

# Data storage

EOSC ask me anything webinar

Debora Testi (CINECA, DICE project coordinator)



with



The EOSC Future, C-SCALE, DICE, EGI-ACE, OpenAIRE-Nexus and Reliance projects are funded by the European Union Horizon Programme calls INFRAEOSC-03-2020 and INFRAEOSC-07-2020.





# Join us on Slido

During this webinar we will be collecting questions and feedback via Slido. Join us there via a mobile device to post your own comments or upvote those of others.



Or

Go to **sli.do**  
Enter event code **#682308**  
and password **EOSC-AMA**



# Agenda

- Welcome and Intro
- Data storage in the EOSC Portal - *Debora Testi (CINECA, DICE coordinator)*:
  - type of resources and brief description
  - how to make available similar services
  - how users can access the services
- Data storage use cases:
  - Using B2Drop as file storage on the cloud for research objects - *Raul Palma (PSNC, RELIANCE coordinator)*
  - DICE storage services for the Geohazards - *Pedro Goncalves (Terradue, RELIANCE)*
- Q&As

# EOSC Marketplace

- The EOSC Marketplace is where researchers can discover, order and access services and resources
- “Categories” is one of the faceted searches available in the EOSC marketplace
  - Meant to help users to narrow down the searches
- The “Ask Me Anything” sessions are providing information on the resources available in the EOSC portal following those categories
- One of the category is “Access physical and e-infrastructures” which include computing, **data storage**, and networking services and resources



The screenshot shows the 'Resources' section of the EOSC Marketplace. At the top is the 'EUROPEAN OPEN SCIENCE CLOUD' logo. Below it is a breadcrumb 'Resources'. The main content is a table of resources.

All Resources	323
CATEGORIES	
Access physical & eInfrastructures	67
Aggregators & Integrators	22
Processing & Analysis	147
Security & Operations	29
Sharing & Discovery	63
Training & Support	32
Other	10
FILTERS	
Scientific Domains	▼
Providers	▼
Dedicated for	▼
Related Infrastructures and platforms	▼
Rating	▼
Order type	▼
Resource availability	▼

