Computing and Data Projects

digital Twins

Jorge Gomes

jornadas.fccn.pt

Digital
Twin

A virtual replica designed to accurately reflect real world objects, systems or processes.

- a high-fidelity model of a system which can be used to emulate the actual real system

A digital twin goes beyond mere simulation:

- is much a richer virtual environment that mimics the physical one
- enables study and prognosis of multiple processes
- can be feed with real data

Apollo Simulators installed at NASA - Front - Lunar Module Simulator (green)
Mid- Mission Effects Projector-Lunar surface (green)
Back - Command Module Simulator (tan)
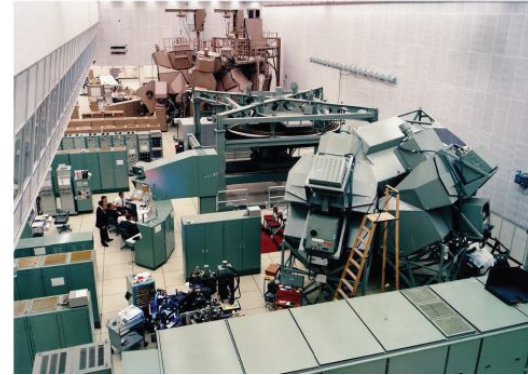
Apollo 13

The ground simulators saved the mission.

High fidelity models enable to explore solutions and predict results.

The procedures carried out in space were first tried out on earth, and rejected if the simulators showed it to be dangerous or impractical
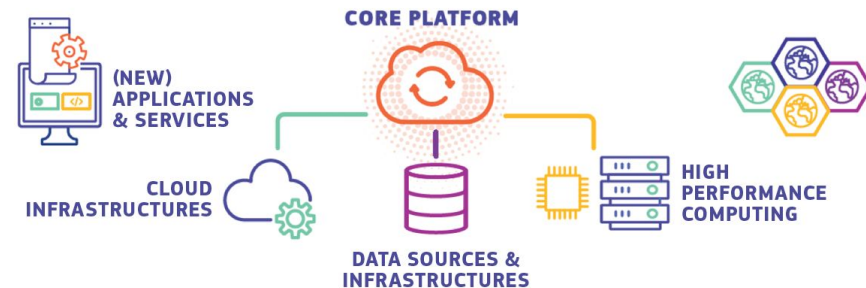
Destination Earth

Is a flagship initiative of the European Commission to develop a highly-accurate digital model of the Earth (**a digital twin of the Earth**).

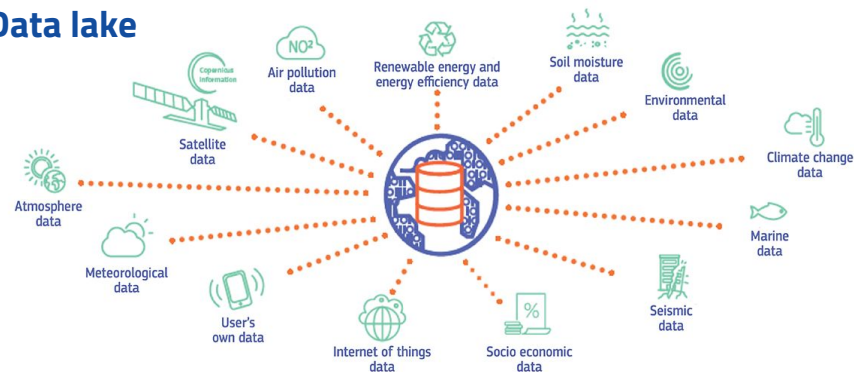Understand and simulate the evolution and behavior of the Earth system components.

Digital modelling of the Earth system.
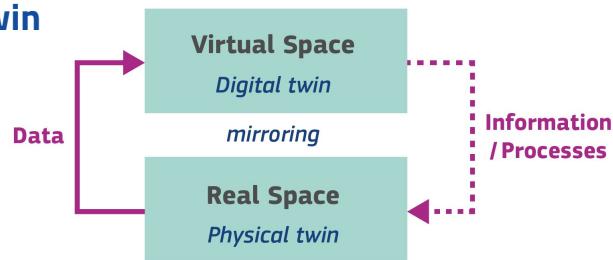
Provide evidence-based decision-making tools.

