



interTwin: an interdisciplinary Digital Twin Engine for Science



Andrea Manzi (EGI Foundation)

BOLZANO 2-4 OCTOBER 2023

eurac
research



Institute of
Atmospheric Pollution
Research
National Research Council of Italy



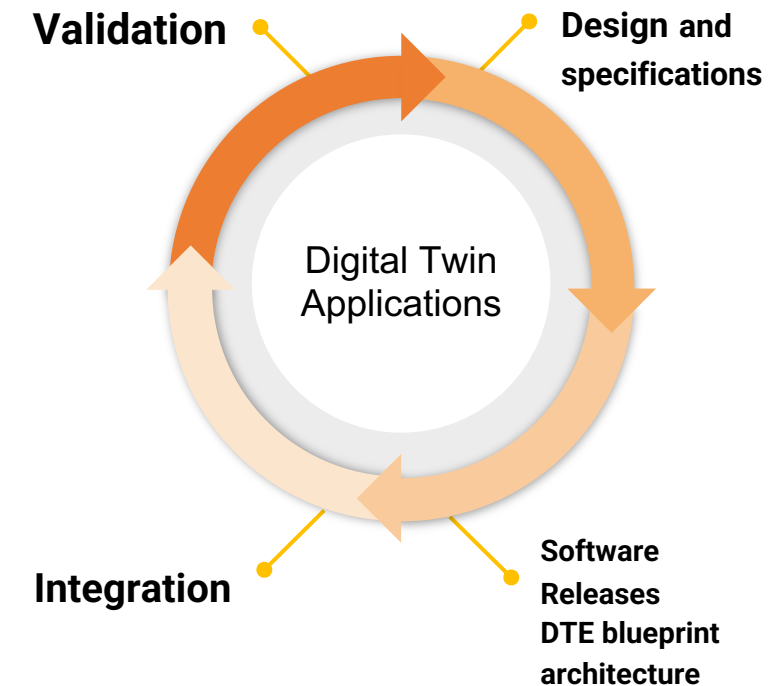
European
Commission



High level overview

Creation of a **prototype digital twin engine**:

- Based on a **DTE Blueprint Architecture**
 - Interdisciplinary
 - Co-Designed (Providers and Communities)
- Resulting **Platform** must be
 - Open Source with
 - TRL 6 (prototype model) to 7 (prototype pilot)
- And it must be based on
 - **Open Standards**
 - with the capability to integrate with **application specific DT's**





Consortium



EGI Foundation as coordinator

29

Participants, including 1 affiliated entity and 2 associated partners

Consortium at a glance

10
Providers

cloud, HTC , HPC
resources and
access to
Quantum systems

11
Technology
providers

delivering the
DTE
infrastructure and
horizontal
capabilities

14
Community
representants

from 5 scientific
areas;
requirements and
developing DT
applications and
thematic modules

1.09.22 - 31.08.25

Budget 11,7 M euro

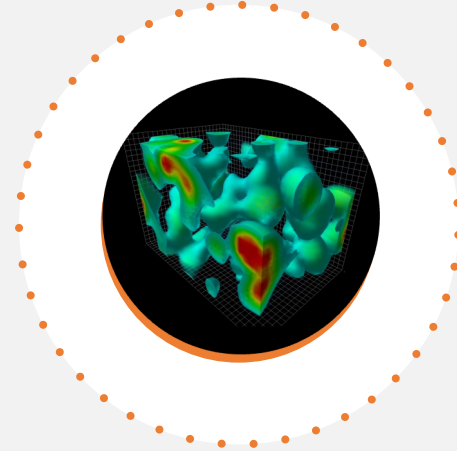
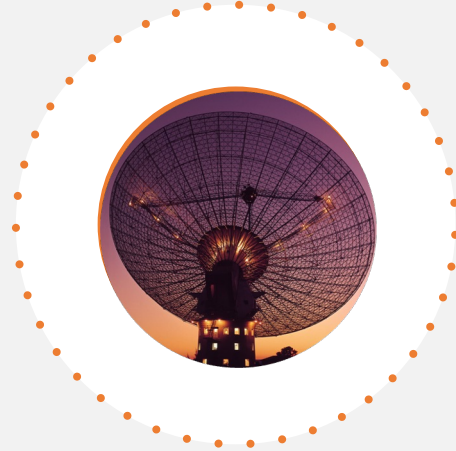
BOLZANO 2-4 OCTOBER 2023