<https://marketplace.eosc-portal.eu/>





# Data storage

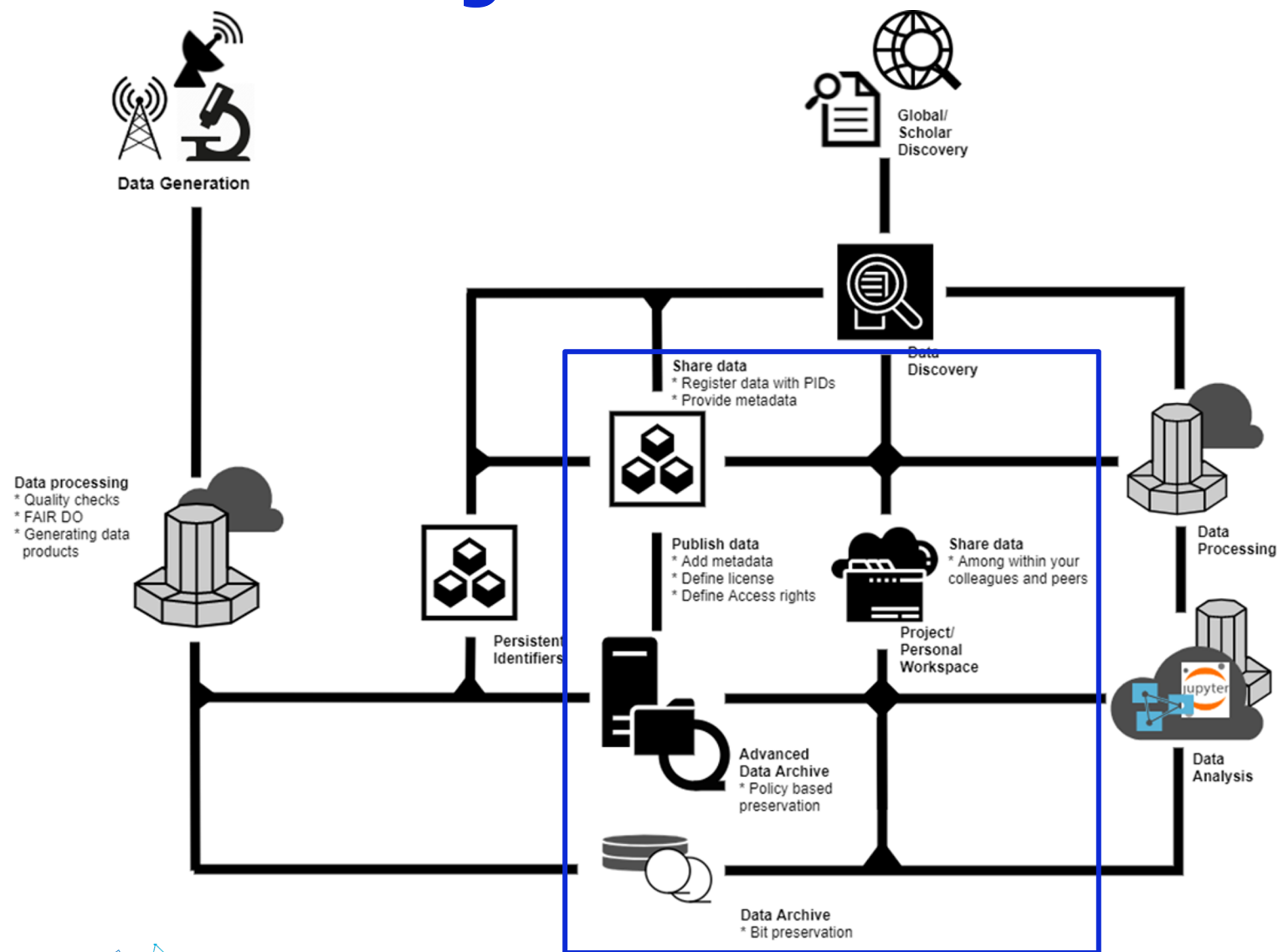
- Services to store data, files and digital objects
- 18 services are available under this category (5% of the total services)
- They are representing the 15% of the total numbers of orders received via the EOSC marketplace
- They are divided into different sub-categories mapping the different features a data storage services can have
  - Online - Archive - Backup
  - Replicated - Synchronised

All Resources 321

## CATEGORIES

Compute	46
<b>Data Storage</b>	<b>18</b>
Archive	4
Backup	1
Data	6
Digital Preservation	1
Disk	1
File	2
Online	3
Queue	0
Recovery	0
Replicated	2
Synchronised	2
Other	3

# Data storage - research data workflow



- Mid-term storage during project (for active data) in good connection to computing facilities
- Mid-term storage resources after/between projects, accessible from computing facilities to store non-active research data
- Mid-long term storage accessible from computing facilities to store non-active stable research data; with optional value-added services (integrity checks, replications, ..)
- Long-term data resources preservation and sharing to store non-active FAIR data



# Services overview



with



# B2SAFE

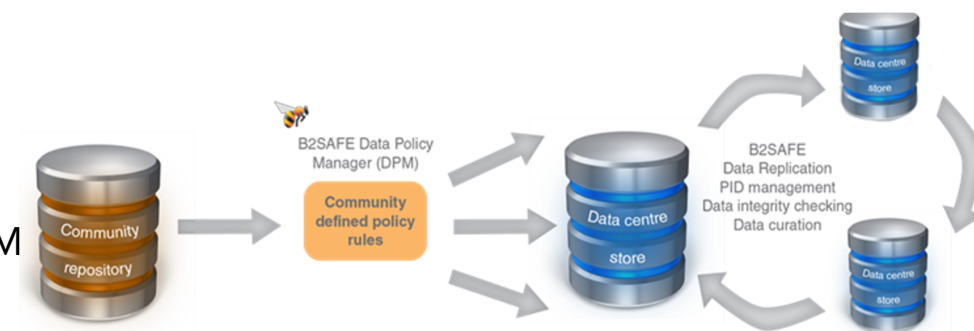


## What

- Provide an abstraction layer which virtualizes large-scale data resources
- Guard against data loss in long-term archiving and preservation
- Optimize access for users from different regions
- Bring data closer to powerful computers

