EOSC-Hub Service Portfolio

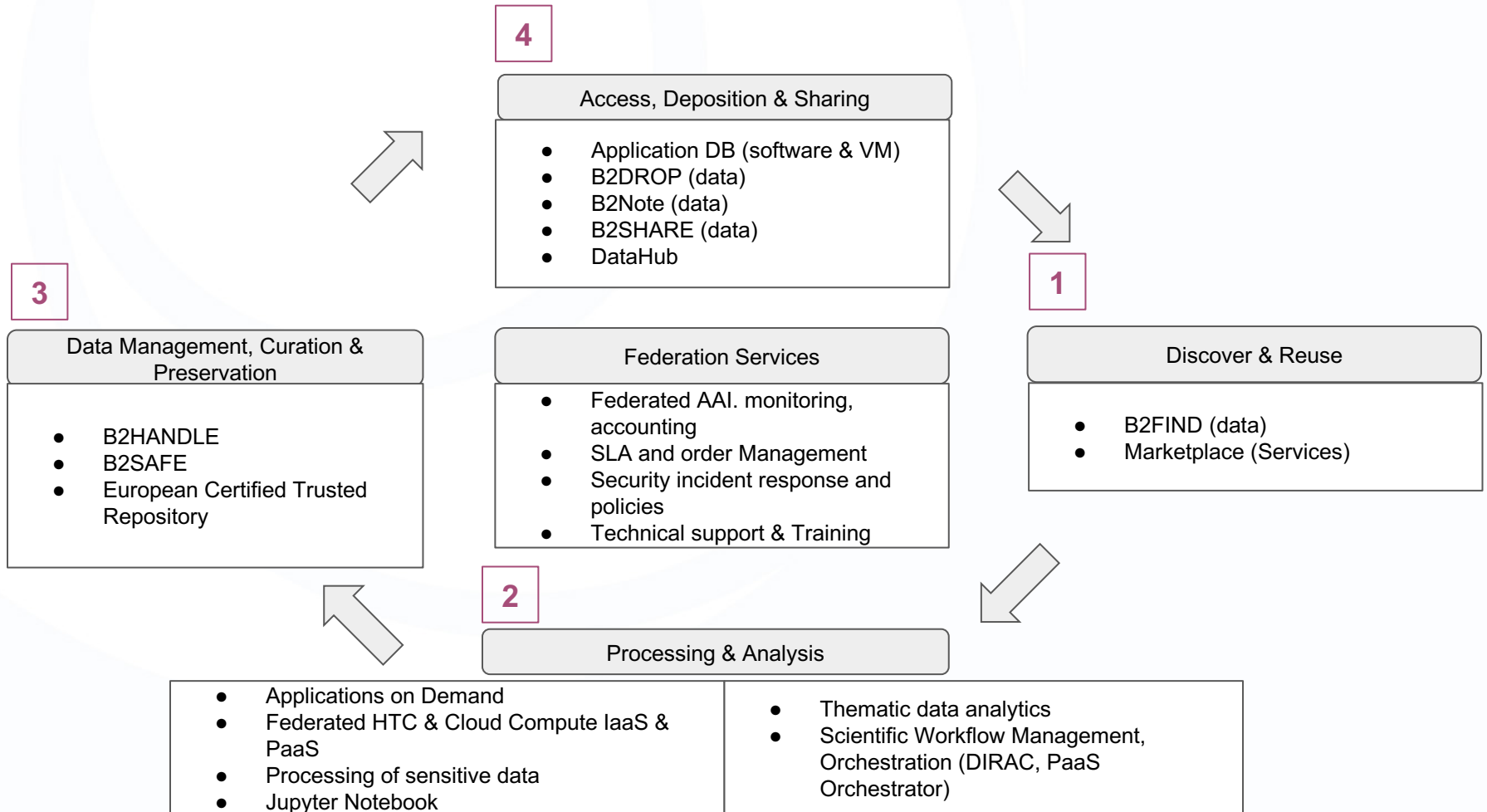🔗 eosc-hub.eu

🐦 @EOSC_eu

EOSC-hub Service Categories

- Discover and Reuse
- Processing & Analysis
- Data Management, Curation & Preservation
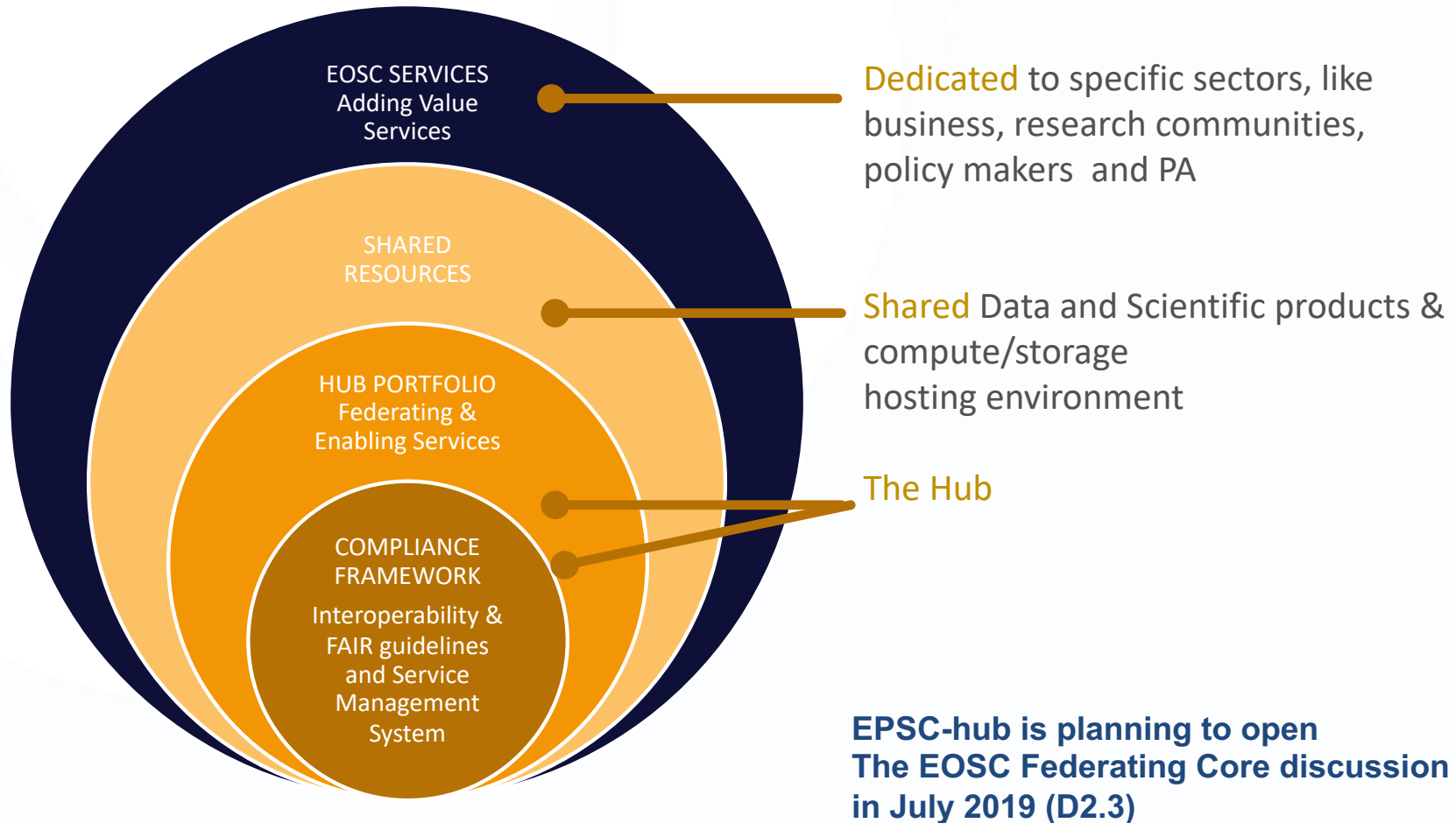- Access, Deposition & Sharing
- Federation services

EOSC-hub Service Catalogue:
https://wiki.eosc-hub.eu/display/EOSC/EOSC-hub+service+catalogue

Services for the Research Data Lifecycle

4

Access, Deposition & Sharing
- Application DB (software & VM)
- B2DROP (data)
- B2Note (data)
- B2SHARE (data)
- DataHub

3

Data Management, Curation & Preservation
- B2HANDLE
- B2SAFE
- European Certified Trusted Repository

Federation Services
- Federated AAI. monitoring, accounting
- SLA and order Management
- Security incident response and policies
- Technical support & Training

1

Discover & Reuse
- B2FIND (data)
- Marketplace (Services)

2

Processing & Analysis
- Applications on Demand
- Federated HTC & Cloud Compute IaaS & PaaS
- Processing of sensitive data
- Jupyter Notebook

- Thematic data analytics
- Scientific Workflow Management, Orchestration (DIRAC, PaaS Orchestrator)

EOSC Federating Core (yellow)

EOSC-hub

EOSC SERVICES
Adding Value
Services

SHARED
RESOURCES

HUB PORTFOLIO
Federating &
Enabling Services

COMPLIANCE
FRAMEWORK

Interoperability &
FAIR guidelines
and Service
Management
System

Dedicated to specific sectors, like business, research communities, policy makers and PA

Shared Data and Scientific products & compute/storage hosting environment

The Hub

EPSC-hub is planning to open The EOSC Federating Core discussion in July 2019 (D2.3)

