TELESPAZIO
a LEONARDO and THALES company

EarthCODE reproducible science platform, influenced by EOEPCA+ workflow and data standards

FOSS4G:UK South West 2024

Richard Conway and Garin Smith

Tuesday 12th November 2024

EOEPCA
BETTER ACCESS TO EARTH OBSERVATION

esa

Context – Exploitation Platform

Transforming Data to Actionable Information – value-adding

Virtual analysis environment

Data

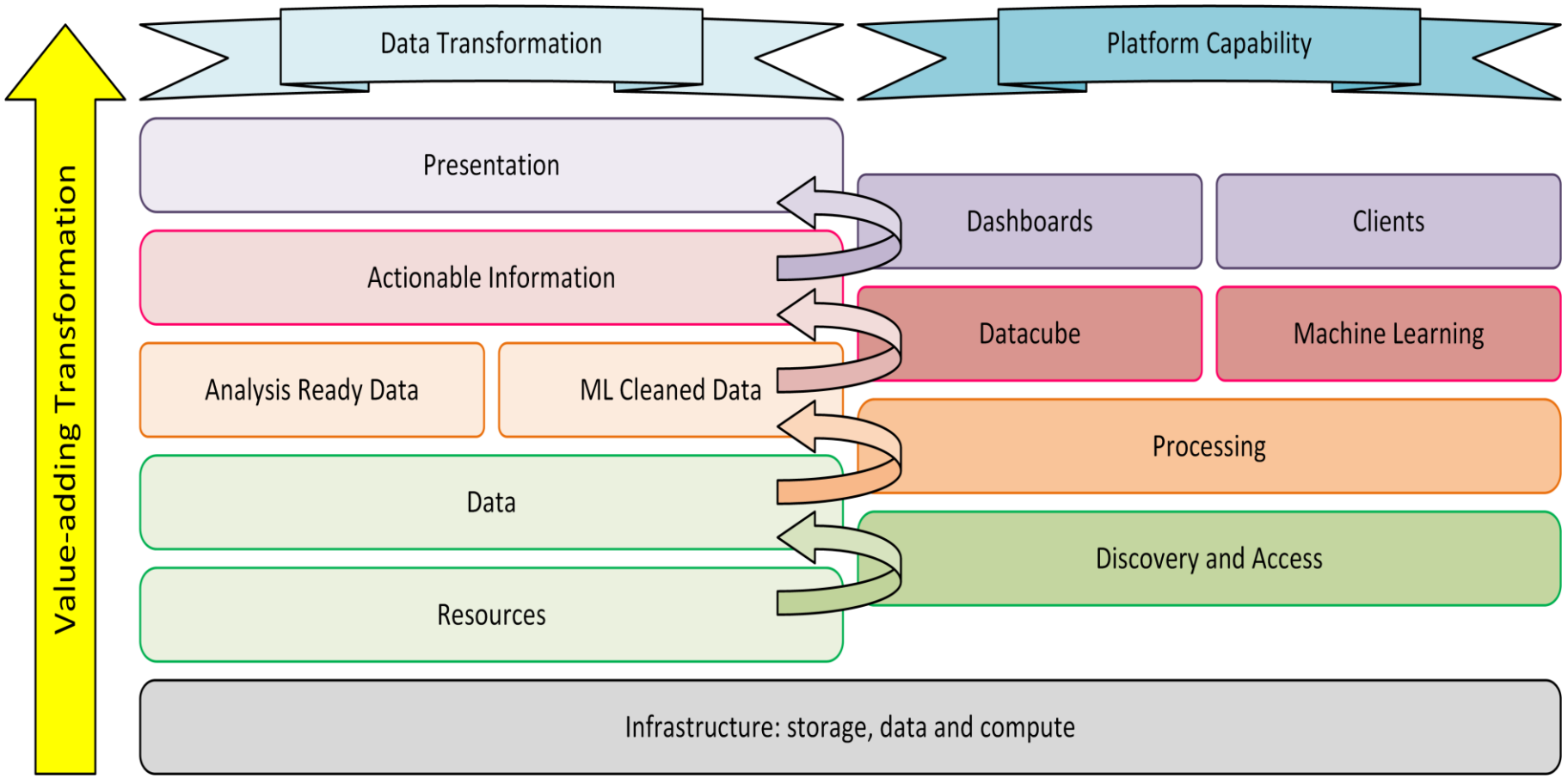
Compute

Tooling

Collaboration

Sharing

Publishing



© 2024 Telespazio

Aims and Objectives

Problem

Many platforms in a fragmented ecosystem

Difficult for users to exploit their complementary offerings

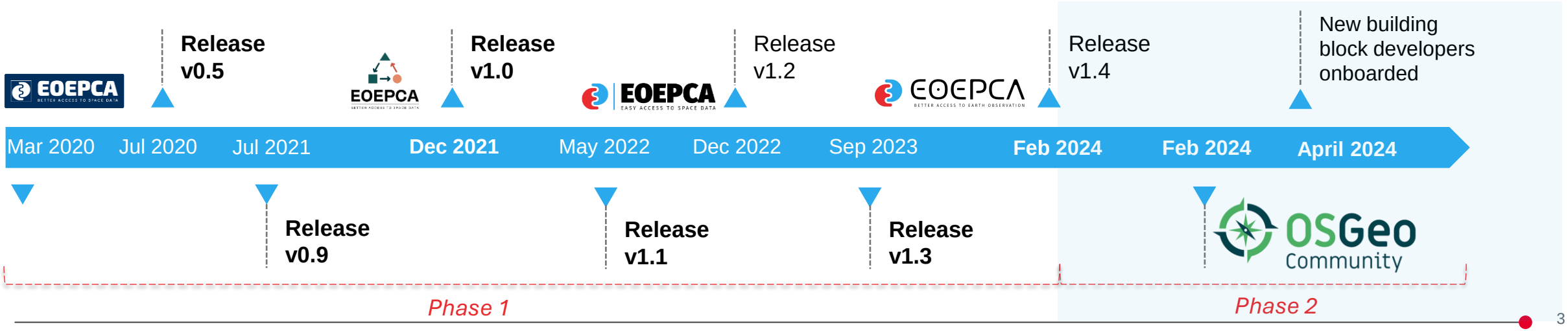
Our Approach

Common Architecture

- **Open Standards**
- Enabling **Federation** among EO cloud platform offerings
- Promote and develop **Interoperability** standards

