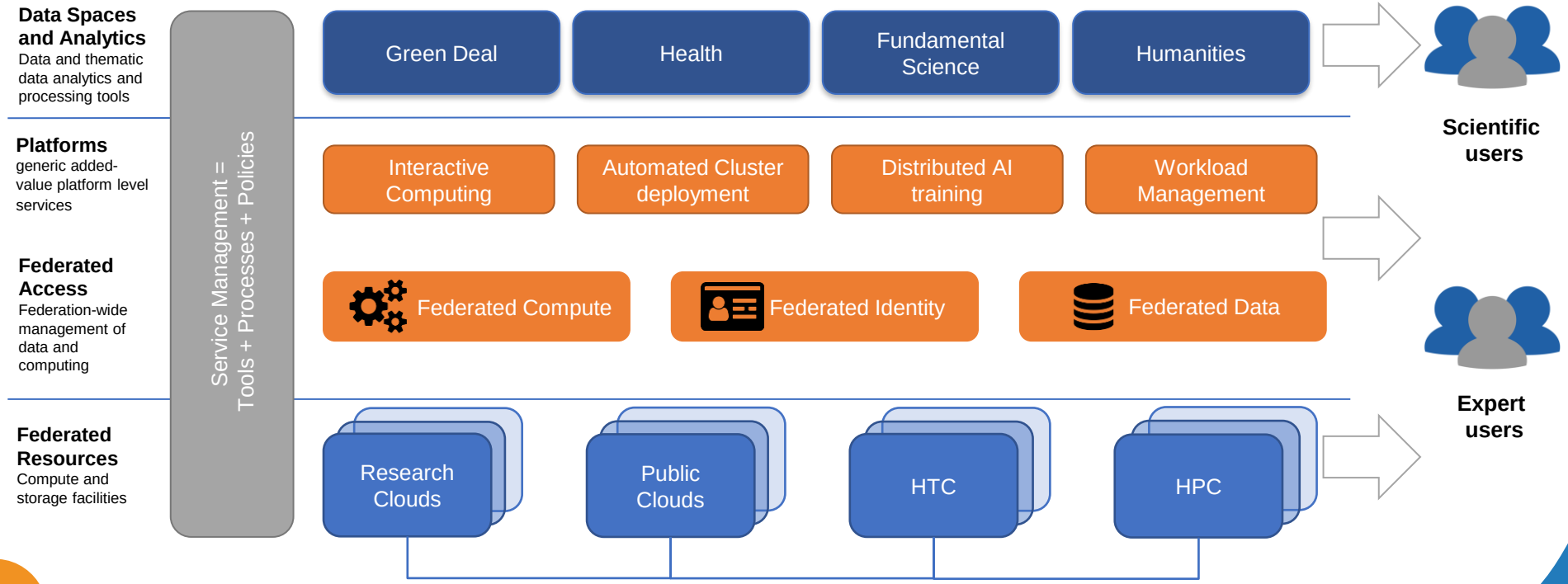
A Compute - Data Continuum for Open Science

Enol Fernández
Cloud Solutions Manager @ EGI Foundation
EGI-ACE Infrastructure and Federation services coordinator