EOSC-hub Service Provisioning Model

- Shared Services offered in multi-tenant mode
  - **EOSC-hub funded**: Management, access provisioning and technical support
  - **Community funded**: Service enabling and integration, customization
- Services requiring the operations of dedicated instances and/or allocated capacity
  - **EOSC-hub funded**:
    - Service capacity for piloting activities (depending on complexity)
    - Brokering (liaison with service suppliers matching RI requirements)
    - Technical support during testing phase
  - **Community funded**:
    - Service customization to the RI's needs and enabling
    - Service capacity for large-scale piloting (testing phase) and/or full service scale provisioning (production phase)
    - Operational costs (e.g. different funding options possible: subcontracting, in-kind contributions provided by third parties against payment, person months)

Engagement levels for service providers

- Governance
    - Participation to service portfolio management and the definition of its policies
- Operational
    - Provisioning of services and resources through the EOSC-hub Marketplace as delivery channel
    - Participation to coordinated operations of the services provided to the service catalogue
- Technical
    - Enabling of access to services and resources through the Hub: AAI integration

EOSC-hub   FAIR Data

F indable   A ccessible   I nteroperable   R eusable

Data has **rich metadata**, specifies **data identifiers**, has a **globally unique persistent identifier** and is registered and/or indexed in **searchable resources**

Data is **retrievable** via their **identifier**, via **standard protocols**, protocols are **open, free** and **universally implementable**, allows **authentication** and **authorization** where necessary and **metadata** are kept **accessible** when **data** is **no longer available**

Data has formal, accessible, shared and broadly applicable **knowledge representation,** use **vocabularies** that flow **FAIR principles** and include qualified **references** to **other (meta)data**

Data has plurality of accurate and **relevant attributes**, is released with clear and accessible **data usage license**, has associated **provenance** and meets domain-relevant **community standards**

EOSC-hub Support to FAIR

- Findable
  - Discovering scientific products, services and resources
- Accessible
  - Enabling access to scientific resources and services
- Interoperable
  - Integrating and maintaining relevant services according to a defined standards roadmap
- Reusable
  - Managing and preserving data (raw data, metadata, software, scientific products, annotation)

(1)Discover and Reuse

Enhance discovery and reuse of scientific products across communities
  B2Find
  Marketplace

B2FIND

Making Open Science findable
(http://b2find.eudat.eu/)

Provided through EOSC-hub

- Cross-disciplinary metadata and discovery service (B2FIND) allowing RI to make their data findable and discoverable in a central catalogue
  - Metadata can be harvested via OAI-PMH. Possibility to use also APIs as JSON-API's and CSW2.0 to collect the metadata from the communities.
  - The project provides support to integrate community data catalogue

To be funded by the community
- Elicitation & mapping of metadata schemas in use within the community
- Definition of FAIR implementation guidelines (input to EOSC-hub)

EOSC-hub Marketplace

Provided through EOSC-hub:

- Marketplace: multi-tenant user-facing platform for service providers to publish their EOSC services and EOSC-compliant data repositories, and collect service orders
  - Mature services and curated data
  - The RI retains control and accountability for the services and data published and participate in the management of the Hub service portfolio
  - Support to usage of common service templates

To be funded by the community:

- Service design to set-up a cluster-wide or RI-specific service portfolio
- Advice on ISO-compliant service portfolio management process, auditing
- Operation of a dedicated customized marketplace (if necessary)

(2) Processing and Analysis

Scale out your computing environment and process & analyse data in a federated environment

- Applications on Demand
- Federated High Throughput Computing
- Federated Cloud Compute IaaS and PaaS
- Processing of sensitive data
- Jupyter
- Scientific Workflow Management, Orchestration (DIRAC, PaaS Orchestration)
- Discipline-specific data analytics tools

Applications on Demand

Online scientific applications and application-hosting frameworks with computing and storage for compute-intensive data analysis (https://marketplace.egi.eu/42-applications-on-demand-beta).

Provided through EOSC-hub:

- Hosting platform, compute and storage, extendible with new applications, application-hosting frameworks, and HTC or cloud resources: custom applications can be executed on EGI Cloud Compute and High-Throughput Compute services and offered as scalable, online services to researchers worldwide
- Application porting support

To be funded by the community:
- Porting of applications and support to end-users

Federated High Throughput Computing (HTC)

Run computational jobs at scale on the EGI infrastructure. It allows you to analyse large datasets and execute thousands of parallel computing tasks. HTC is provided by a distributed network of computing centres, accessible via a standard interface and membership of a virtual organisation (https://marketplace.egi.eu/32-high-throughput-compute)

Provided through EOSC-hub:
- Technical support
- Capacity via brokering to national HTC providers
- Services to federate community-owned HTC clusters

To be funded by the community:
- Capacity for compute-intensive applications
- Service enabling via porting of community applications

Federated Computing IaaS and PaaS

Execute compute- and data-intensive workloads (both batch and interactive)
Host long-running services (e.g. web servers, databases or applications servers)
Create disposable testing and development environments on virtual machines and scale your infrastructure needs (https://marketplace.egi.eu/31-cloud-compute).