Based on an open, flexible, and secure cloud-based computing infrastructure with access to data and HPC.
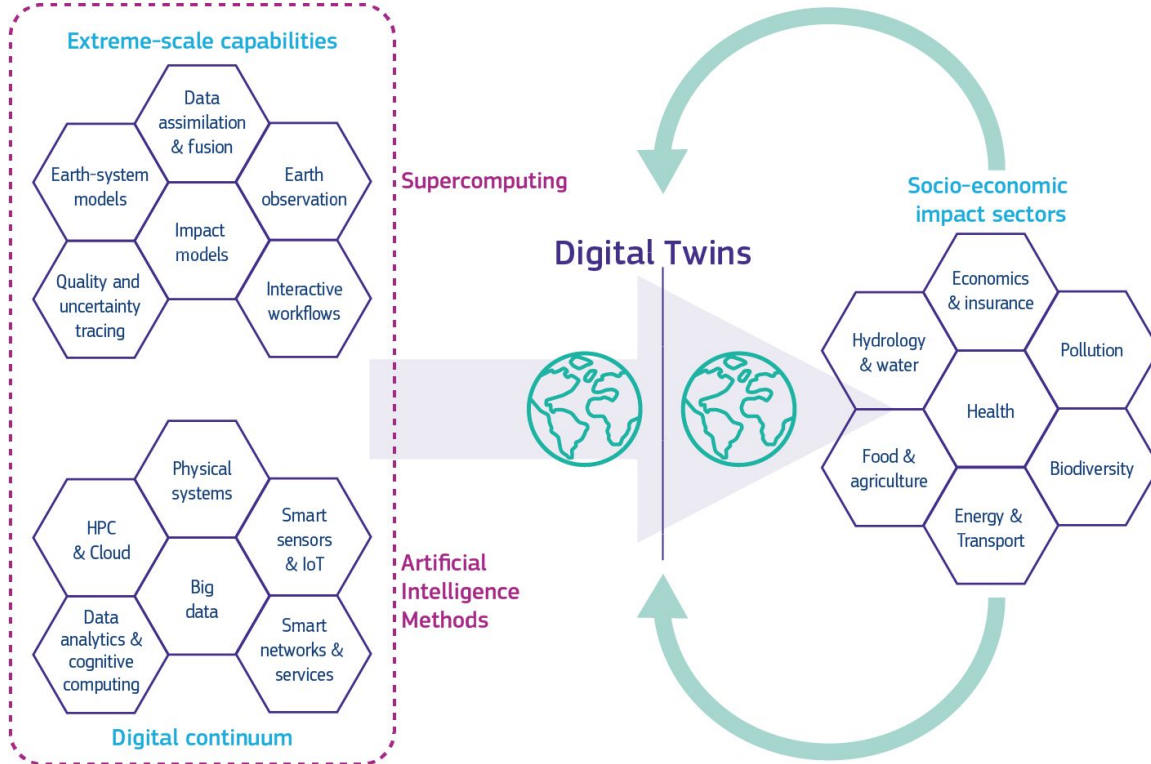

CORE PLATFORM

(NEW) APPLICATIONS & SERVICES

CLOUD INFRASTRUCTURES

DATA SOURCES & INFRASTRUCTURES

HIGH PERFORMANCE COMPUTING

Data lake

Copernicus information

Air pollution data

Renewable energy and energy efficiency data

Soil moisture data

Environmental data

Satellite data

Climate change data

Atmosphere data

Marine data

Meteorological data

Seismic data

User's own data

Internet of things data

Socio economic data

Digital twin

Virtual Space

Digital twin

mirroring

Real Space

Physical twin

Data

Information / Processes

Destination Earth Twins

Model, monitor and simulate natural phenomena, hazards and the related human activities. Anticipating climate-related events.

Digital replicas of the Earth System in thematic areas:

- **Earth science**
- **Climate change**
- **Oceans**
- **Biodiversity**
- **Others**

Integrated into a comprehensive digital twin of the Earth system

GEOphysical extremes
DT-Geo

- **DT-Geo**
  - Digital Twin of geophysical extremes.
  - Analyse and forecast the impact of geohazards from **earthquakes, volcanoes, tsunamis and anthropogenic seismicity.**
  - For integration in **Destination Earth** Initiative.
  - Urgent computing, early warning forecast and rapid post-event assessment.
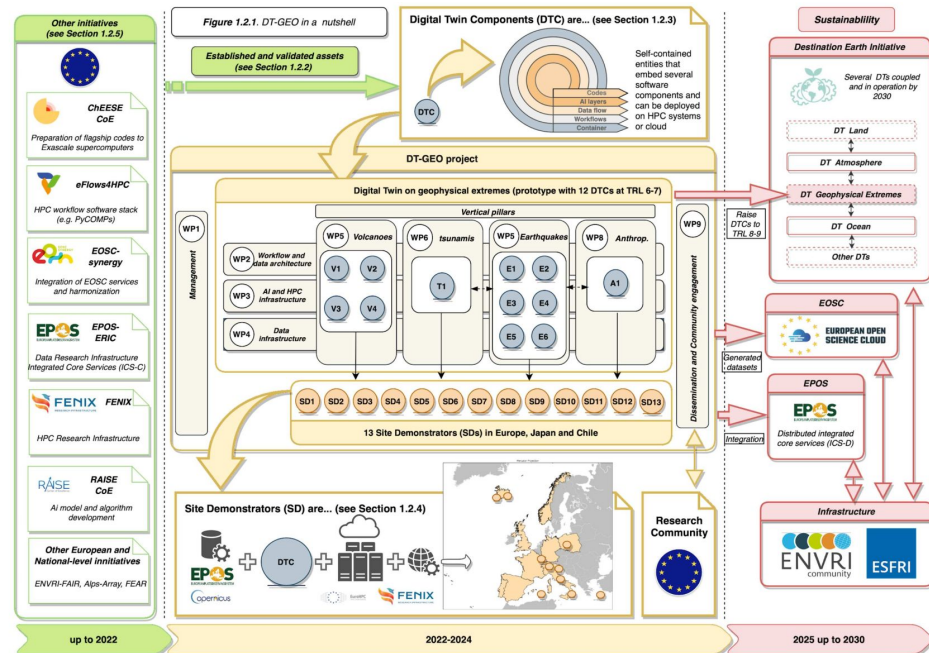 - Started in September of 2022.