Reference Implementation

- **Open Source**
- Avoid further fragmentation
- **Reusable Building Blocks**
- Reduce development costs



© 2024 Telespazio

Architecture

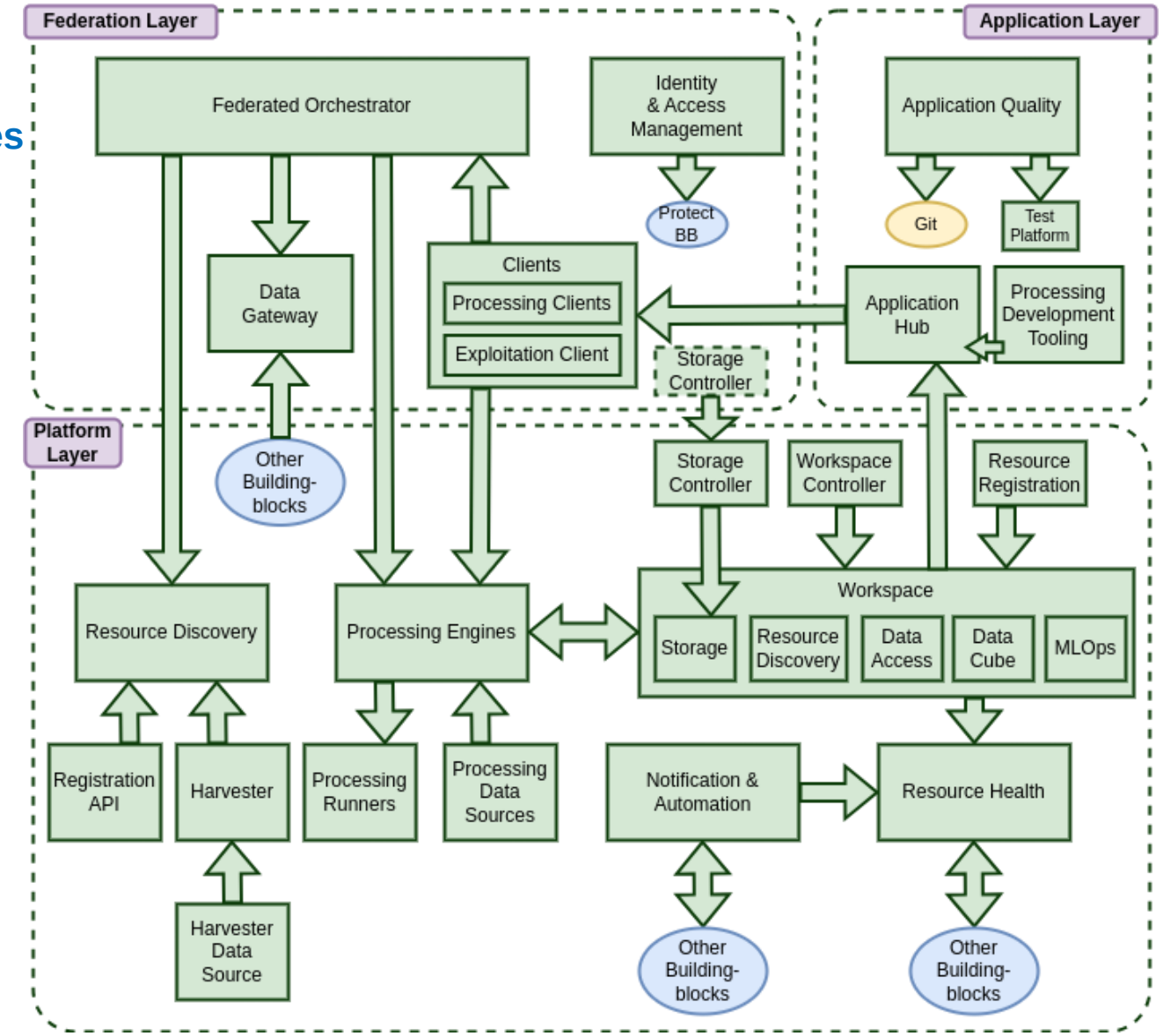
Reference Implementation of the Common Architecture defined by Building Blocks with open standard interfaces

What is a Building Block (BB)?

- Software component that implements a specific platform capability
- Typically provides a service interface (REST API) -> standards
- Dedicated helm chart for each BB – for Kubernetes deployment
- Designed to be used on their own, or in combination as a system
- Open Source

Community Oriented

- Open invitation to engage
  - Use cases and Requirements Definition
  - Co-design and Co-development
  - Adopter
- OSGeo Community project
- OGC – Working Group participation and Testbeds



© 2024 Telespazio

User-defined Processing – Processing BB

Bring the processing to the data

Two Approaches

General Purpose Processing

- **OGC API Processes**
- Containerised workflows
- **OGC Application Package**
- Suited to batch processing

Client-oriented Semantics

- **openEO Process Graphs**
- Python, R, JavaScript clients
- Close data integration
- Emerging support for Application Packages

OGC Application Package

- **OGC Best Practice**
- Portable processing workflows
- Common Workflow Language
- Steps: Containerised algorithms

Workflow Runners

- Backend workflow execution
- Extensible integrations…
  - Kubernetes
  - HPC
  - Dask, …

zoo

GeoTrellis

© 2024 Telespazio

Platform Resources

stac-fastapi

resources

pycsw

Tiler

datasets

eoAPI

TiPg

pgstac

elasticsearch

Resource Discovery for Reproducible Open Science

- **Resource Discovery BB**
- Workflows, datasets, notebooks, ML models, training data, source repo, documentation, …