## Why

- It offers support for
  - data management policies (e.g. registration of PIDs, cross-site replication, data integrity checks)
  - policies customised to community and organisational needs
  - both archival and active data
  - large scale storage resources (e.g up to PB-scale)
  - integration with different kind of storage systems (e.g. Tape based HSM, POSIX filesystems, Object storage)
- Accessible via GridFTP, Webdav, iRODS commands
- Service offered by a network of EUDAT service providers

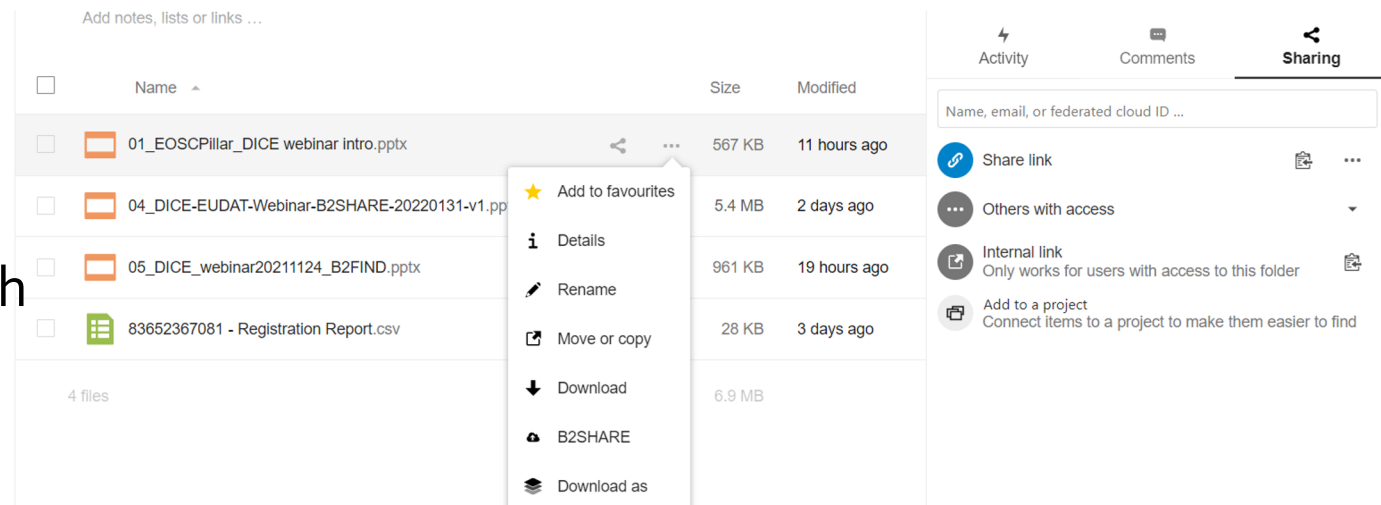


<https://marketplace.eosc-portal.eu/services/b2safe>



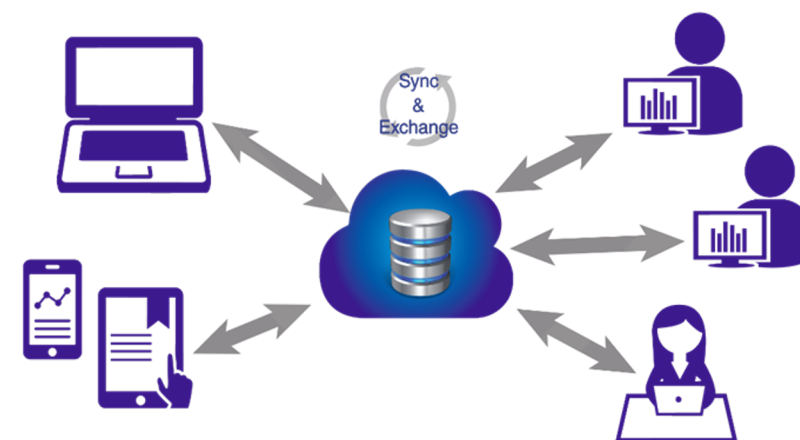
## What

- Low-barrier, user-friendly and trustworthy storage environment which allows users to synchronise their active data across different desktops and to easily share this data with peers
- Default quota of 20GBs (Basic), high quotas optional (Premium)