- **Contribution**
  - Software and Service Quality assessment
  - Integration of workflows and containerisation
  - Providing tools for containerisation (udocker)

DT-Geo Software stack

Figure 1.2.2. DT-GEO baseline software stack ordered by categories (white boxes). The following colour code is used for the WPs (see Sec. 3.1.1).
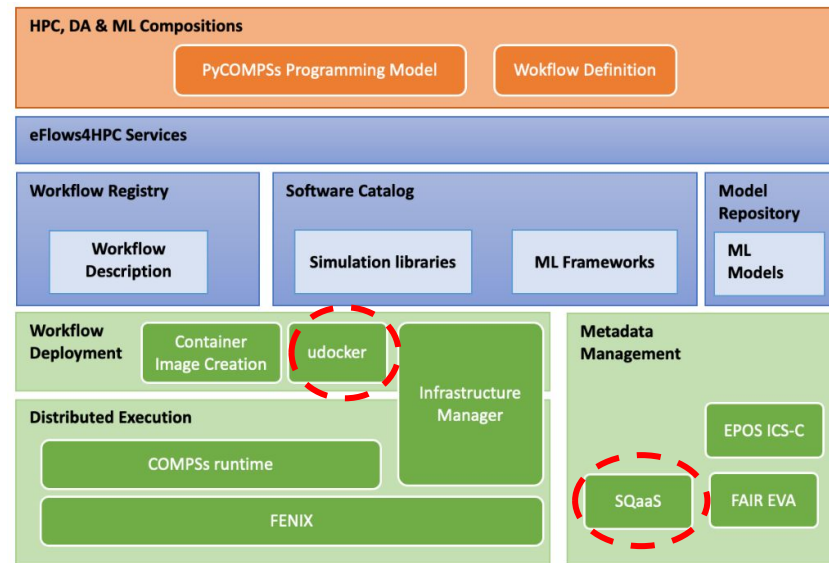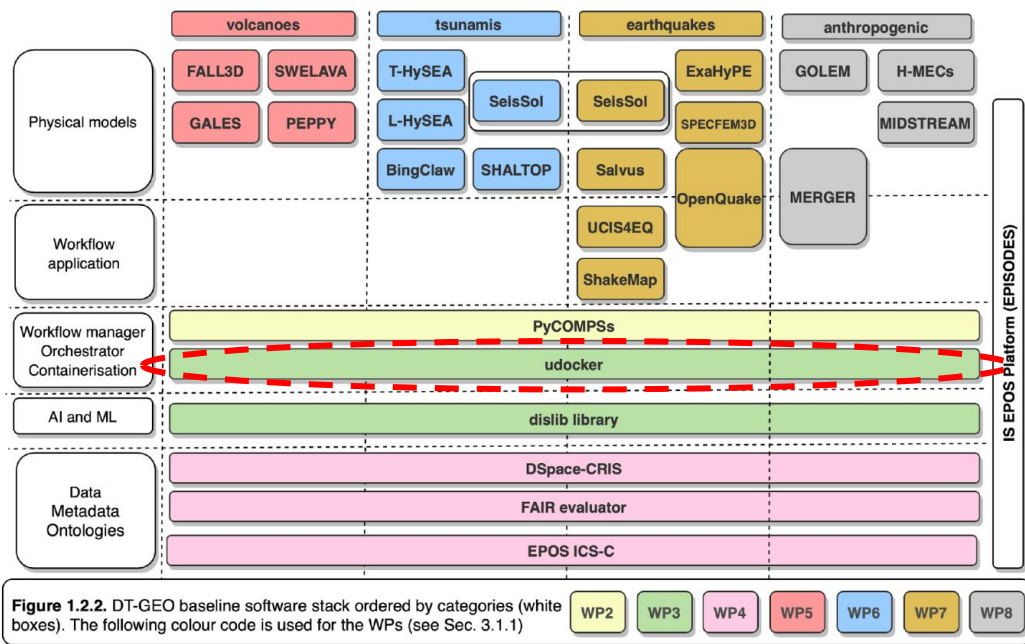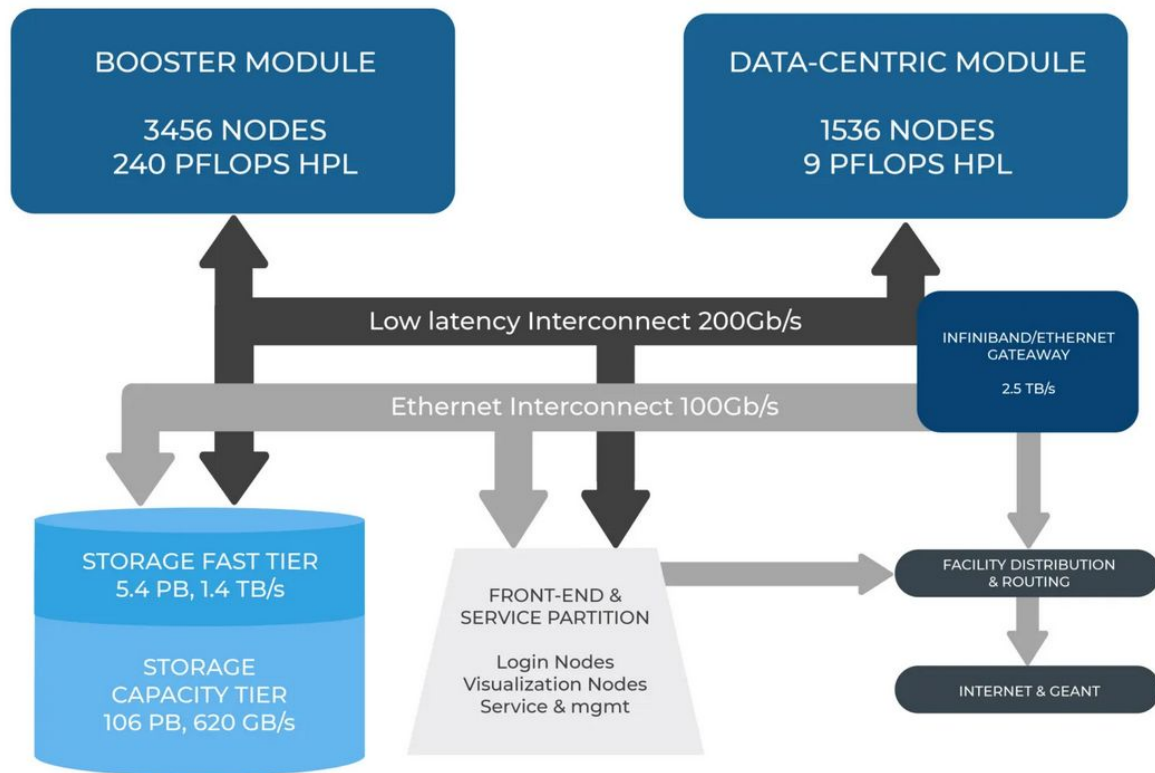
DT-Geo @ CINECA - Leonardo main components