Data Discovery & Access

- **Data Access BB**
- Discovery - STAC
- Retrieval and Visualisation - OGC Features, Maps, Coverages

Analysis Ready Data

- **Datacube Access BB**
- Pixel-based access to multi-dimensional data
- **OGC GeoDataCube API**
- Consolidated API for data discovery, access and processing

pygeoapi

Flowable

Ingestion & Harvesting

- **Resource Registration BB**
- API for adding resources of all types
- Data harvesting from external sources

User/Team Resources

- **Workspace BB**
- Collaboration for users and teams
- Storage and shared services
- Integrates with other BBs

vcluster

EOxHub

File Browser

© 2024 Telespazio

User Analysis and Exploitation

SharingHUB GitLab jupyterhub DVC STAC

Machine Learning

- **MLOps BB**
- Model Training & Management
- Training data management

Interactive Analysis

- **Application Hub BB**
- Notebooks for Scientific Storytelling
- Dashboards to showcase research outcomes

Application Best Practice

- **Application Quality BB**
- Best practices for Reproducible Open Science
- Static analysis
- Performance tuning

Grafana

Event-driven Behaviour

- **Notification & Automation BB**
- Automated behaviour
- E.g. systematic processing
- Building Block decoupling

Operational Outcomes

- **Resource Health BB**
- Added-value Outcomes
- **Datasets, workflows, applications**
- Automated monitoring and alerts

sonarqube

APACHE JMeter™

Apache Airflow Gatling LOCUST pytest

© 2024 Telespazio

Platform Federation

Federated Workflows

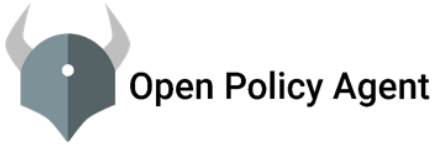
- **Federated Orchestrator BB**
- Cross-platform workflow execution
- Hybrid workflows, combining
  - OGC API Processes
  - Application Packages
  - openEO Process Graphs

Abstract Data Access

- **Data Gateway BB**
- Data source abstraction
  - Data access protocol
  - Authn / Authz
- Python library
- Extensible data providers

Federated User Identity

- **IAM BB**
- Single Sign-on
- External Identity Provider integration



© 2024 Telespazio

EOEPCA+: Partner Organisations

DLR

EOX

EOFARM

development SEED

Resource Discovery

Resource Registration

Data Access

Datacube Access

Workspace

TERRADUE

vito

GeoLabs

eurac research

Matthias Mohr Softwareentwicklung

Processing

Application Hub

Federated Orchestrator

Notification & Automation

NOVASPACE

Communications & Outreach

13 Building blocks (and counting) by a variety of developers

werum SOFTWARE & SYSTEMS

IAM

CS a Sopra Steria company

MLOps

Data Gateway

spaceapplications SERVICES

SENSmetry

Application Quality

Resource Health

© 2024 Telespazio

EOEPCA+: Partner Organisations

DLR

EOX

EOFARM

development SEED

TERRADUE

vito

GeoLabs

eurac research

Matthias Mohr Softwareentwicklung

NOVASPACE

Communications & Outreach

Resource Discovery

Resource Registration

Data Access

Datacube Access

Workspace

Processing

Application Hub

Federated Orchestrator

Notification & Automation

werum SOFTWARE & SYSTEMS

IAM

CS a Sopra Steria company

MLOps

Data Gateway

spaceapplications SERVICES

SENSmetry

Application Quality

Resource Health

EarthCODE

© 2024 Telespazio

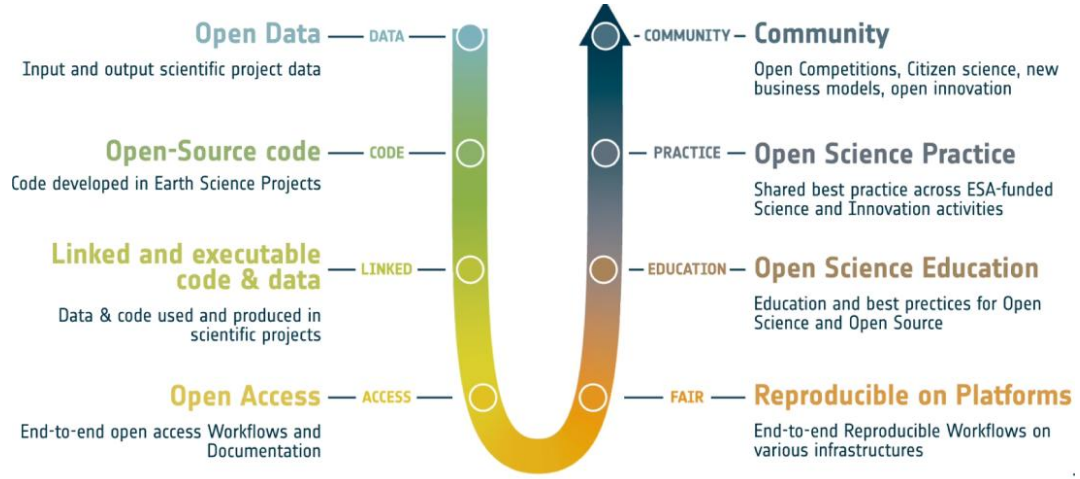
EarthCODE

FutureEO Independent Science Review 2022 recommendations:

More visibility, discoverability and useability of science results, leveraging interoperability and Open Science Tools, and improving communication and community building.

Scientific community feedback at Science Strategy workshop (2023):

ESA to support FAIR and reproducible Open Science practices and complementing the scientific process by Software Development best practice.