## Why

- Ease of Use
- Trusted European Service
- Access via Web GUI, desktop clients and Webdav
- Multiple versions of files are kept
- Integration with community services can be supported



<https://marketplace.eosc-portal.eu/services/b2drop>

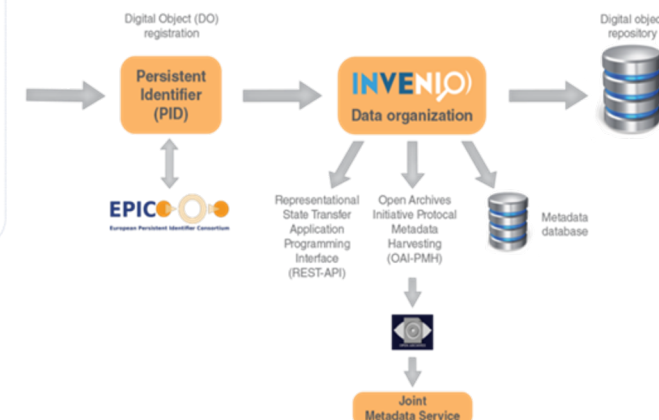


## What

- a user-friendly, reliable and trustworthy way for researchers to store, publish and share research data in a FAIR way
- preserves (small-scale) research data for long-term
- supports community domains with metadata extensions, access rules and publishing workflow

## Why

- Register Data for Publications
- Make known and share registered research data to wider community



Select a domain or project

Linguistics Herbarium Biodiversity Generic **Ontology** RDA Hydro-Meteorology Biomedical Research

Generic

Title \* Title of the resource

Description \*

Creator author

Open Access ☒

Embargo Till

Licence

Keywords keyword1, keyword2, ...

Contact Email contact email

Discipline None selected

<https://marketplace.eosc-portal.eu/services/b2share>

# Data archive services



- CESNET operates storage facilities available through various means of access
  - file system, object storage, sync'n'share, repositories, special applications
  - exceeding 100 PB of physical media in total
  - fully integrated into the national e-infrastructure
- Service offered through DICE
  - Ceph-based object storage
  - available through S3 and block (RBD) interfaces
  - detailed configuration adjustable to user's needs
- Redundancy, geographical replicas, ... to be discussed when storage is requested

<https://marketplace.eosc-portal.eu/services/datacare-object-based-storage>



- INFN and CSCS both offer Object Storage as part of DICE
- Object Storage organizes the way data is written to a storage system so that data is written into self-contained entities called objects. It gives each object a unique ID, which is managed in a flat index, not in a tree-based structure.
- Object storage systems may be accessed for example via a command line interface (CLI) or via a REST API interface

- <https://marketplace.eosc-portal.eu/services/cscs-object-storage>
- <https://marketplace.eosc-portal.eu/services/inf-n-cloud-object-storage-service-for-dice>