EGI Conference 2023, Poznań

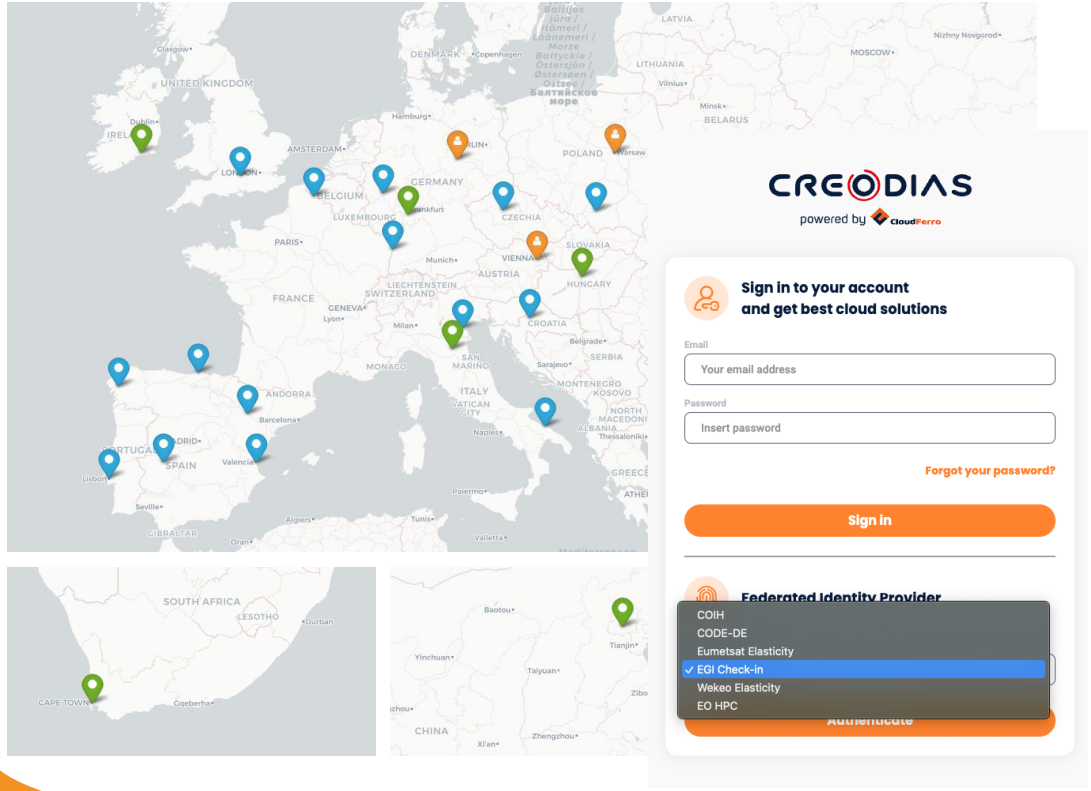
EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.

The EOSC Compute Platform



Expanding EGI cloud

New providers and integration models



7 new fully integrated providers

- *CNAF, GSI-LCG2, GRNET-OpenStack, ILIFU, CNIC, WALTON, ELKH-Cloud*

New integration points for enabling Open Science

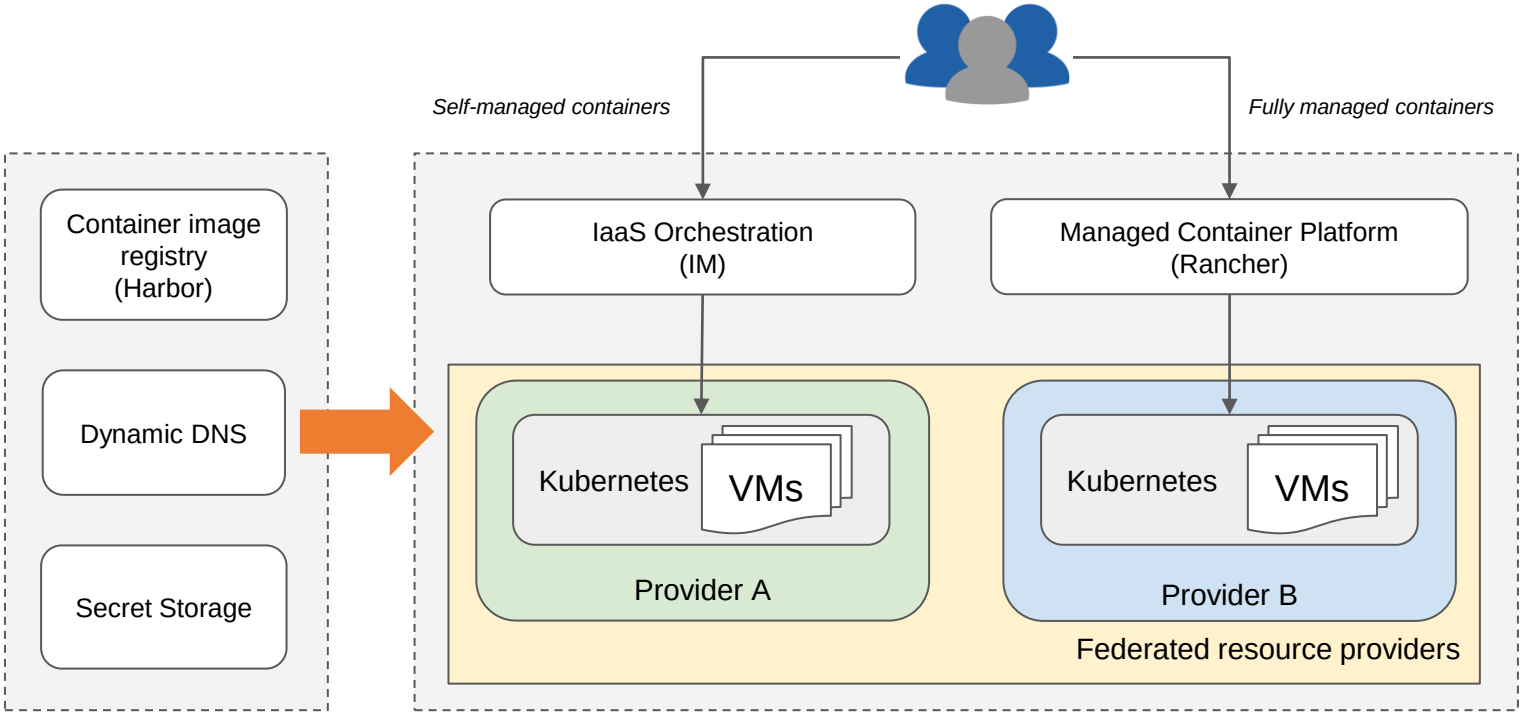
- *AAI*
- *Application sharing - AppDB*
- *Data sharing - DataHub*