Earth Science Collaborative Open Development Environment (EarthCODE) responds to these recommendations

© 2024 Telespazio

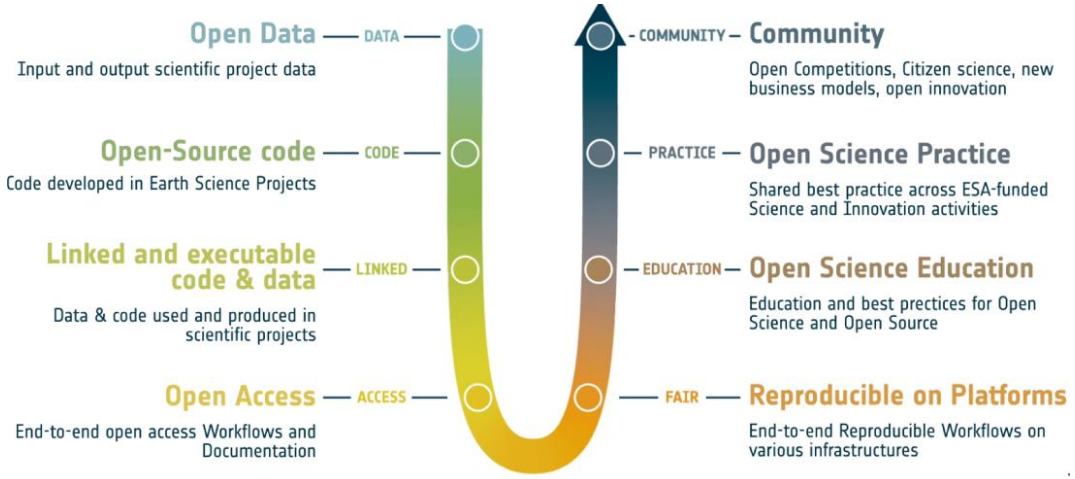
EarthCODE

FutureEO Independent Science Review 2022 recommendations:

More visibility, discoverability and useability of science results, leveraging interoperability and Open Science Tools, and improving communication and community building.

Scientific community feedback at Science Strategy workshop (2023):

ESA to support FAIR and reproducible Open Science practices and complementing the scientific process by Software Development best practice.

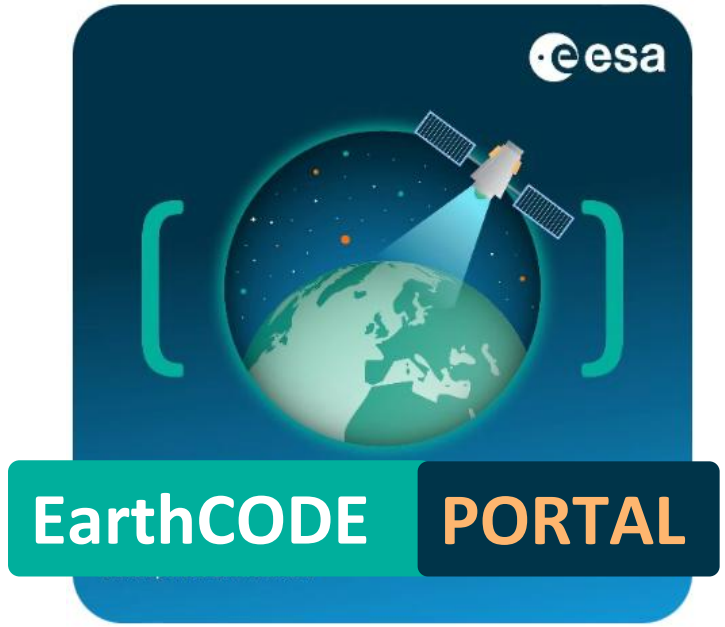


Overall aim of the EarthCODE portal is to help users create scientific workflows for their experiments and publish them to a repository after they have been verified, so that experiments can seamlessly be re-used with the specified data and necessary infrastructure according to FAIR Open Science Principles, engaging with the relevant communities to further science.

© 2024 Telespazio

EarthCODE Strategy

- EarthCODE is designed to provide Earth system scientists with computational resources, data management tools, and open science support.



Contributor

user

Catalogue Resources Manage Community

Catalogue

- **Science Data** → Visualise on a web map / import in a dashboard
- **Methods & Workflows** → View / Modify / Use / Execute in a platform
- Papers & Documentation

NoR Sponsorship

Resources

- **EO Platforms** → Develop new methods
- **FAIR Management Tools** → Package your algorithm as a FAIR workflow
- Learning resources