Physics domain DTs



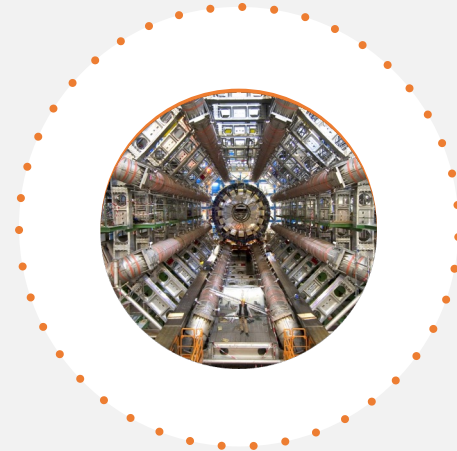
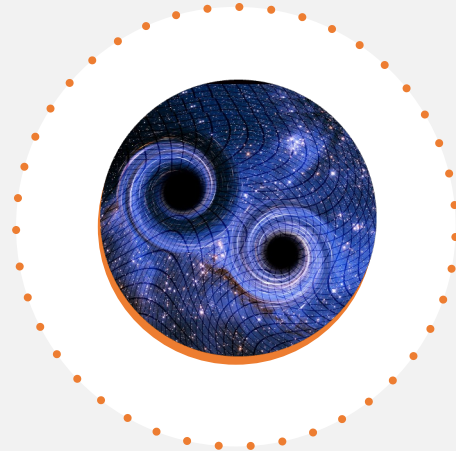
Radio
Astronomy