Expansion beyond EU

- *South Africa, China, Eastern countries*

Container support

Two options to access Kubernetes



Container support

Container execution

HTC

Tokenization process

Transition from X.509 certificates to tokens:

- *From low-level authentication/authorisation based on GSI to a plugin based architecture capable of handling different token formats*
- *From VOMS to Check-in as Attribute Authority*
- *Started with compute services*
  - First implementations tested in HTCondor-CE and DIRAC
  - ARC-CE testing coming soon
  - Next: grid storage services

Once completed users will no longer need to hold a certificate to access HTC services

HPC

New resource types in the federation

*Pilot **HPC integration** into the federation:*

- *Access with **Check-in** credentials*
- *Container execution with **udocker***
- *Data sharing via **DataHub***
- *Accounting and monitoring using already existing tools in the federation*

4 sites and use cases with different approaches:

- **HPC as a Service** - build HPC-like environments on demand using cloud techniques
- **OpenID Connect access** to login nodes - ssh-oidc
- **HTC middleware** - reuse existing tools
- **Higher-level gateways** - access via JupyterHub

Data Management services in EGI-ACE

Existing EGI portfolio services

EGI Data Transfer: Schedules and manages data transfers between facilities.

EGI DataHub: An integrated platform for data federation, remote access, caching

New community services

openRDM: Integrated data management service with interactive access via Notebooks. Digital notebook, Data management, Inventory management

Spider: Meant for processing of big data, supporting several storage backends

Rucio: Delivers large scale and policy-driven data management in a distributed infrastructure.

Expanding the computing capabilities of EGI

INDIGO-PaaS Orchestration cloud resources using TOSCA standard with automatic selection of providers..

DODAS On-demand Distributed and customisable data analytics platform

DEEPaaS Automated training of ML/AI models on the distributed infrastructure

EC3 Deploy Virtual Elastic Clusters on the Cloud

IM A cloud orchestrator to deploy and configure complex virtual infrastructures

Replay Reproduce and share research on a notebooks-based platform

Enhancing the existing compute services

DynamicDNS Friendly hostnames for services/VMs running on EGI **AppDB** FAIR software repository

Workload Manager Manage and distribute your computing tasks in an efficient way while maximising the usage of computational resources.