NoR Sponsorship

Manage

- As a PI, manage your project outcomes

Publish your data and workflow

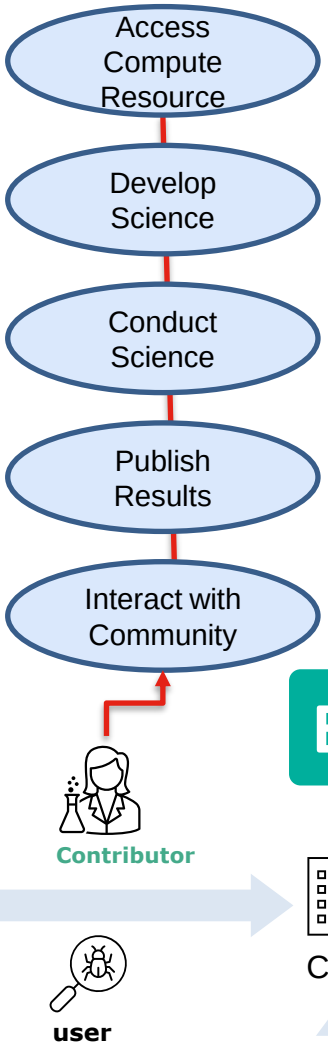
Community

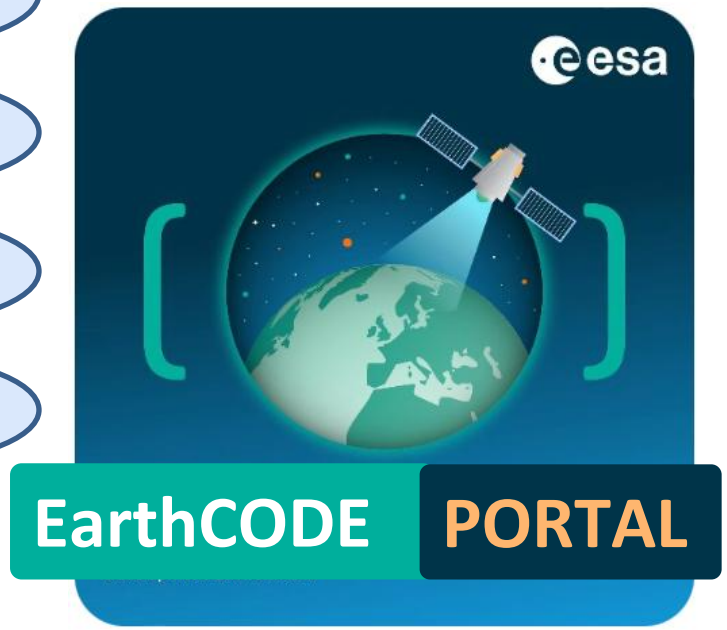
- Discuss on the forum about datasets, methods, etc

Long-term storage in Project Results Repository on **ESA Infrastructure**

© 2024 Telespazio

EarthCODE Use Cases





Catalogue
- **Science Data** → Visualise on a web map / import in a dashboard
- **Methods & Workflows** → View / Modify / Use / Execute in a platform
- Papers & Documentation

NoR Sponsorship

Resources
- **EO Platforms** → Develop new methods
- **FAIR Management Tools** → Package your algorithm as a FAIR workflow
- Learning resources

NoR Sponsorship

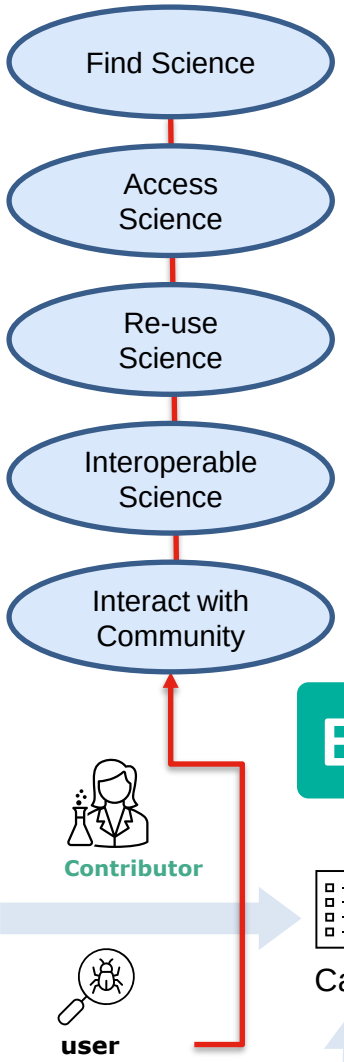
Manage
- As a PI, manage your project outcomes → Publish your data and workflow

Community
- Discuss on the forum about datasets, methods, etc

Long-term storage in Project Results Repository on **ESA Infrastructure**

© 2024 Telespazio

EarthCODE Use Cases



Catalogue

- **Science Data** → Visualise on a web map / import in a dashboard
- **Methods & Workflows** → View / Modify / Use / Execute in a platform
- Papers & Documentation

NoR Sponsorship

Resources

- **EO Platforms** → Develop new methods
- **FAIR Management Tools** → Package your algorithm as a FAIR workflow
- Learning resources