Quantum Field Theory



Gravitational Wave
Astronomy

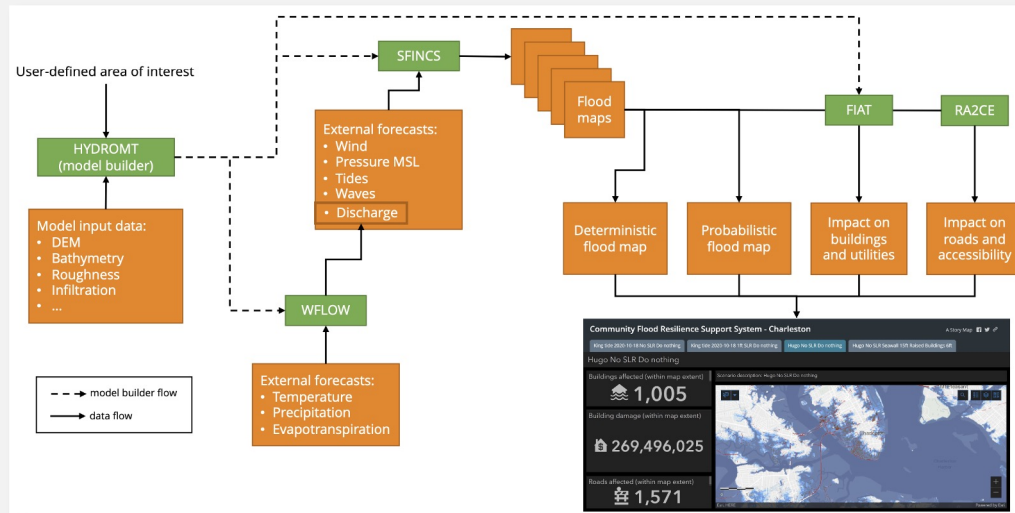


High Energy Physics

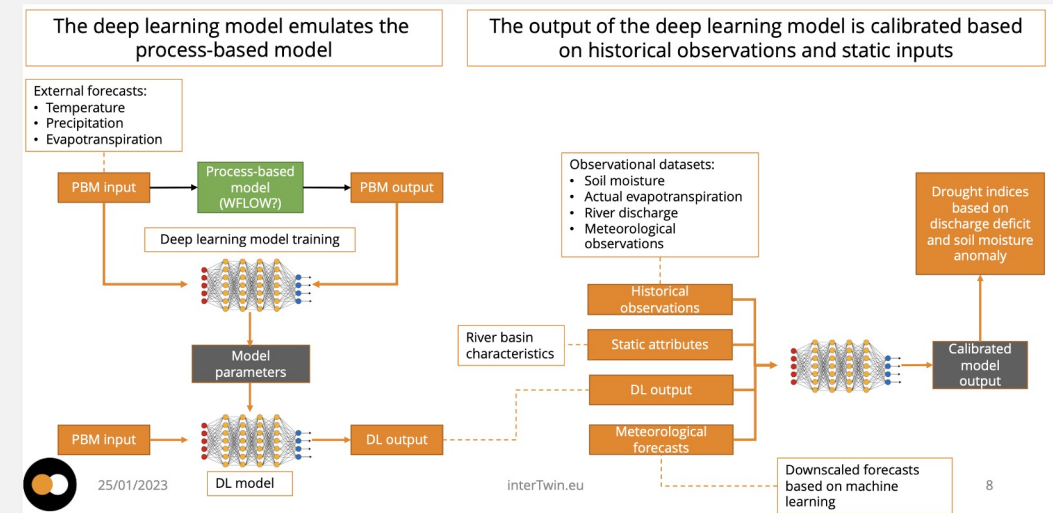


6 DTs for Extreme Events on the Earth

- Climate Change Future Projections of Extreme Events as **storms and fire**;
- Early Warning for Extreme Events of **floods and droughts**;
- Climate Change Impacts of Extreme Events of **storms, fire, floods and droughts**.