Notebooks Create interactive documents with live code, visualisations and text

Software Distribution Publish and access software efficiently across multiple sites

Beyond EGI-ACE

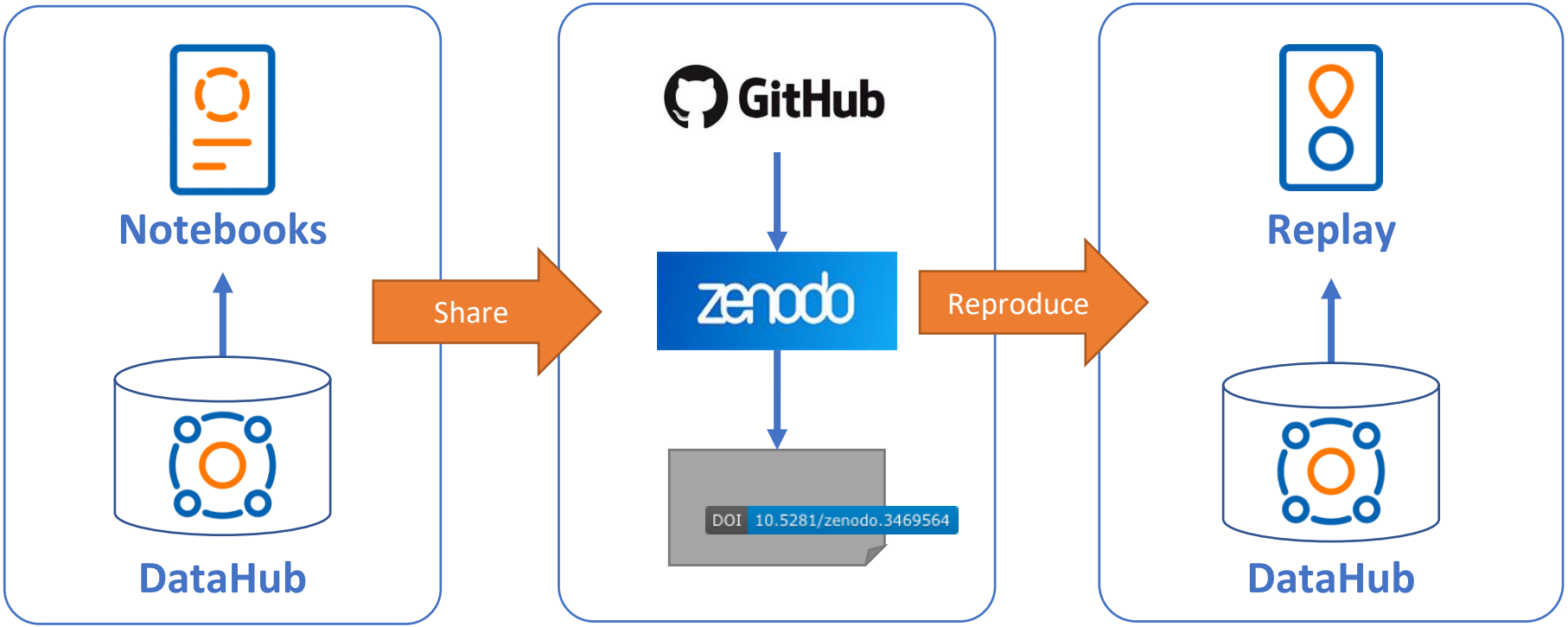
Consolidating the EGI-ACE services into the EGI portfolio