Provided through EOSC-Hub:

- Multi-cloud IaaS with Single Sign-On (EGI Federated Cloud)
- Run Docker containers (deploy and scale Docker containers on-demand)
- Appliance Library to share and automatically distribute applications Orchestration to easily move applications across providers.
- Unified web dashboard to interact with all providers.
- Services to federate community-owned cloud resources
- Technical support

To be funded by the community:

- Capacity for compute-intensive applications
- Service enabling via porting of community applications

Processing of sensitive data

Provided through EOSC-Hub:

- If access to data is restricted by National or European regulations or by other confidentiality policies, the sensitive data services provide:
  - A secure IT platform to store, process, analyse and share data in a secured environment
  - Provide secure, separated and private environments enforced via strong access rules
  - Provide consultation and technical support to make use of the Sensitive Data Service

To be funded by the community:

- Service provisioning and capacity on the Sensitive Data Services

Jupyter Notebook

Share documents with live code, equations, visualisations and explanatory text.

Provided through EOSC-hub:

- Jupyter Notebook service with
    - AoD integration
    - Persistent storage
    - Customized notebook environments
    - Access to other EOSC services from the notebooks

To be funded by the community:

- Community customisation & specific services instances.

Scientific Workflow Management and Orchestration

- DIRAC4EGI: Workload management service to distribute jobs and manage centrally thousands of computational tasks on cloud and HTC
- TOSCA-based deployment orchestration on multiple IaaS

Provided through EOSC-hub:

- Operations of of workflow management system and orchestrator
- Technical Support
- Compute infrastructure, brokering to national compute providers

To be funded by the community:

- Community customisation & specific instances.

Who	Service
WeNMR. A worldwide e-Infrastructure for NMR spectroscopy and Structural biology	**Amber** is a suite of programs that allow users to perform molecular dynamics simulations on biological systems **HADDOCK** is an information-driven flexible docking approach for the modelling of biomolecular complexes. The **CS-ROSETTA** web server generates 3D models of proteins. **DISVIS** allows visualising and quantifying the information content of distance restraints between macromolecular complexes. **FANTEN** is a user-friendly web tool for the determination of the anisotropy tensors and residual dipolar couplings. The **GROMACS** web server is an entry point for molecular dynamics on the grid. **POWERFIT** performs a full-exhaustive 6-dimensional cross-correlation search between the atomic structure and the density. The **UNIO** web server is an entry point for molecular dynamics on the grid. Besides the application software, the services also provide automated pre- and post-processing, the compute, storage and job scheduling and monitoring for running the application.
ENES. Services for Climate Modeling in Europe	**The ENES Climate Analytics Service (ECAS)** will enable scientific end-users to perform data analysis experiments on large volumes of climate data, by exploiting a PID-enabled, server-side, and parallel approach
Compact Muon Solenoid (*CMS*)	**Dynamic On Demand Analysis Service (DODAS)** provides dynamic generation of scalable, monitored HTCondor-based batch system clusters and Spark/Hadoop-based Big Data clusters instantiated on-demand over IaaS clouds

Who	Service
CLARIN (European Research Infrastructure for Language Resources and Technology)	**The Component MetaData Infrastructure** provides a framework to describe and reuse existing metadata blueprints
INCD (Portuguese National Infrastructure for Distributed Computation that provides scientific computing services for science)	**On-demand Operational Coastal Circulation Forecast Service (OPENCoastS)** builds on-demand circulation forecast systems for selected sections of the Portuguese coast
Earth Observation Data and Adding Value Services	**MEA** is a geospatial data analysis tool empowered with OGC standard interfaces. **EPOSAR** allows for a systematic generation of ground displacement maps and time series. **Sentinel Playground** - provide access to complete archive of Sentinel-2 data and ESA Archive of Landsat 5,7 and 8. **Datacube Data Analytics Service** proposes a multi-sensor, -scale and -purpose datacube approach. Geohazards Exploitation Platform is focused on the integration of Ground Segment capabilities and ICT technologies to maximise the exploitation of EO data. **OSS-X Sentinel Service** is a web based system designed to provide EO data users with Search - Cataloguing - Order and Dissemination capabilities for the Sentinel products. **EO Cloud** is a cloud processing platform based on open source OpenStack technology. **EODC SDIP** provides cloud, high performance computing and data storage facilities.