Digital Twin for Flood Early Warning in coastal and inland regions



Digital Twin for Drought Early Warning in alpine regions

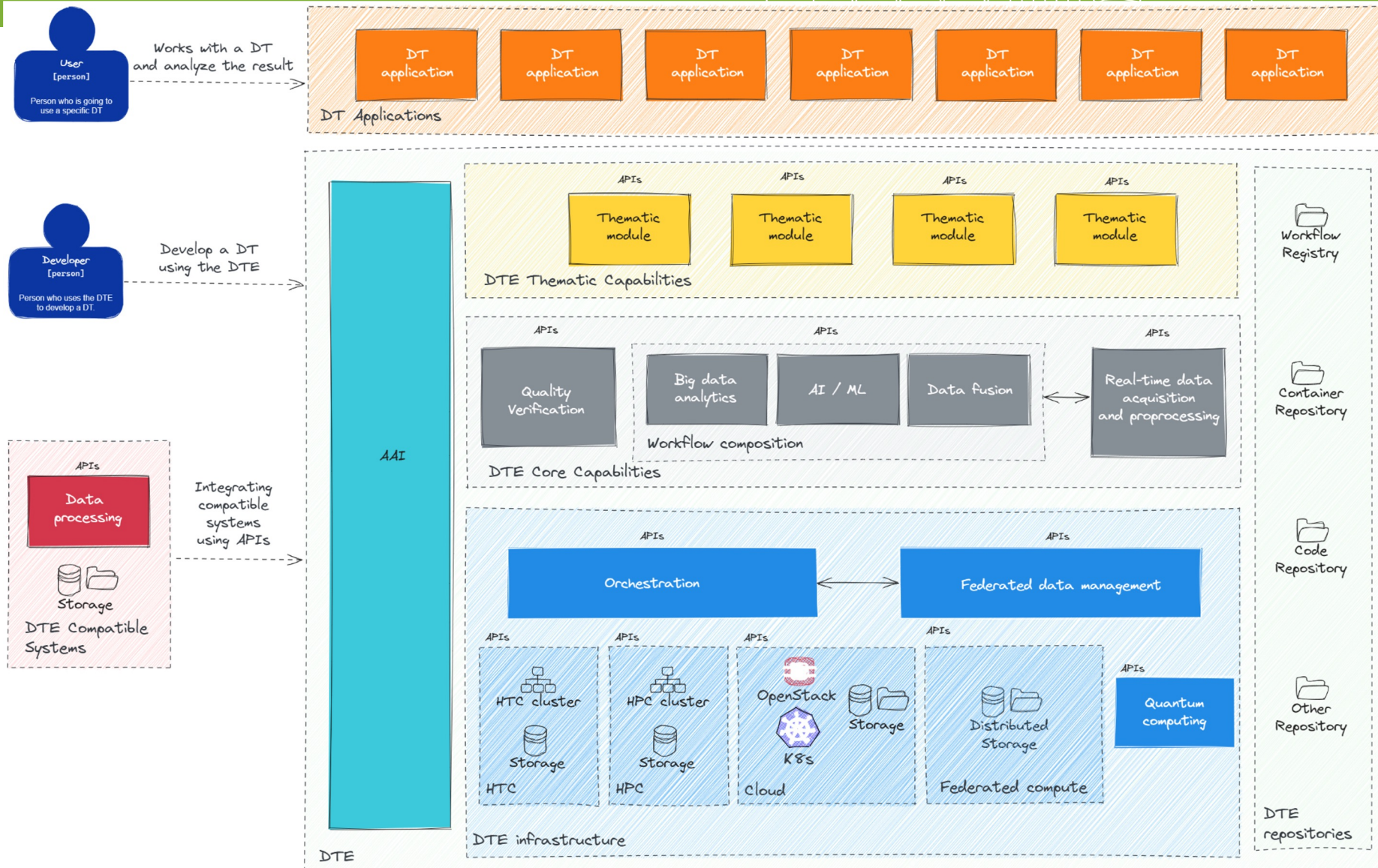


Blueprint and co-design

- First version of the Blueprint architecture and design specifications are available in Github and Zenodo. Second and Third versions are planned in 2024
 - <https://github.com/interTwin-eu/architecture-diagrams/tree/main/Blueprint%20architecture>
 - <https://zenodo.org/communities/intertwin/>
- It also considers other relevant initiatives and projects (*DestinE, EOSC, ESCAPE, C-Scale, Digital Twin Consortium, Gaia-X, and other EU Data Spaces*) to identify potential architectural components that can be incorporated within the interTwin context, and to identify where interoperability is desirable.