- 1.536 CPU-based nodes
  - 172.032 cores
- 3.456 GPU-based nodes
  - 13.824 GPU
  - 110.592 cores
- 155 Racks
  - 16 CPU racks
  - 116 GPU racks
  - 12 I/O racks
  - 1 System racks
- Power Requirements
  - HPL: ~ 8.0 MW
  - Operational: ~ 6.0 MW

BOOSTER MODULE

3456 NODES
240 PFLOPS HPL

DATA-CENTRIC MODULE

1536 NODES
9 PFLOPS HPL

Low latency Interconnect 200Gb/s

Ethernet Interconnect 100Gb/s

INFINIBAND/ETHERNET GATEAWAY

2.5 TB/s

STORAGE FAST TIER
5.4 PB, 1.4 TB/s

STORAGE CAPACITY TIER
106 PB, 620 GB/s

FRONT-END & SERVICE PARTITION

Login Nodes
Visualization Nodes
Service & mgmt

FACILITY DISTRIBUTION & ROUTING

INTERNET & GEANT

GT-Geo application containerisation HPC ready
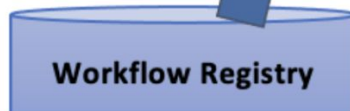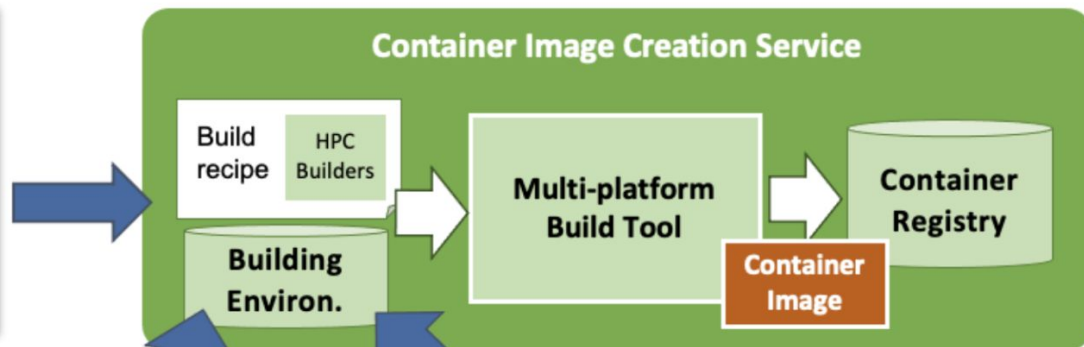
DT-Geo uses
udocker

UDOCKER

User tool to execute docker containers in user space. Developed at LIP:

- Fully user space.

- No root privileges required to use or install.

- Does not require compilation.

- Download and execution of docker containers by non-privileged users.

- Suitable for Linux batch systems and interactive clusters managed by other entities such as grid infrastructures.

- Does not require Linux namespaces.