Software Distribution

Replay

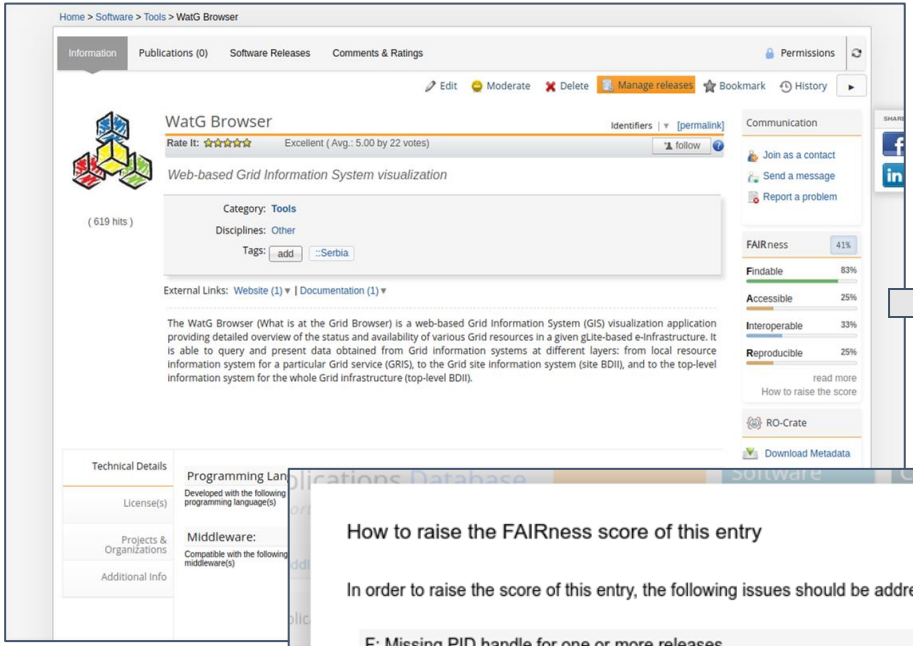
Infrastructure Manager

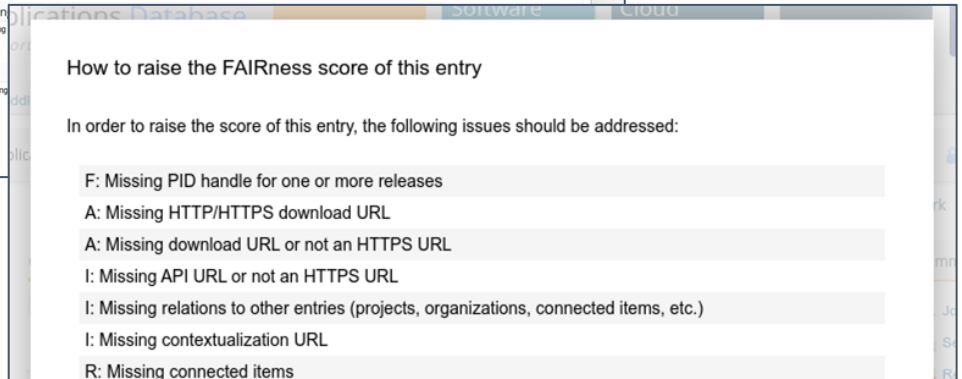
Reproducible Open Science in EOSC

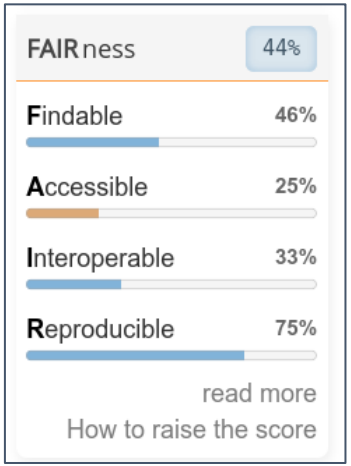


AppDB and FAIR: FAIR scoring

Supporting FAIR research software

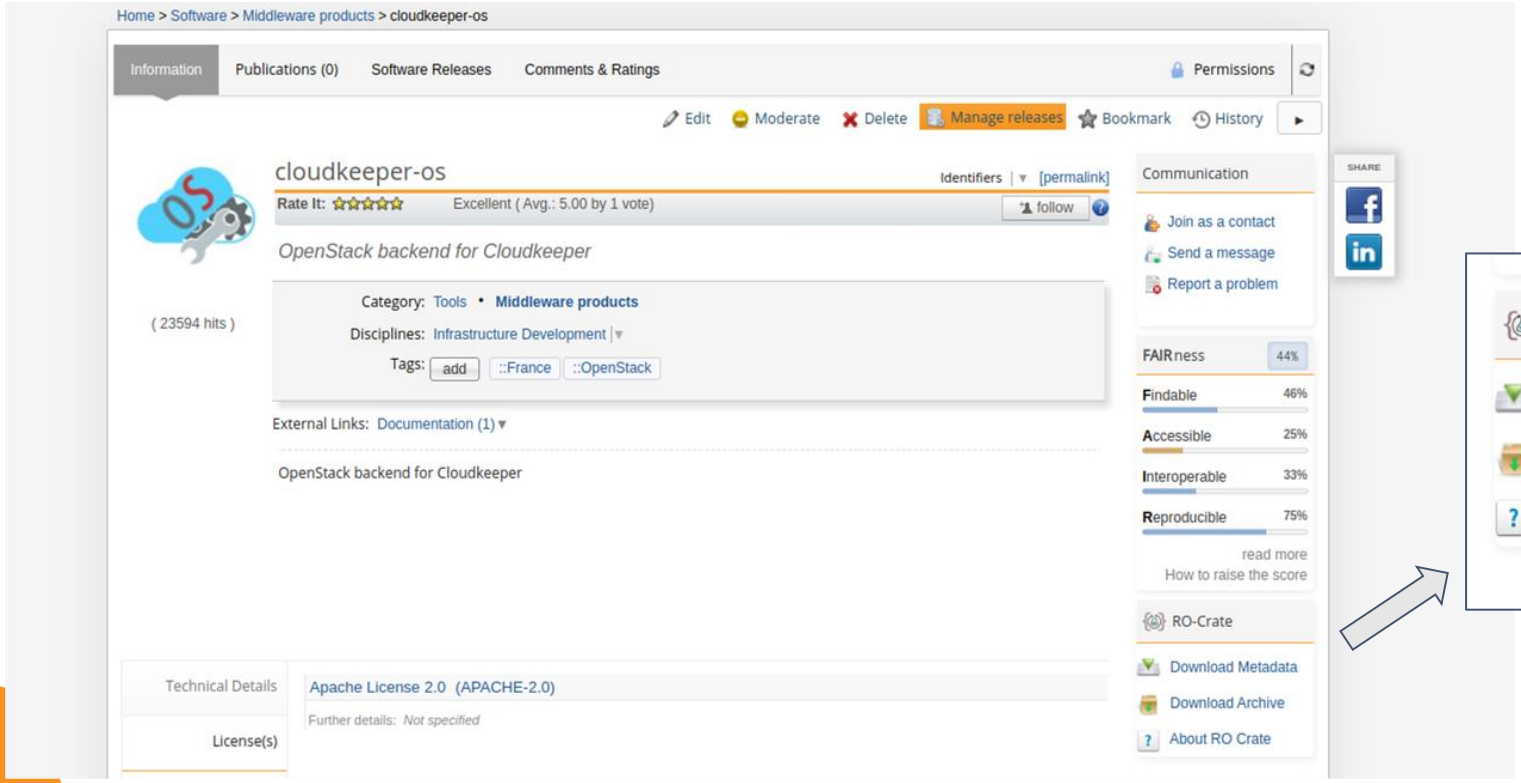