NoR Sponsorship

Manage

- As a PI, manage your project outcomes

Publish your data and workflow

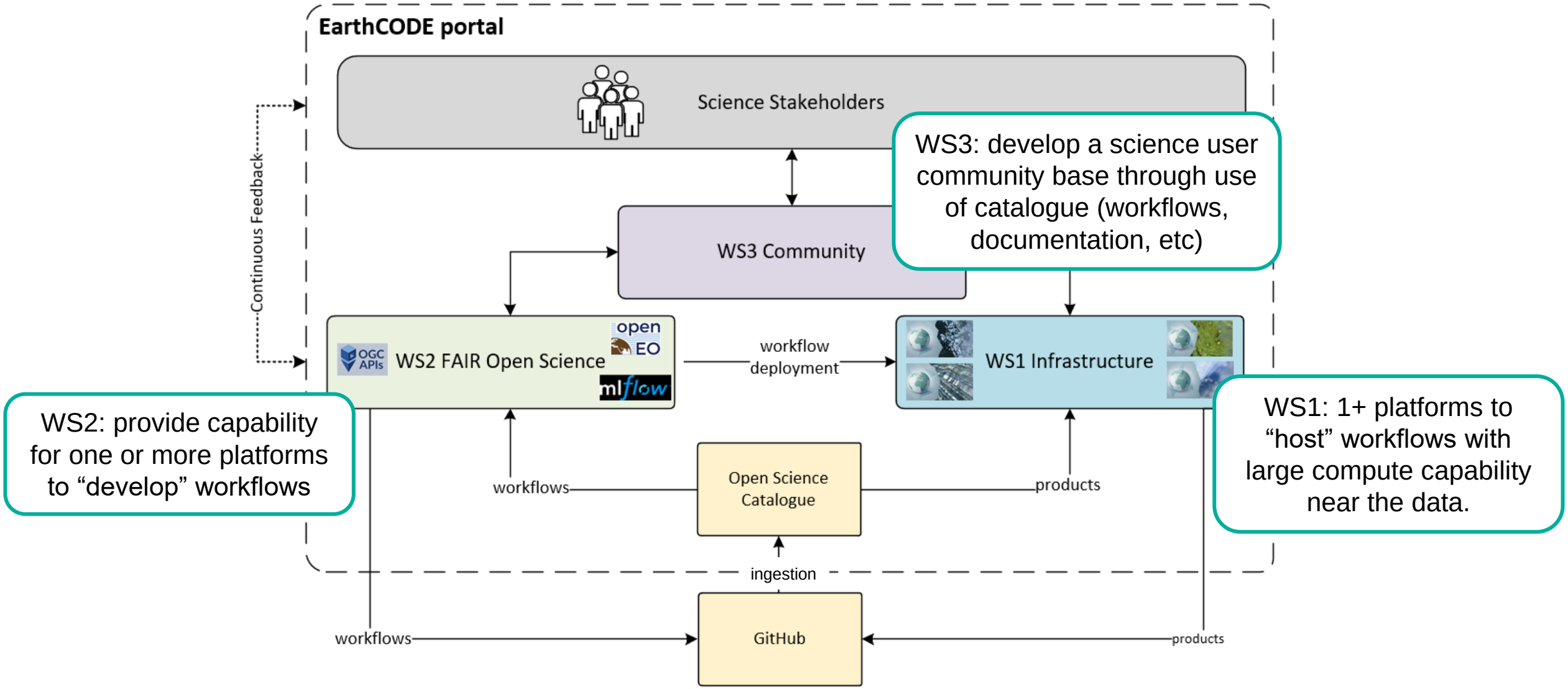
Community

- Discuss on the forum about datasets, methods, etc

Long-term storage in Project Results Repository on **ESA Infrastructure**

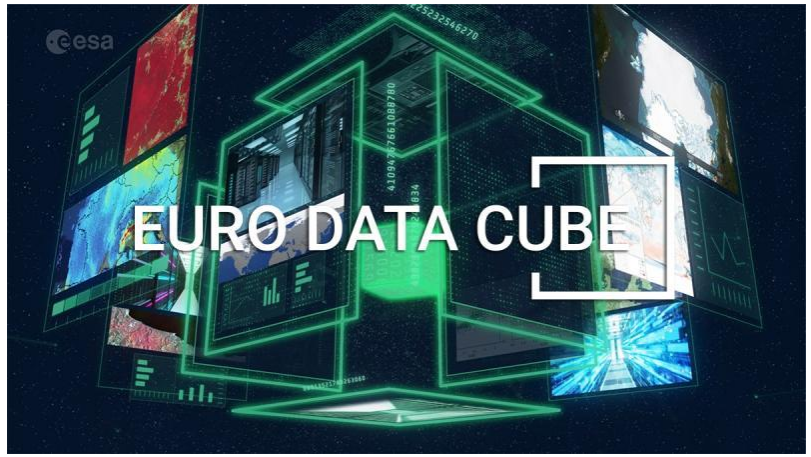
© 2024 Telespazio

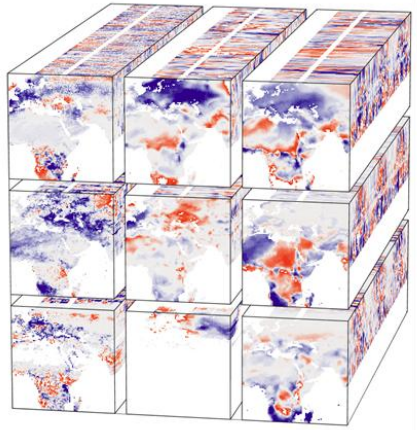
Conceptual Architecture (in support of reproducibility and interoperability)