# EGI Online Storage

Includes 3 categories of storage services available in the EGI Federation

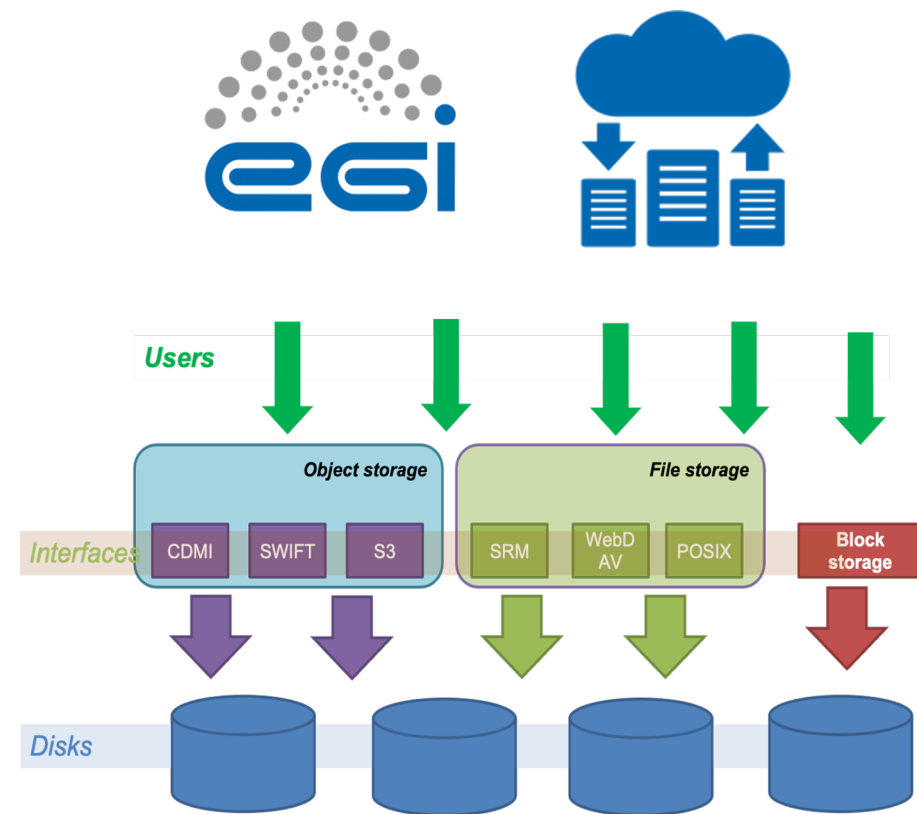
**File storage** can be used for storing and accessing files on the infrastructure as input/output to EGI HTC computations

- Technologies: [DPM](#), [dCache](#), [StoRM](#), [EOS](#)

**Block storage**, allows to **expand** the storage capacity of instances in the EGI Federated Cloud, offering the lowest possible latency for applications

- Technology: Openstack Cinder

<https://marketplace.eosc-portal.eu/services/egi-online-storage>



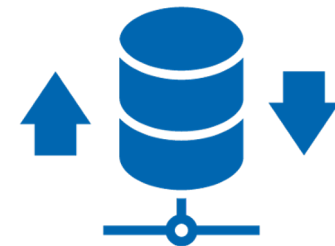
**Object Storage**, allows relatively inexpensive, scalable and self-healing retention of massive amounts of unstructured data