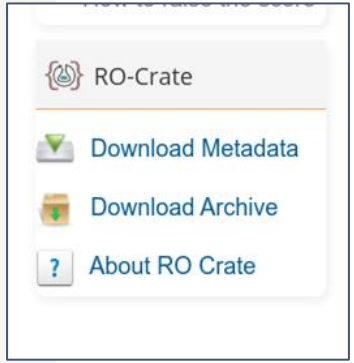




AppDB and FAIR: RO-Crate

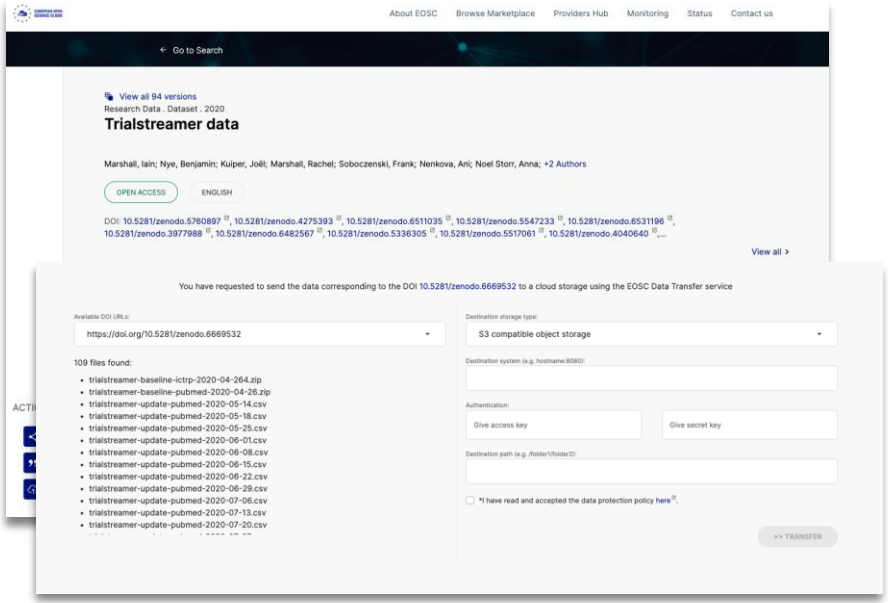
FAIR packaging of research software metadata with RO-Crate

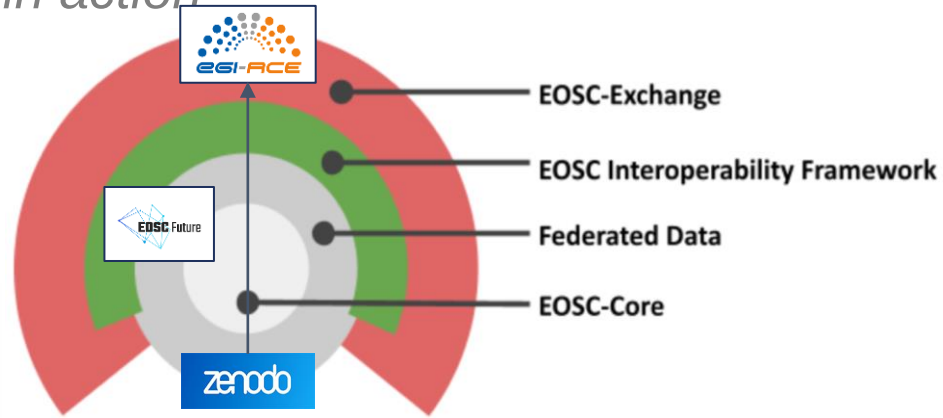




EOSC Data Transfer

EOSC Interoperability Framework in action





Collaboration with OpenAIRE-Nexus in the context of EOSC-Future:

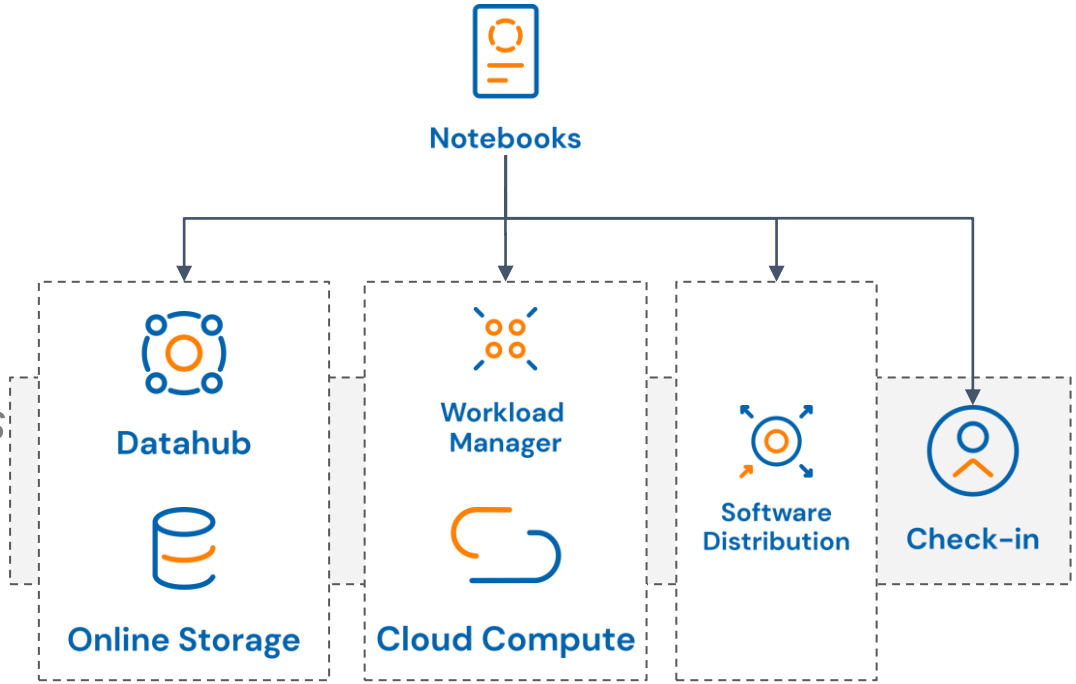
- ***Data Transfer GUI*** *integrated with EOSC Portal*
- ***EOSC Data Transfer API*** *abstracting over Data transfer services*
- ***EGI Data Transfer*** *first compatible implementation*
- *Data Staging from EOSC Data repositories to EGI-ACE providers*