© 2024 Telespazio

Platforms and Partners





EARTH SYSTEM DATA LAB

EOX

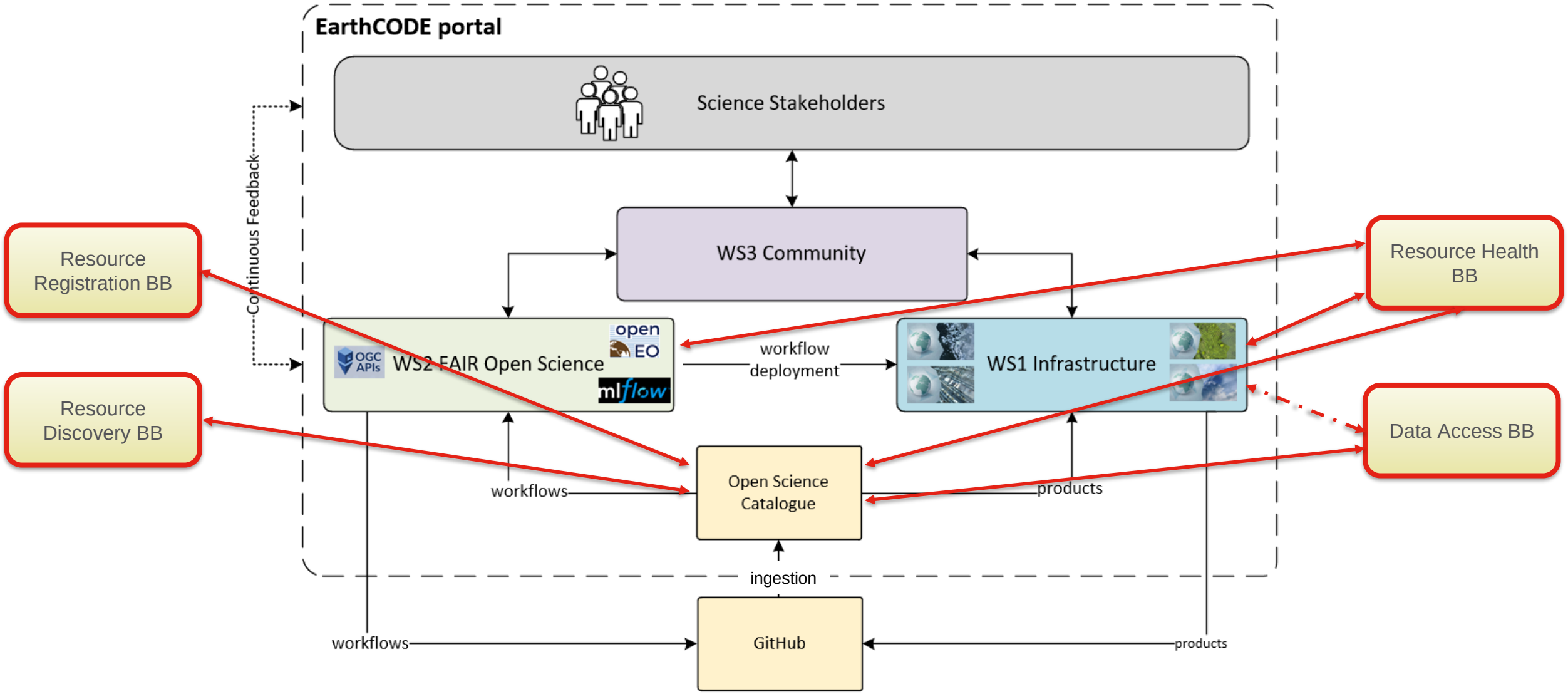
vito Remote Sensing

BROCKMANN CONSULT

Best practise ITT is planned for 2025 to allow more platforms

© 2024 Telespazio

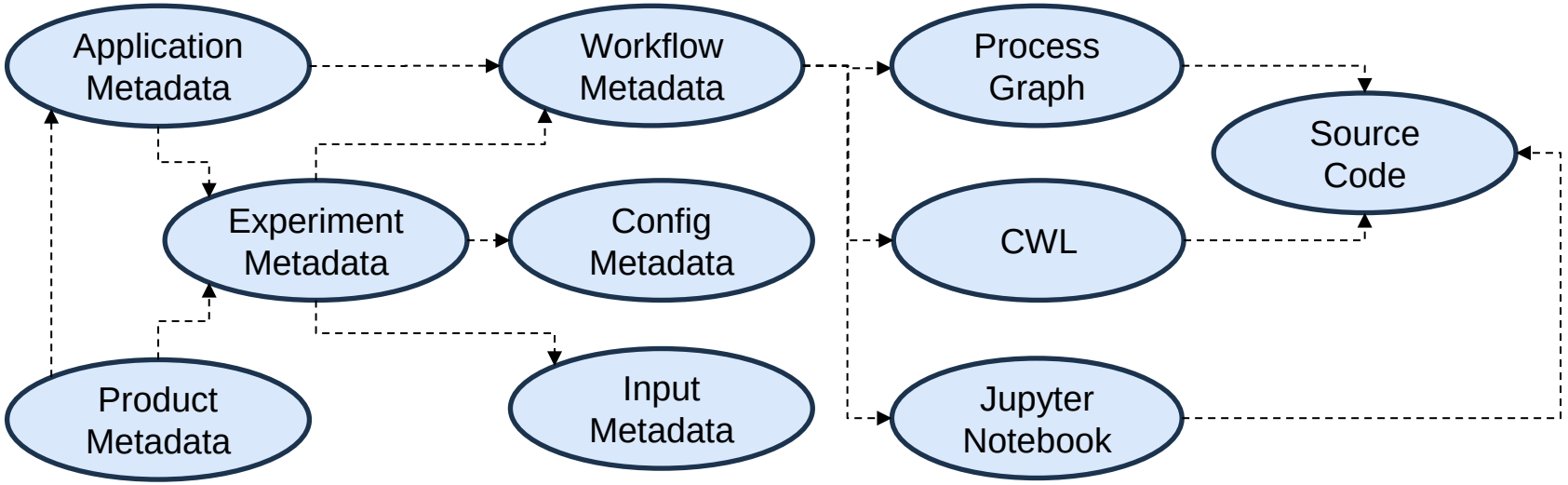
Conceptual Architecture (in support of reproducibility and interoperability)