https://github.com/indigo-dc/udocker
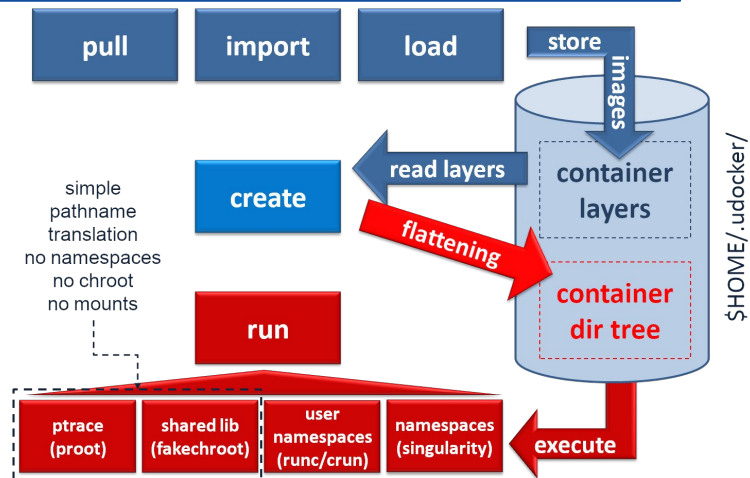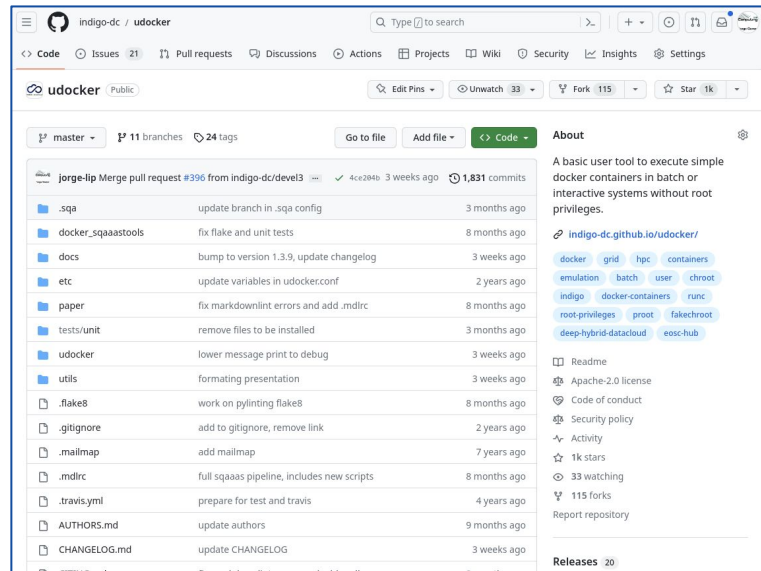
Interdisciplinary DTs
interTwin

- **InterTwin**
  - Develop a common approach to the implementation of Digital Twins (digital twin engine - DTE)
  - Applicable across the whole spectrum of scientific disciplines
  - Open-source interoperable platform
  - Software components for modelling and simulation to integrate application-specific DTs
  - Blueprint architecture for DTs
  - Liaison with Destination Earth
  - Started in September 2022

- **Contribution**
  - Software release and management
  - Quality and validation for applications, models and services

10
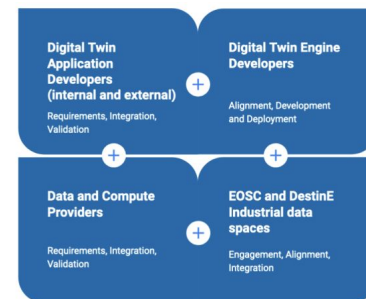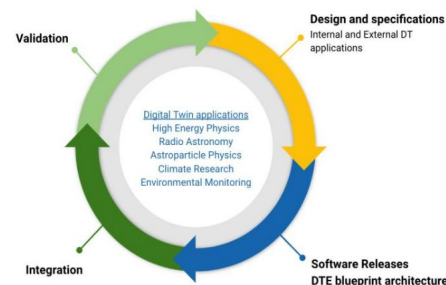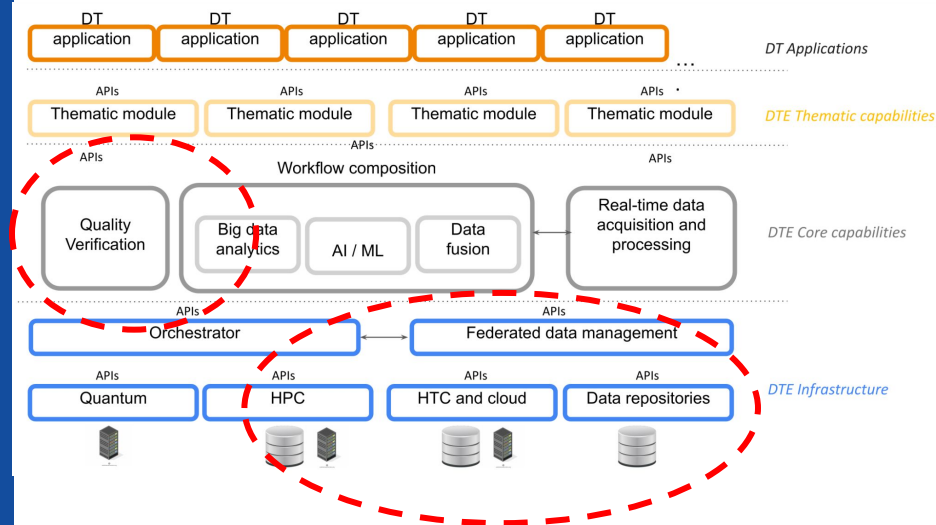
Core capabilities 1/2
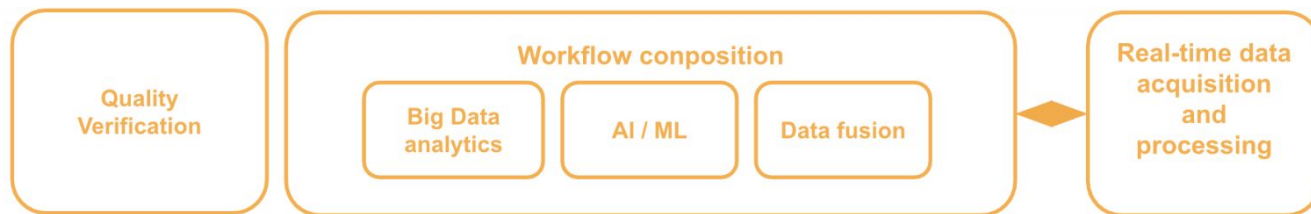
The DTE Core Modules offer horizontal capabilities to facilitate the creation and the operations of data-intensive and compute-intensive DT applications:

Advanced workflow composition: executes DT workflows that can invoke other Core module capabilities. An **interdisciplinary processing graph** will serve as a link and API for the supported workflow engines guaranteeing a **common user experience** and facilitating the integration of discipline specific tools.