EOSC Data Transfer: Architecture and Interoperability Guidelines

https://doi.org/10.5281/zenodo.8052845

An Open Platform for research

- ***Common AAI***: *use same account across all services and providers, homogeneous authorisation*
- ***Composable***: *use services from different layers together to build new solutions*
- ***API driven***: *allow users to create complex workflows and support new scenarios*
- ***Interoperable with EOSC***: *EGI core services*



Learn more!

Check demos during the conference

This evening:

- *19:15 FedCloud client and FedCloud generic services: Dynamic DNS and Secret management*
- *19:40 OpenRDM: FAIR research data management*
- *20:05 FAIR Software in the EGI Infrastructure with AppDB*
- *20:30 Effortless service deployment on distributed cloud environments through the PaaS Orchestrator*

Wednesday:

- *10:30 EGI & C-SCALE: Notebooks for Earth Observation*
- *10:30 EGI Workload Manager Service: activities in the EGI-ACE project*

Thank you!

Contact: **egi-ace-po@mailman.egi.eu**
Website: **www.egi.eu/projects/egi-ace**

EGI Foundation

@EGI_eInfra

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.

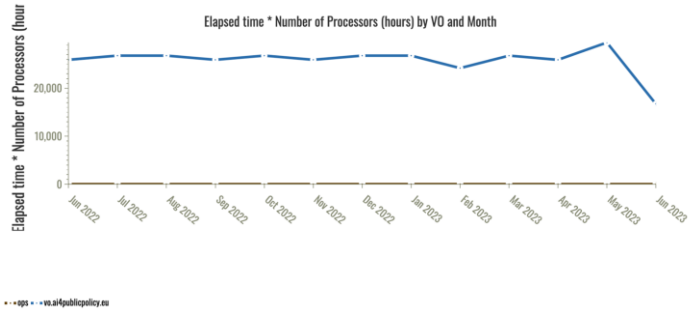
Two examples: ELKH-Cloud & CloudFerro

ELKH Cloud

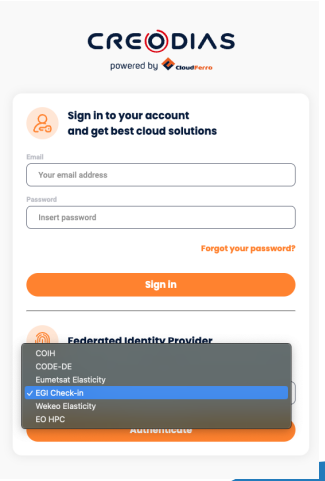
For Hungarian researchers from the Eötvös Loránd Research Network (ELKH) AND their international collaborators.

AI4PublicPolicy

OpenBioMaps
sustainable diversity



CloudFerro

Commercial cloud provider focused on Earth Observation, expanding their user base with Check-in.

C-SCALE