Discipline-specific data analytics 3/3

Who	Service
DARIAH (pan-European infrastructure for arts and humanities)	**DARIAH Science Gateway offers cloud-based services and applications to the humanities research communities**
IFREMER (operator of tools for observing and monitoring oceanographic databases)	**The INFRAMER platform** provides users with marine data collections from state-of-the art integrators in the world. Data collections provided on the platform are public but might require specific license or citation agreement from the users.
EISCAT (next generation incoherent scatter radar system)	**The EISCAT_3D portal** provides services for data cataloguing, discovery and pre-defined analysis

(3) Data Management, Curation and Preservation

Manage, preserve and curate data according to domain specific policies and provide access to HTC, HPC and Cloud for the processing and analysis of data

- B2HANDLE
- B2SAFE
- European Trusted Digital Repositories

B2HANDLE

Making science referenceable (https://www.eudat.eu/services/userdoc/b2handle)

Provided through EOSC-Hub
- Distributed service for storing, managing and accessing persistent references (PIDs) to scientific products
  - Unified technical interface for minting PIDs and PID namespaces (prefixes)
  - Replicated PIDs for high availability and resolution, including reserve lookups
  - Easy integratable and client-side application support through a Python library

To be funded by the community:
- Provisioning of community dedicated PID prefixes
- Provisioning of B2HANDLE service for minting PIDs

B2SAFE
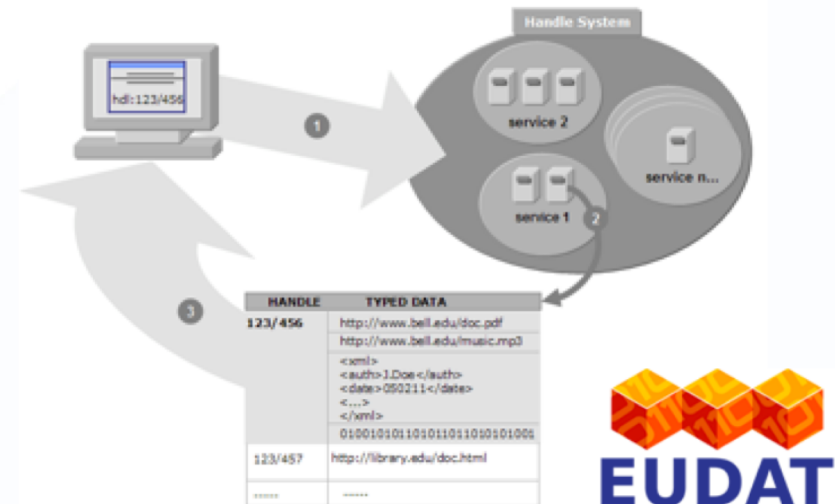
Supporting data Management Policies
(https://www.eudat.eu/b2safe)

Provided through EOSC-Hub:

- Service to implement data management policies in a distributed and federated data infrastructure
    - Enabling access to large scale storage and archiving facilities
    - Replication, persistent identifier and data curation policies to secure data for long term preservation according to domain specific policies;
    - Staging of data to HTC/HPC resources (EGI FedCloud, PRACE HPC, etc.)
    - Technical support on data management policies

To be funded by the community:

- Procurement of large pledges of storage infrastructure to be federated in EOSC
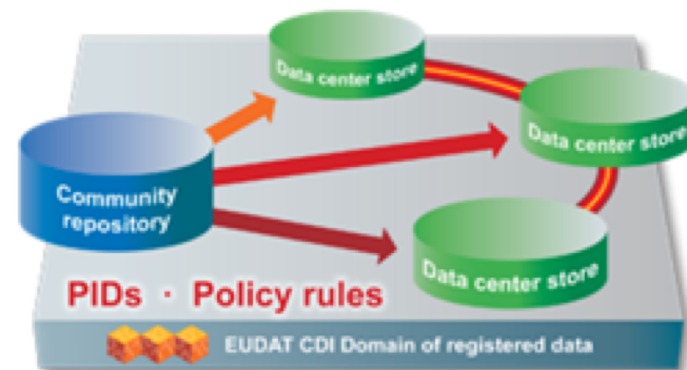
European Trusted Digital Repositories

Provided through EOSC-hub:

- European Trusted Digital Repositories (ETDR) allowing RIs to publish and manage data in:
  - Deposit data in repositories which are certified according to a requirements for trusted digital repositories (e.g. CoreTrustSeal, Data Seal of Approval, Nestor Seal or ISO16363)
  - Ensure reliability and durability, manage, share and curate data in a FAIR way
  - Consultancy to become an ETDR

To be funded by the community:

- Technical support to ingest data in ETDR
- Service provisioning and resources in an ETDR

(4) Access, Deposition, Sharing

Make digital objects (data, software and applications) identifiable and share them with other researchers

- Application Database
- B2DROP
- B2NOTE
- B2SHARE
- DataHub

Application Database

Share/Discover and Use of community-specific scientific software, applications and cloud virtual appliances (https://appdb.egi.eu/)

EOSC-hub funded:

- Application Database platform operations
- Use to the service, including dashboard for managing VAs
- Technical support
- Support to packaging virtual appliances following security best practices

B2DROP

Sync and share research data (https://www.eudat.eu/services/b2drop)

Provided through EOSC-hub:

- Store and share data with colleagues and team members, including research data not finalised for publishing
    - Cloud storage to share data with fine-grained access controls
    - Synchronise multiple versions of data across different devices, including workflow and computing environments
    - Publish data via B2SHARE

