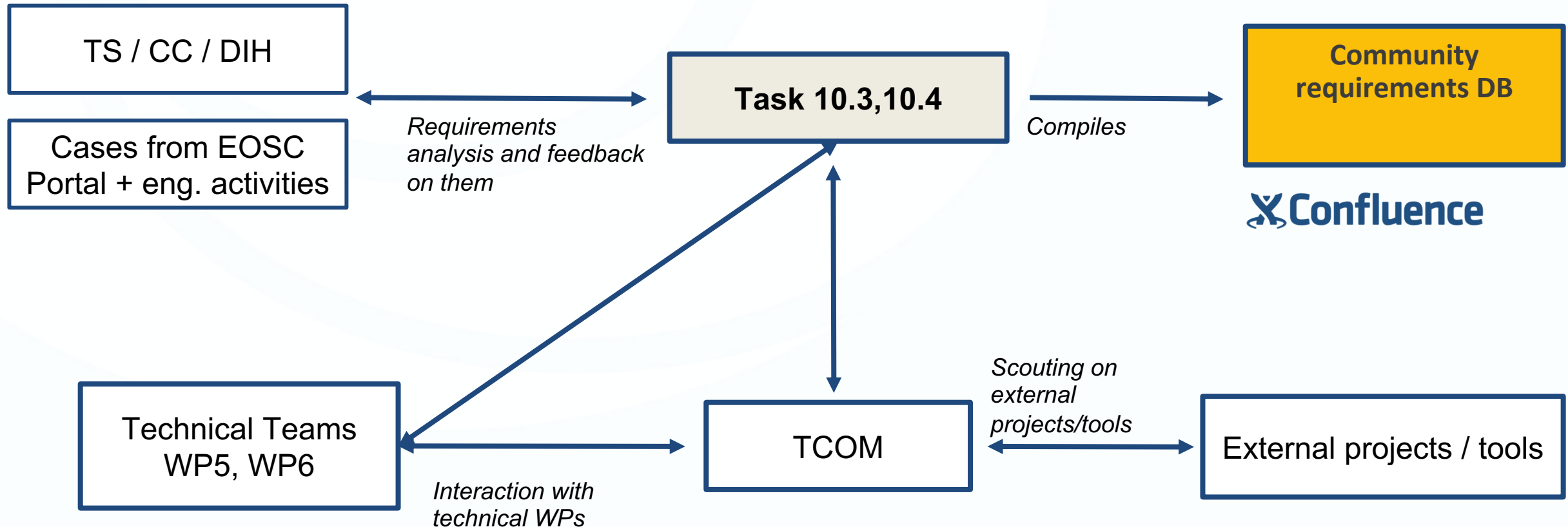
- 22 Communities recorded in the Community Requirements DB
  - Thematic services, competence centers, business pilots, cases from EOSC Portal or engagement activities
  - Around 100 user stories and 120 use cases recorded and analysed
  - 109 technical requirements for technical teams (WP5, WP6, WP7) recorded → new service releases

- T10.3 is supporting the Early Adopter Programme
  - Technical reviews of proposals
  - Shepherding of shortlisted proposals
    - Identify service and resource providers
    - Define a technical plan for integration

Highlights and main results

- First EOSC Technical Architecture (proposal)
  - Input to the "EOSC Architecture WG" → EOSC Interoperability layer
  - Feedback collection on-going → survey in the EOSC Secretariat platform
  - Cross-project Workshop on technical architecture with OpenAIRE and GEANT (June 2019)
- Initial specifications for a number of building blocks
  - Federation services: AAI, Monitoring, Accounting, etc
  - Common Service: Cloud IaaS, Cloud Container, PaaS, Data Repository, etc
- First EOSC Hub Technical Roadmap
- Ongoing collaborations with RDA
- 22 internal & external user communities supported
  - > 100 user stories & use cases analysed
  - > More than 100 technical requirements defined for WP5/WP6,WP7

Future plans

- Enhance the proposal for the EOSC Technical Architecture
  - Feedback from relevant stakeholders
  - Liaison with EOSC Architecture WG
  - Organise a Workshop to discuss an enhanced version of the EOSC Technical Architecture
- Release the first set of EOSC Technical Specifications for external comments
  - Involve other key actors in the definition of the EOSC Technical Specifications
- Updated Technical Roadmap
- Continue the collaboration with RDA
- Continue Technical Support activities

Thank you for your attention!

Questions?

Contact

donvito@infn.it

EOSC-hub

🔗 eosc-hub.eu   🐦 @EOSC_eu