Data Fusion: implements and integrates processes for merging datasets from different sources. This includes **linking of observational and modelled data**, and the **harmonization of different types of observational data** like gridded datasets with vector based datasets like point streams of data from ground stations.

19

Core capabilities 2/2
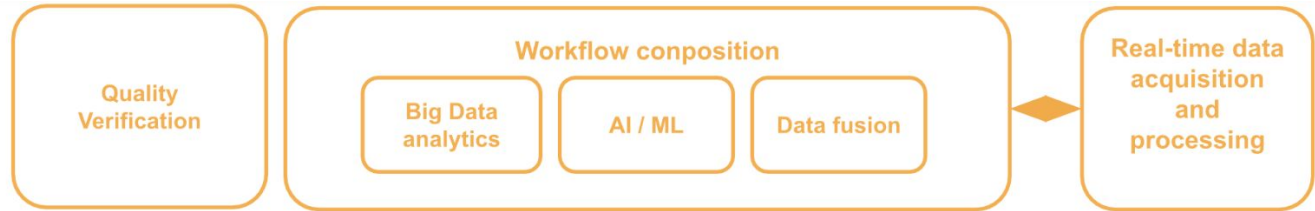
- **AI workflow and method lifecycle management**: toolbox for realizing complex AI setup

- **Real time acquisition and data analytics**: delivers high performance data ingestion by applying the paradigm of "**serverless**" computing to DTs. This module expands the advanced workflow composition module to trigger on the fly processing upon data acquisition.

- **Validation, verification, and uncertainty tracing for model quality**: toolkit that provides developers with the possibility to **design and validate DT Models guaranteeing quality and reliability of the DT applications outputs**. Offered as a "Model Validation as a Service" enabling customisations of best practices and standard quality measures for scientific disciplines and applications.