To be funded by the community:
- Provisioning and operation of a dedicated customized B2DROP instance (if required)

B2NOTE

Use annotations to structure your data (https://b2note.eudat.eu/)

Provided through EOSC-hub:

- Manage and share annotations on data with colleagues and team members
  - Annotations are keywords or commentaries attached to a object, that explains or classifies it.
  - B2NOTE annotation service is integrated with the B2SHARE service and technology
  - B2NOTE can be easily integrated with other community data repository services
  - Provide training on semantic annotations
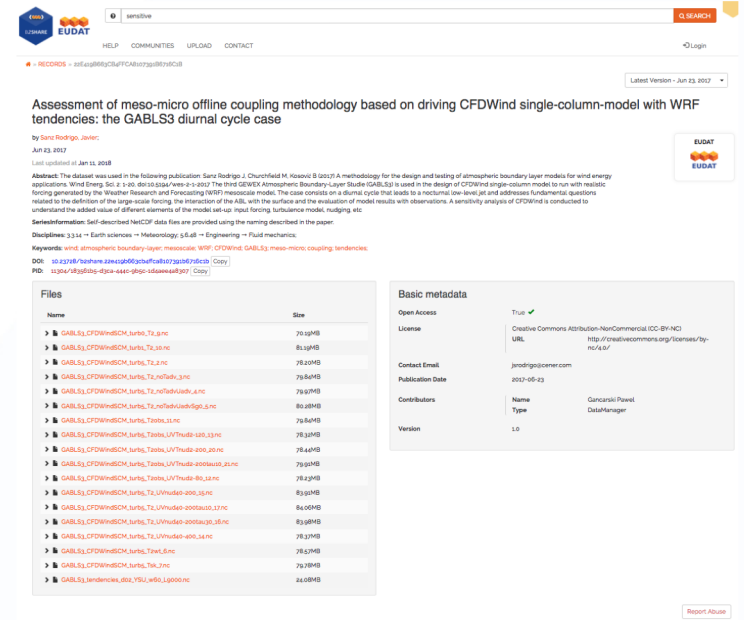
To be funded by the community:
- Technical support on the integration of B2NOTE annotation service into community services
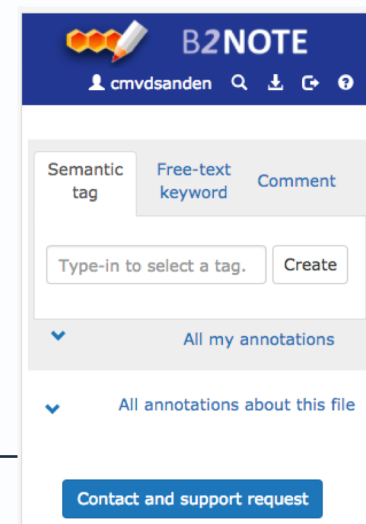
EOSC-hub B2SHARE

Store and publish data (https://b2share.eudat.eu/)

Provided through EOSC-hub:

- Data repository & publishing service (B2SHARE) allowing RIs to publish and manage data in a persistent way
  - Use of DataCite DOIs & EPIC PID
  - Domain specific metadata extensions
  - Manage the publish life cycle with version control
  - Community defined authorisation rules
  - Annotations via defined ontologies

To be funded by the community:
- Customization and provisioning of a (dedicated) B2SHARE instance
- Definition of FAIR implementation guidelines (input to EOSC-hub)

EUDAT

DataHub

Discovery, access and usage of reference open datasets and user data

- Federates existing data sources and data storage providers into one name space
- Brings data to the multiple hybrid clouds and HTC
  - Increased accessibility of data to users → Bring data to computing
  - Scalable federation of distributed data providers
  - Publishing/DOIs
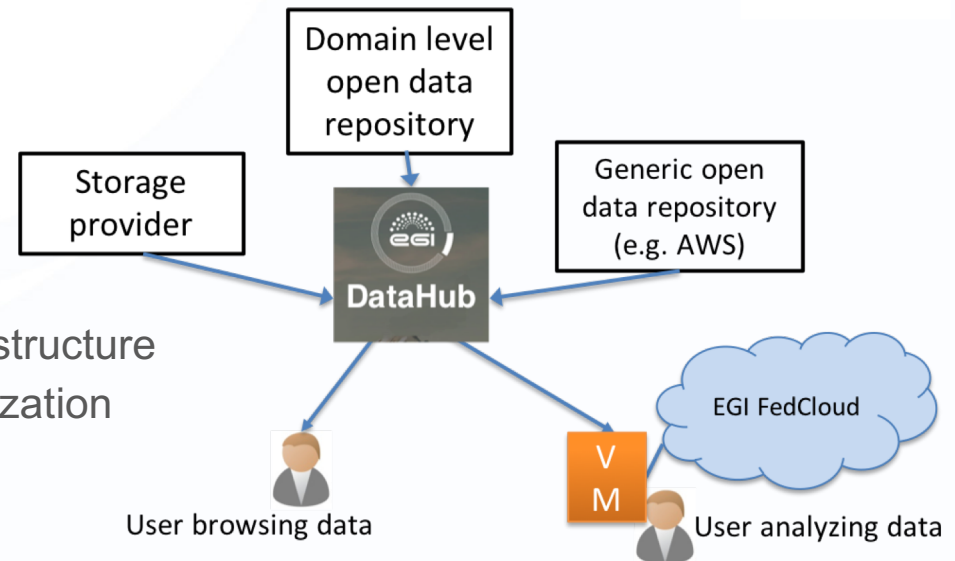 - Use of eduGAIN and federated AAI