interTwin components

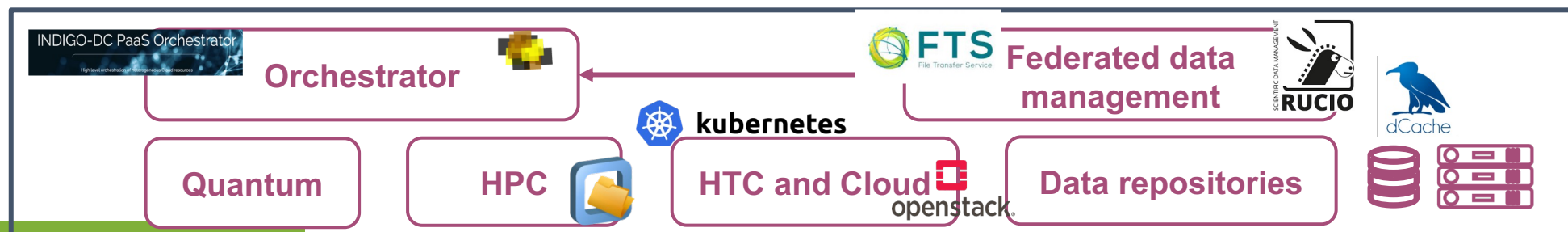




DTE components - Infra

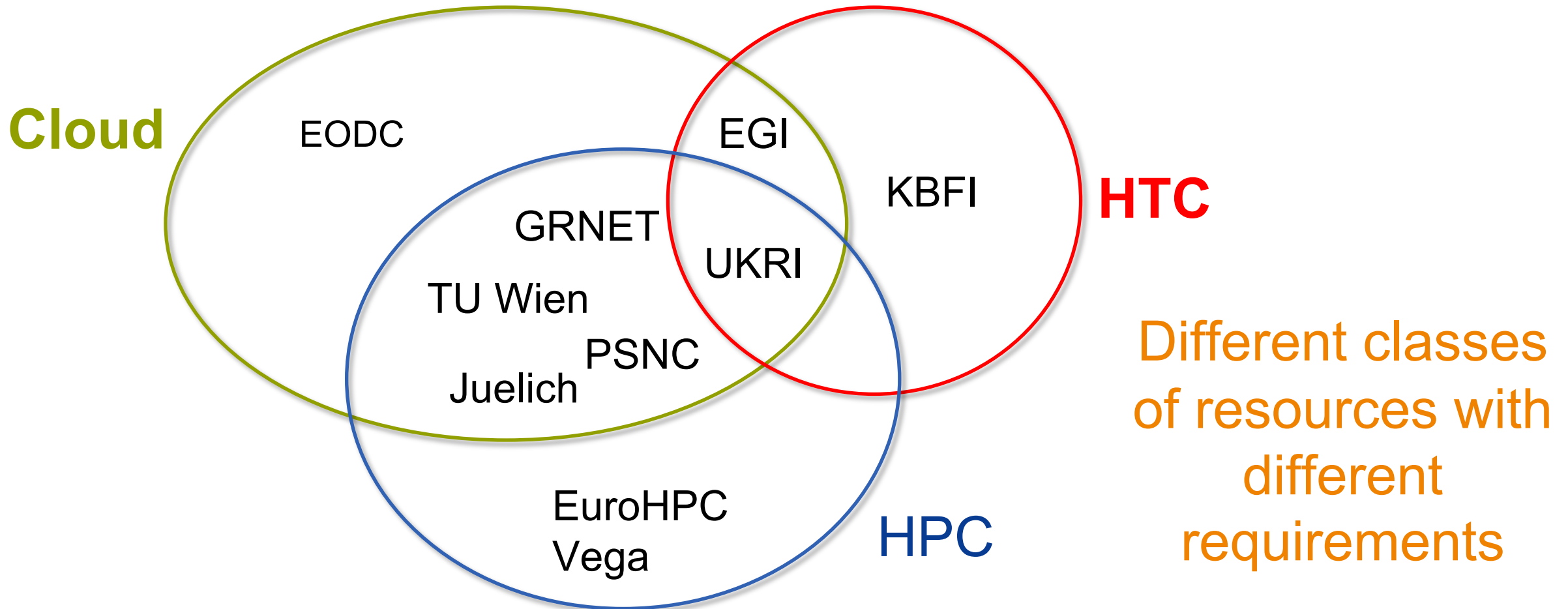
<p>Orchestrator</p>	<p>PaaS Orchestrator + Infrastructure Manager elaborating deployment requested expressed in TOSCA with planned extensions for AI based orchestration</p>
<p>Federated data management</p>	<p>Based on ESCAPE Data Lake architecture and services, Rucio, FTS and HTTP accessed caches/storages. Data lake concept extended to HPC facilities which also integrates STAC Catalogues.</p>
<p>Federated Compute</p>	<p>Use of single-sign-on in complex simulation and modelling tasks to access data and different compute facilities, including offloading to HPC. Automated modelling and simulation fused with data repositories and computation with containers on HTC, Cloud and HPC</p>