20

# Climate change impact of coastal/inland flooding events
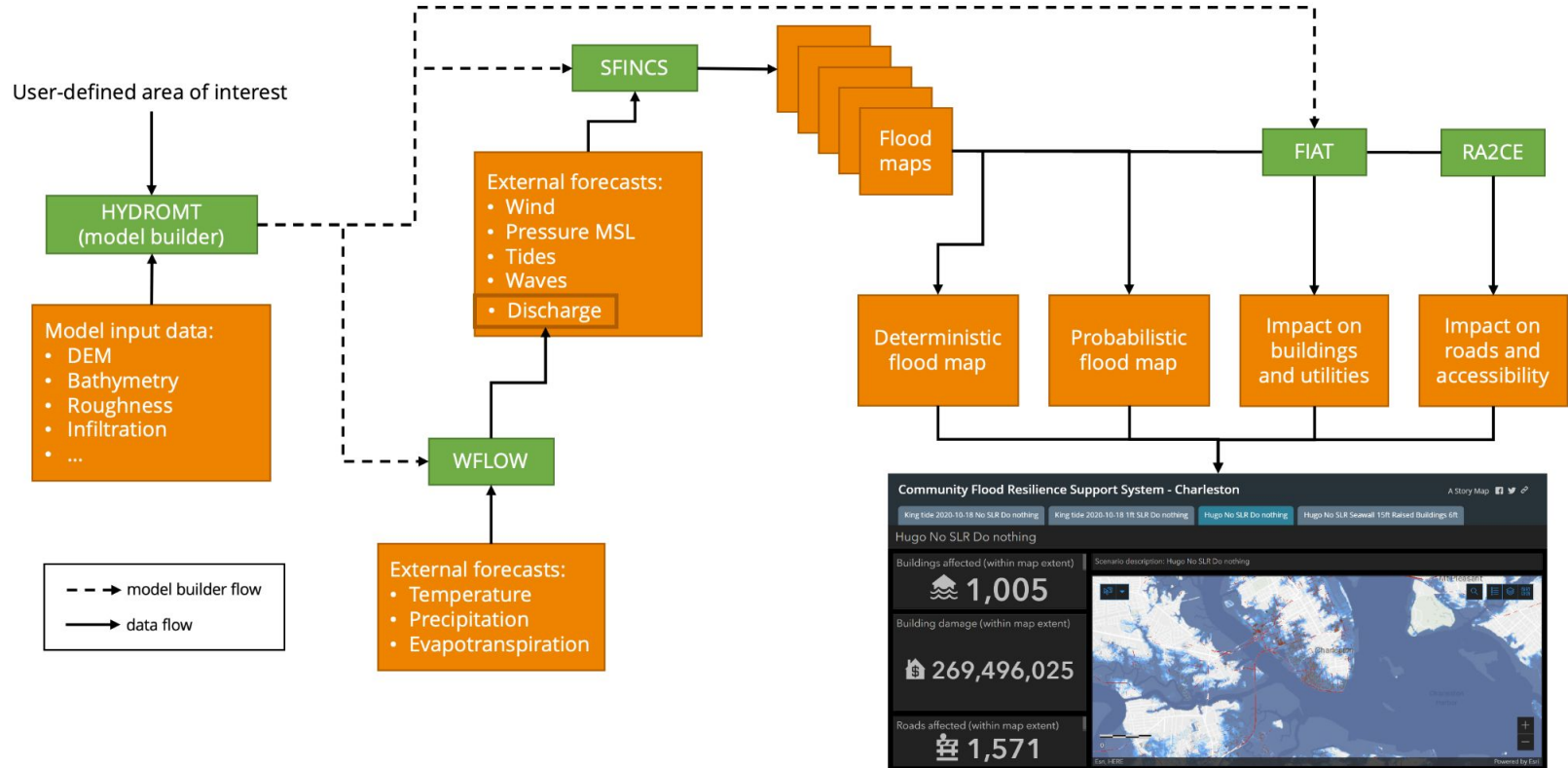
C-Scale
Copernicus

European Open Science Cloud project.

C-Scale (2021-2023) aims to federate European EO infrastructure services:

- Copernicus DIAS and others.
- Capitalise on the European Open Science Cloud (EOSC) capacity and capabilities.
- Support Copernicus research and operations with large and easily accessible European computing environments.

Supporting EO use cases and providing cloud computing resources.
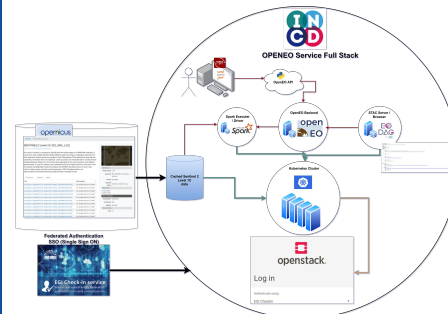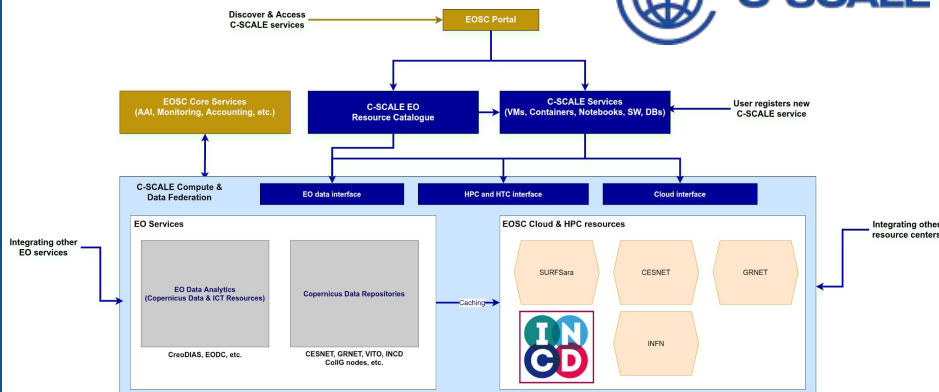
Aqua Monitor detects how the Earth's surface water has changed during the last 30 years.

Changes are detected in real-time using satellite imagery for any place on Earth.

Porting the application from the Google Earth Engine platform to the open C-SCALE infrastructure, providing an interactive (zoomable) map that displays land use changes (wet vs dry).

Relies on the top-of-atmosphere reflectance images from Landsat 4,5, 7, and 8 and will be extended to use Sentinel-2 MSI Level-1C data.

Kubernetes on INCD cloud.
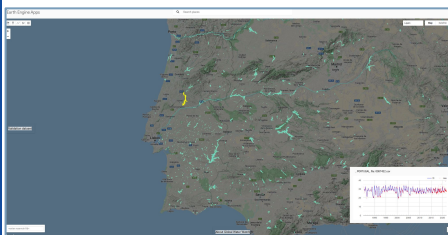
EOSC–Synergy
Quality for research sw

European Open Science Cloud (EOSC) project.

EOSC-Synergy (2019-2022) expanding national e-infrastructures and services in EOSC. Strong focus on quality, infrastructure, and thematic services. Participation on:

- **Coordination of Quality activities for software, services and data.**

- **Quality baselines for software & services.**

- **Development of Jenkins Pipeline Library.**

- Support to thematic services integration.

- Data repositories and computing services integration in EOSC.
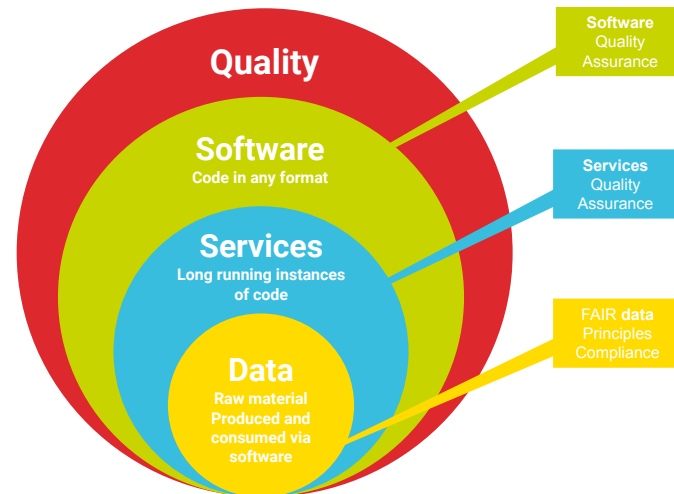
EOSC–Synergy
Quality baselines

Good practices aimed at improving *research* software and services. Set of quality conventions oriented to DevOps.