- Technologies: OpenStack Swift, CEPH

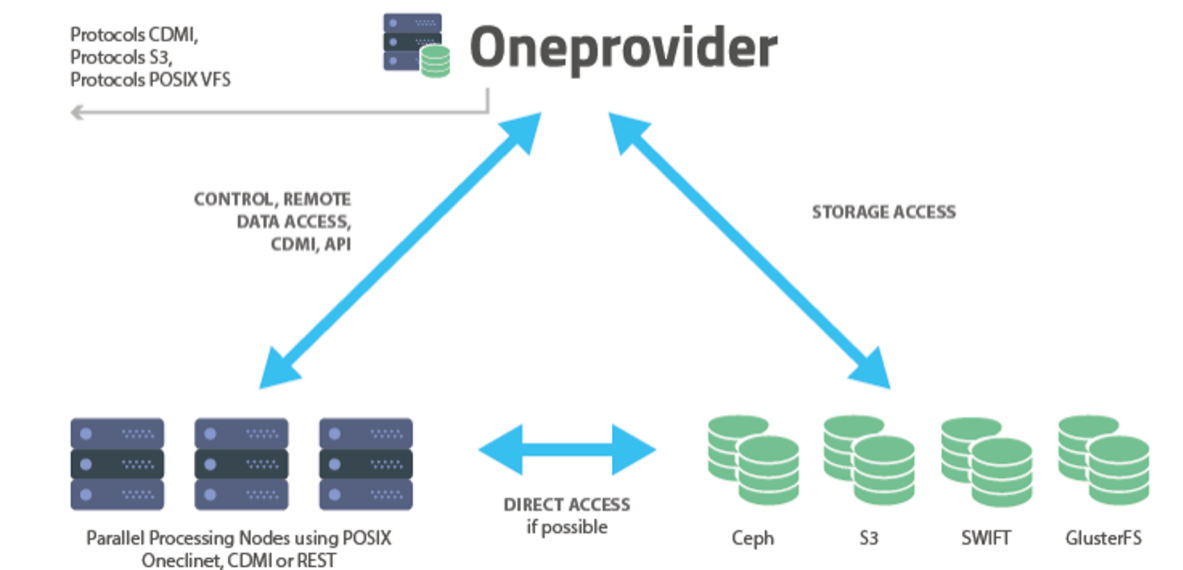
# EGI Data Hub

- **Data-collections aggregation service**
  - Aggregates Data providers (based on Oneproviders)
- **Opendata Repository**
  - Easy sharing data for anonymous access
  - OAI-PMH for metadata harvesting
  - PID/DOI minting
- **Metadata indexing service**
- **Webportal for data discovery**
  - Pluggable portals for data discovery
- **Platform for hybrid cloud data management and processing environment**

<https://marketplace.eosc-portal.eu/services/egi-datahub>



## ONE DATA



with



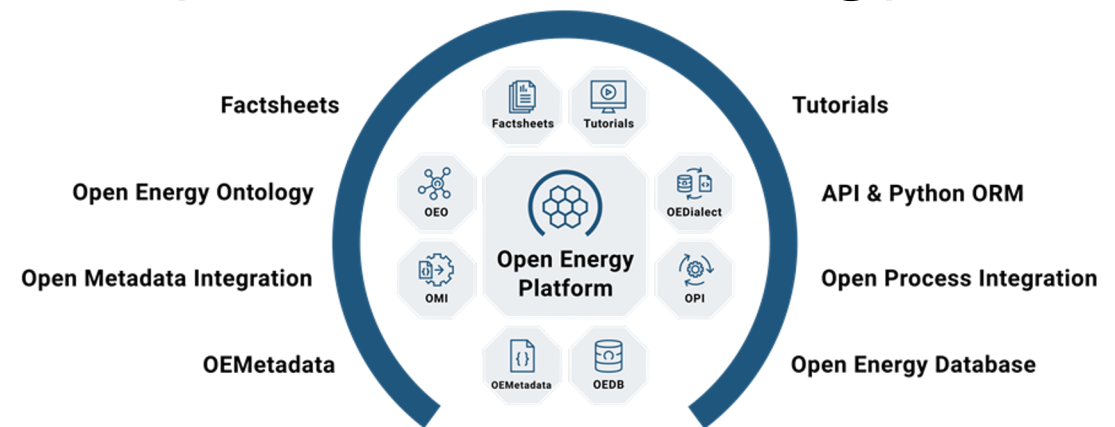
# Open Energy Platform

An open and modular framework for research data management in energy system analysis and a community database for energy data

- Cross-tier community project
  - Since 2015, funding until 2025+
  - Several research projects
  - 10+ institutions and universities

- Code available at GitHub

- Transparent development process
- Existing: Community database, API, Factsheets, Data Review ...
- Outlook: Data Versioning, Scenario Explorer, LOD, Knowledge Graph



## Open Energy Family

<https://marketplace.eosc-portal.eu/services/open-energy-platform>



# C-SCALE Earth Observation Data Federation

- C-SCALE: Copernicus – eoSC AnaLytics Engine
- Bringing together EO data providers
  - Sentinel collaborative ground segments (CESNET, GRNET)
  - CreoDIAS (CloudFerro)
  - National institutes (EODC, VITO)
- Keeping vast EO data archives
  - so that you don't have to do it yourselves
  - Sentinel-1, -2, -3, -5P, PROBA-V, SPOT-VGT and other



with





# C-SCALE Earth Observation Data Federation

- Providing storage capacity for your results
  - and the processing power to achieve them in the first place
  - available for EOSC users through the Virtual Access mechanism
- Bottom line: the right way to store EO data in EOSC is not at all.
  - It's already there!
- C-SCALE Earth Observation Data Federation is expected to be onboarded to EOSC Marketplace in August 2022!