© 2024 Telespazio

EarthCODE High-level Architecture - Metadata

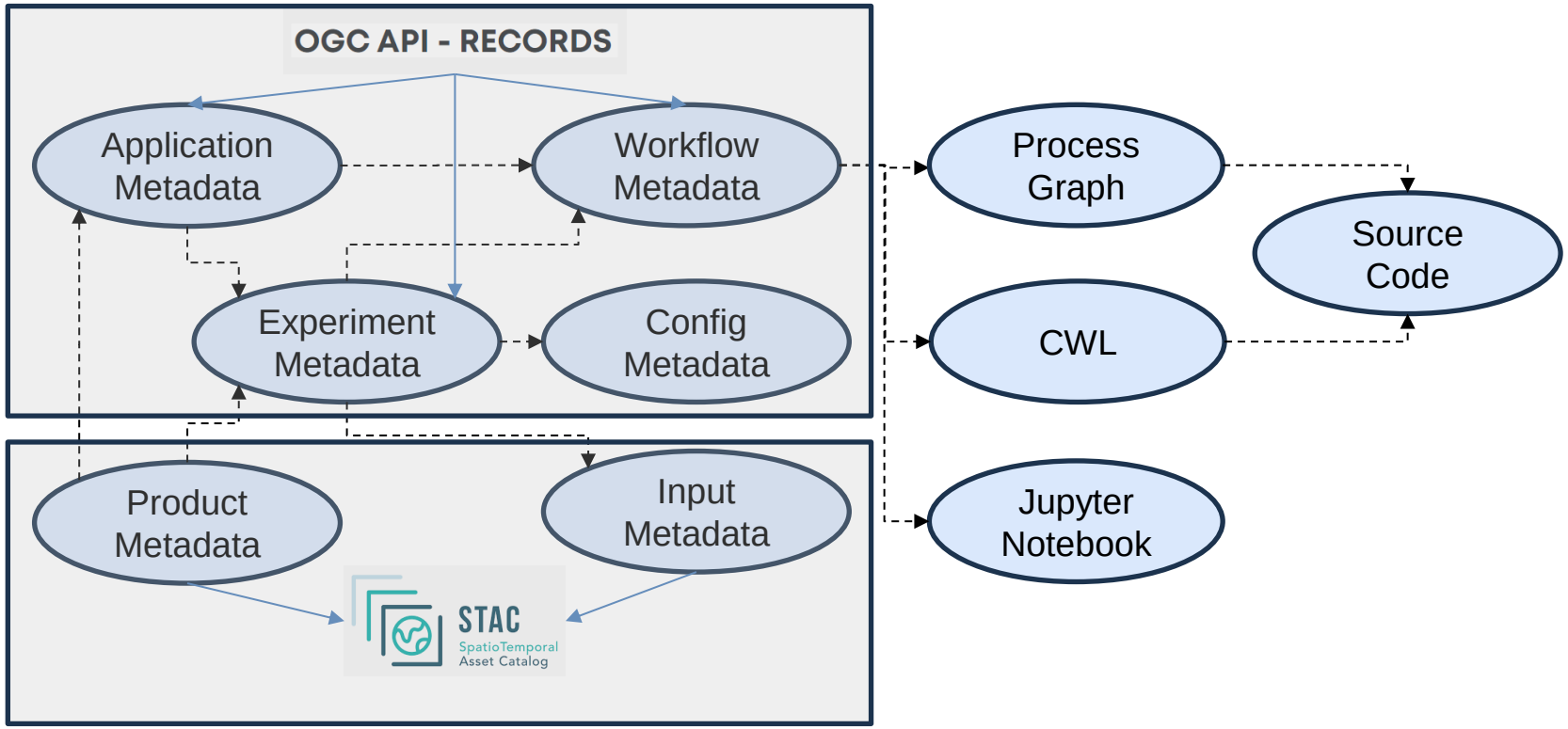
- Workflow
- Experiment
- Product



© 2024 Telespazio

EarthCODE High-level Architecture - Metadata

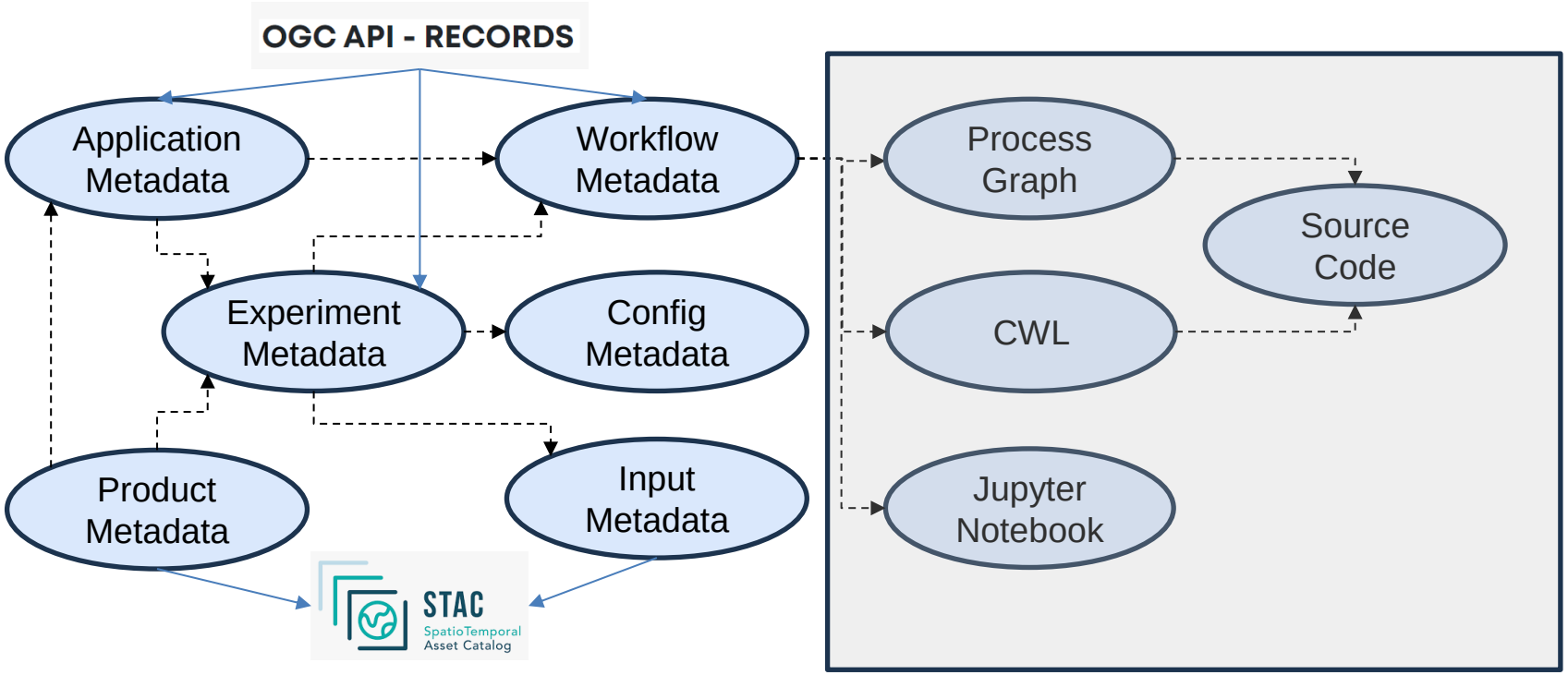
- Describing Workflows, Experiments and Applications
- Describing Products and Input Data



© 2024 Telespazio

EarthCODE High-level Architecture – Workflow Support

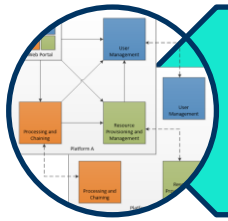
- Find novel ways to convert source code into CWL
- Find novel ways to convert source code into a Process Graph
- Find novel ways to convert source code into other workflows
- Find novel ways to manage data types when using workflows



© 2024 Telespazio

Open Source Software

EOEPCA
BETTER ACCESS TO EARTH OBSERVATION



Web Portal
https://eoepca.org/

GitHub
https://github.com/EOEPCA

Documentation
https://eoepca.readthedocs.io/

EarthCODE

https://earthcode.esa.int/ - EarthCODE Portal

© 2024 Telespazio

TELESPAZIO
a LEONARDO and THALES company

esa

THANK **YOU**
FOR YOUR ATTENTION

telespazio.co.uk