- Software baseline
  - Criteria meant to enhance the visibility, accessibility and distribution of source code.
  - Encouraging good coding practices to improve quality, reliability and security.
- Services baseline
  - Minimum set of principles for reliable and fit-for-purpose services such as web services, web applications, platforms etc.
  - Provides common coherent quality attributes aimed to ensure functional suitability and strengthening of the services reliability and stability.

The baselines are openly developed on github.

10.20350/digitalCSIC/12543

A set of Common Software Quality Assurance Baseline Criteria for Research Projects

A DOI-citable version of this manuscript is available at http://hdl.handle.net/ 10261/160086.

This manuscript (permalink) was automatically generated from indigo-dc/sqa-baseline@a9c34fa on April 29, 2020.

Institute of Physics of Cantabria (IFCA)

Institute of Physics of Cantabria (IFCA)

10.20350/digitalCSIC/12533

A set of Common Service Quality Assurance Baseline Criteria for Research Projects

A DOI-citable version of this manuscript is available at http://hdl.handle.net/.

This manuscript was automatically generated on 29-04-2020.

Authors

- **Pablo Orviz**
  0000-0002-2473-6405 · orviz
  Spanish National Research Council (CSIC); Institute of Physics of Cantabria (IFCA)
- **Mario David**
  0000-0003-1802-5356 · mariojmdavid
- **Jorge Gomes**
  0000-0002-9142-2596 · jorge-lip
- **Joao Pina**
  0000-0001-8959-5044 · jopina
- **Samuel Bernardo**
  0000-0002-6175-4012 · samuelbernardolip
- **Isabel Campos**
  0000-0002-9350-0383 · isabel-campos-plasencia
  Spanish National Research Council (CSIC); Institute of Physics of Cantabria (IFCA)

EOSC-Synergy
SQAaaS

Quality Assurance as-a-Service platform (SQAaaS)

- Enables the on-demand creation of CI/CD pipelines making quality verification and validation easily accessible to developers.
  - The **Pipeline as a Service** building block allows you to compose and test customized CI/CD pipelines in accordance with reference criteria.
  - The **Quality Assessment & Awarding** building block analyses, the level of compliance to the quality baselines.
- Integrates a wide range of quality verification tools that are made easily available through a friendly web interface.
- **LIP developed the core component that powers the SQAaaS** platform and enables creation and execution of the CI/CD pipelines.

EOSC-Synergy
SQA as a Service

The SQAaaS will be provided as a cloud service.
Making adoption and usage easier.

- No need to deploy and setup the components, Jenkins, API, web, containers.
- No need to create the yaml configurations.
- No need to provide IT resources.
- No need to manage the platform.

Basis for EOSC quality assessment
- Can issue digital badges to reward and highlight the quality achievements.
- Based on OpenBadges specification.
- Produce detailed quality reports.

https://sqaaas.eosc-synergy.eu/

https://www.eosc-synergy.eu/technical-areas/quality/
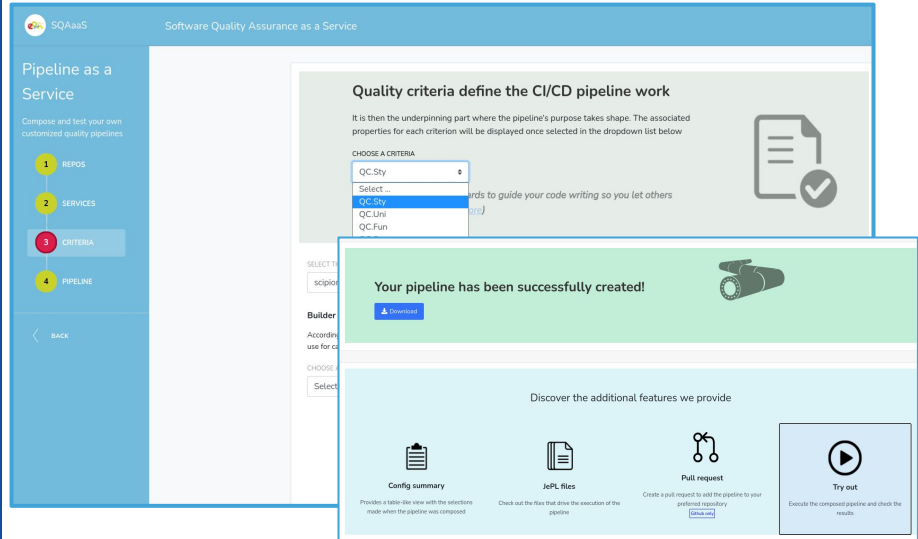
IBERGRID 2023

Benasque, Spain 24th to 30th September 2023

- Design and implementation of **Digital Twins**
- Development, implementation and operation of Data Spaces
- Cooperation between Iberian research communities
- Research applications in advanced Digital Infrastructures
- Development of innovative software services
- R&D for computing services, networking, and data-driven science
- Quality of software, services and data
- Enabling and fostering Open Science adoption in EOSC
- Datacenter and digital infrastructures development and status reporting

Submission of abstracts is open !
https://www.ibergrid.eu/2023-ibergrid-benasque