C-SCALE Contact information:

<http://c-scale.eu/> | @C\_SCALE\_EU | [info@c-scale.eu](mailto:info@c-scale.eu) |

Zdeněk Šustr, CESNET <[sustr4@cesnet.cz](mailto:sustr4@cesnet.cz)>

Charis Chatzikyriakou, EODC <[charis.chatzikyriakou@eodc.eu](mailto:charis.chatzikyriakou@eodc.eu)>



with





# How to join? Service onboarding in the EOSC portal

- You can onboard your own data storage services in the EOSC portal
- Added values:
  - Promote their adoption outside your traditional user groups, reaching a wider user base
  - Get statistics about access requests and customer feedback
  - Get a free online platform where you can manage service requests, interact with users and provide support to them, and agree the most suitable service levels
  - Contribute to the definition and maintenance of EOSC service provisioning policies and the portfolio roadmap
  - Join the group of providers that meet EOSC quality standards

# How to join? Service onboarding in the EOSC portal

- Rules of participation:
  - The service is accessible to users outside its original community
  - At least one service instance is running in a production environment
  - Release notes and sufficient documentation are available
  - Helpdesk channels are available for support, bug reporting and requirements gathering
- How to do it:
  - <https://providers.eosc-portal.eu/home>
  - Register your institution as EOSC Provider
  - Add resources being offered by your institution



## Welcome to the EOSC Providers Portal

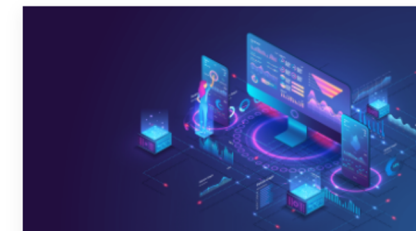
A single platform for providers to onboard their organization into EOSC portal, register and manage their resources catalogue



### Onboarding of Resources

Self service, Workflow based

[Learn more](#)



### Provider Dashboard

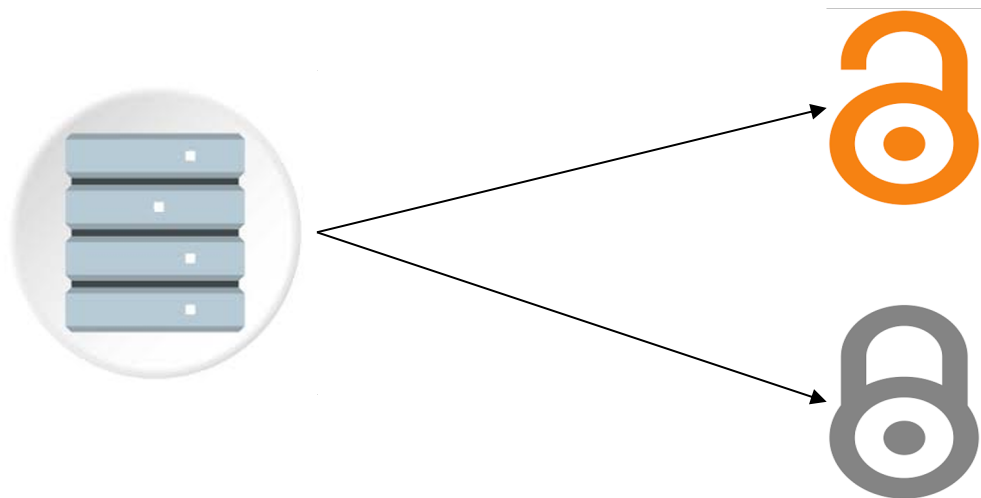
Gain full control of your resources in the catalogue

[Learn more](#)



# How can users access the services?

## Types of user access



- Direct accessible (registration might be needed)
- Link provided to the service website

- Orderable to the service provider
- Link provided to the provider web form

- You can place the order via the EOSC Marketplace by adding it (including also other services if needed ) to a project
- Requests are managed by the EOSC team and will be forwarded to the provider (useful if you would like to use service from more than one provider)

## Pricing

- Usually storage services are offered at the price
- Opportunity (now!) to have access for free until June 2023 (thanks to EC funding via the INFRAEOSCo7 projects) - For standard configurations of the services

# See you next time!

Thanks for joining us today.  
Don't forget to attend our next ask  
me anything webinar

**Software**

**01 March 2022**

Sign up at  
[www.eoscfuture.eu/events](http://www.eoscfuture.eu/events)



with



The EOSC Future, C-SCALE, DICE, EGI-ACE, OpenAIRE-Nexus and Reliance projects are funded by the European Union  
Horizon Programme calls INFRAEOSC-03-2020 and INFRAEOSC-07-2020.