Provided through EOSC-Hub:
- Technology and its support
- Test infrastructure

To be funded by the community:
- Storage capacity to scale up data infrastructure
- Dedicated support and service customization

EOSC-hub service catalogue mapping to FAIR

F indable   A ccessible   I nteroperable   R eusable

Service	F	A	I	R
B2Handle	X	X		X
B2Find	X		X	X
Marketplace	X	X		X
Application Database				X
B2SHARE	X	X	X	X
B2DROP		X		
DataHub	X	X	X	X
European Trusted Digital Repositories	X	X	X	X

Service	F	A	I	R
Federated AAI		X		
B2SAFE		X		
B2NOTE			X	X
Federated Cloud/HTC				
Jupyter Notebook				
Applications on Demand				
Workflow Management and orchestration				
Sensitive Data Services				

Federation services

- Federated AAI
- Monitoring and accounting
- SLA and order management
- Security incident response and security policies
- Technical support and training

EOSC-hub AAI

Services for Trust and Identity

Provided through EOSC-hub:

- Multi-tenant service for federated authentication and authorization supporting all main standards
  - Only one account needed for federated access to multiple heterogeneous (web and non-web) service providers using different technologies (SAML, OpenID Connect, OAuth 2.0, X509)
  - Use of federated IdPs in eduGAIN
  - Identity linking enables access to resources using different login credentials (institutional/social)
  - Aggregation and harmonisation of authorisation information from multiple sources
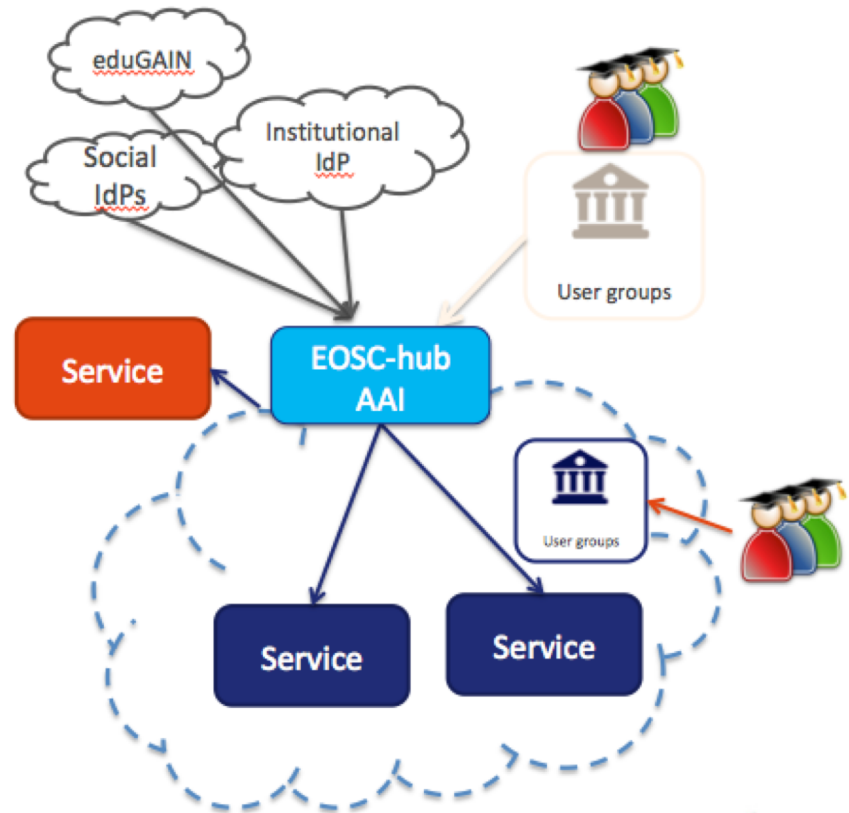
To be funded by the community:

- Dedicated support
- Service customization
- Operations of a dedicated AAI infrastructure (if required)

EOSC-hub AAI Architecture

Secure: operates under strict security policies

Simple: hides the complexity

Low overhead: Easy integration of multiple IdPs and AAs

Interoperable: AARC blueprint, eduGAIN
REFEDS R&S and Sirtfi policies

Supports SAML 2.0, OpenID Connect, OAuth 2.0 and X.509 credentials

Capabilities

- Only one account needed for federated access to multiple heterogeneous (web and non-web) service providers using different technologies (SAML, OpenID Connect, OAuth 2.0, X509)
- Identity linking enables access to resources using different login credentials (institutional/social)
- Aggregation and harmonisation of authorisation information from multiple sources
- AAI is offered in two configurations
  - As a service: EOSC-hub AAI catch-all instance
  - Dedicated instance

Federated service management

- Management of shared services of the Hub and related processes
    - Federated authentication and authentication
    - Marketplace (discovery, order management, SLA management)
    - Helpdesk for incident and problem management
        - can be offered and branded as front desk of RI users
    - IT Security Management (e.g. coordinated incident response, security forensics/monitoring)
    - Service quality assurance (auditing)
        - accounting and monitoring infrastructure
    - Service portfolio management
- Maintenance of the EOSC-hub corpus of policies
    - data policies ← input from INFRAEOSC-04
    - security policies ← input from INFRAEOSC-04
    - standards roadmap ← input from INFRAEOSC-04
    - FAIR implementation guidelines ← input from INFRAEOSC-04

- Data and services to be published in the EOSC-hub marketplace
  - Terms of use of contributed data and services are defined by the RIs
- Definition/Maintenance of
  - FAIR implementation guidelines
  - Community-specific data policies for compliance to applicable laws and technical interoperability
  - Input to EOSC-hub Principles of Engagement (PoE) and standards roadmap

- Participation to EOSC-Hub service portfolio management for all cluster services contributed to the Hub
  - Definition and maintenance of cluster service descriptions
  - Self-certification of compliance of cluster data and services to the Hub
- Participation as service provider to the governance of the Hub (definition and maintenance of the overall Hub Service Management policies and processes) together with other providers e.g.
  - Definition/Maintenance of service level targets and performance reporting
  - Development of accounting probes and publishing of accounting records
  - Definition of policy and security requirements
  - IT security management concerning RI services

Examples of use cases - Competence Centres

	Challenges	Relevant EOSC-hub services
ELIXIR CC	• Establish a federation of cloud providers to replicate ELIXIR Core Datasets and Applications • Bring these cloud-data providers into EOSC • Enable federated AAI across life science and EOSC services	• Federated Cloud • AAI • Operational policies • Experience with Virtual Access • Code of Conduct for sensitive data
Fusion CC (ITER)	• Port fusion workflows to federated compute environment • Federate and enable access to distributed datasets • Application and data provenance	• AAI • Data and compute federation (containers) • Workflow management • Monitoring, helpdesk, • PIDs, ...
Marine CC (Ifremer, Euro-Argo)	• Cloud-based data subscription and data delivery service (from EMSO, Argo, SeaDataNet, etc.) • Enabling users' simulations to run within custom environments on 'subscribed data' • Operate the setup as an EOSC service	• Storage and compute clusters • Data discovery and staging (B2Find, B2Stage) • Jupyter service (24/7) • FitSM (IT service management)
EISCAT_3D CC	• Make EISCAT data accessible via a 'data webshop' • Enable online data analytics for researchers (based on community and custom applications) • Operate the data-compute portal as an EOSC service	• B2Share (metadata schema management) • DIRAC file catalogue and application manager for researchers • Federated compute sites (cloud) • User authentication, authorisation (VOMS, Perun) • FitSM (IT service management)

Examples of use cases - Competence Centres (cont)

	Challenges	Relevant EOSC-hub services
EPOS - ORFEUS CC	- Effective transfer and staging of big data - Harmonised data management policies at seismic observatories - Enable researchers to analyse data, create and publish 'user-defined data products'	- AAI - Data staging and transfer services - Federated compute services - Jupyter (data access and analytics environment) - Data Management good practices
Radio Astronomy CC (LOFAR→ SKA)	- Enable researchers to find, access and process data from the LOFAR Telescope - Support application developers in deploying workflows for researchers - Support research users in data analytics on federation of compute clusters	- B2Find - B2Share - B2Stage - B2Safe - B2Handle (PIDs) - Jupyter - Federated compute and storage resources
ICOS-eLTER CC	- Enable researchers to find, access and process data from ICOS and eLTER - Integrate ICOS Portal and analysis tools with large-scale storage and compute resources - Support research users in data analytics on federation of compute clusters	- B2Safe - B2Find - B2Stage - Federated cloud
Disaster Mitigation Plus CC	- Setup simulation web portals for simulation of natural hazards (storm surge, dust transportation, forest fire, flood) - Federate environmental data from agencies in Asia-Pacific region	- Federated compute resources (HTC and cloud) - Operational tools - FitSM

Thank you for your attention!

Questions?

Contact

🔗 eosc-hub.eu   🐦 @EOSC_eu