DTE
Infrastructure





Pilot Compute Resources for HPC, HTC and Cloud

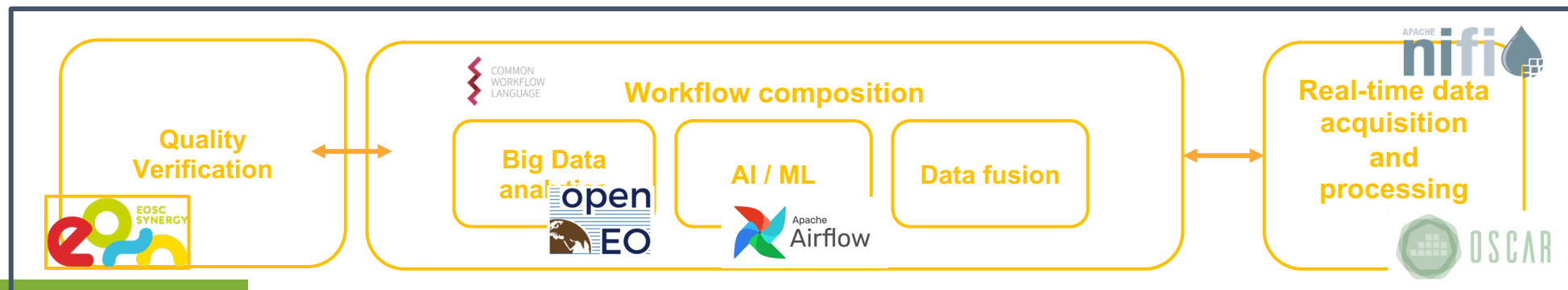




DTE components - Core

<p>Workflow composition</p>	<p>Workflow definitions based on CWL, run on a workflow execution system (e.g. StreamFlow, AirFlow) and able to combine self-contained execution steps from other workflow engines (e.g . ecFlow, Ophidia), different back-ends, distributed big data analysis tools (e.g. openEO, Dask, Spark) and ML/DL training platforms (e.g., Horovod, HeAT, PyTorch DDP). Data Fusion as one of the workflow steps to merge observational and modeled data and different data sources.</p>
<p>Real-time data acquisition and processing</p>	<p>Generic framework for real time acquisition and processing that builds on event-triggered execution of workflow engines and exploit serverless computing. Based on Apache NIFI and OSCAR Framework</p>
<p>Quality Verification</p>	<p>Specific module for quality assurance (QA) that aims at tackling the early validation of the DTs, before being deployed as a “living DT”. Based on the SQaaS developed in EOSC Synergy project</p>

**DTE
Core Modules**





DTE components - Thematic

The **interTwin DTE Thematic modules** are addons providing capabilities tailored to the needs of specific groups of applications (i.e. of general applicability to multiple 'adjacent' communities) developed with the aim to be "**promoted**" as Core modules following the successful adoption by multiple resource communities from different domains:

- **Lattice QCD simulations**
- **Noise simulation for radio astronomy**
- **GAN-based modules to manage noise simulation, low-latency de-noising and veto generation**
- **Climate analytics and data processing**
- **Earth Observation Modelling and Processing**
- **Hydrological model data processing**
- **Fast simulation with GAN**



Project Year 1

Project Year 2

Project Year 3

06.2023

DTE Design and specifications

Deliverables: Report on requirements for all use cases

DTE blueprint architecture

Deliverable: DTE blueprint architecture, functional specifications and requirements analysis v1

11-.2023-
02-20024

Software Releases

Deliverables: Software releases for all use cases and modules

01.2024

DTE blueprint architecture II

Deliverable: DTE blueprint architecture, functional specifications and requirements analysis v2

04.2024

Validation

Deliverables: DT application development and integration report

09.2024

Design and specifications II

Deliverables: Updated report on requirements for all use cases

10.2024

DTE blueprint architecture III

Deliverable: DTE blueprint architecture, functional specifications and requirements analysis v3

01-04.2025

Software Releases II

Deliverables: Final Software releases for all use cases and modules

07-08.2025

Validation II

Deliverables:
DT application development and integration report
Report on software architecture concepts based on DestinE and InterTwin
Final Architecture design of the DTs capabilities

2023

2024

2025

Timeline

New or extended components in our Community in Github

<https://github.com/interTwin-eu/>



Interoperability and technology exchange

- interTwin is planning pilot activities with **DestinE** to **design a compatible architecture** that addresses the requirements of the largest set of user communities, thanks to **ECMWF** partner of the project
- DG-Connect has also started an activity of technology exchange, **harmonization of architectures** and common DT glossary. It includes DestinE, interTwin and other HE projects dealing with Earth DT (BioDT, DT-GEO and EDITO-Infra)





Future plans and long term support



- The Project is building services to be operated by the EG-Federation and linked to EOSC
 - Aim at extending the technical capabilities of the European Open Science Cloud with modelling & simulation tools integrated with its compute platform.
- In parallel it aims at integrating technologies in the DestinE Ecosystem. Early discussions and interest on:
 - Computing offloading component ([interLink](#))
 - Intelligent Cloud orchestration subsystem
- Finally the project is facilitating its Environmental DTs prototypes to be hosted in the context of DestinE if the 2 DTEs are interoperable.
 - **Early warning** DT from Deltares as a common DT.

Thank you!



www.intertwin.eu



info@intertwin.eu



[intertwin_eu](https://twitter.com/intertwin_eu)



[intertwin](https://www.linkedin.com/company/intertwin)



Backup slides



Mission of the EGI Federation

Deliver open solutions
for advanced computing and data
analytics in research and innovation

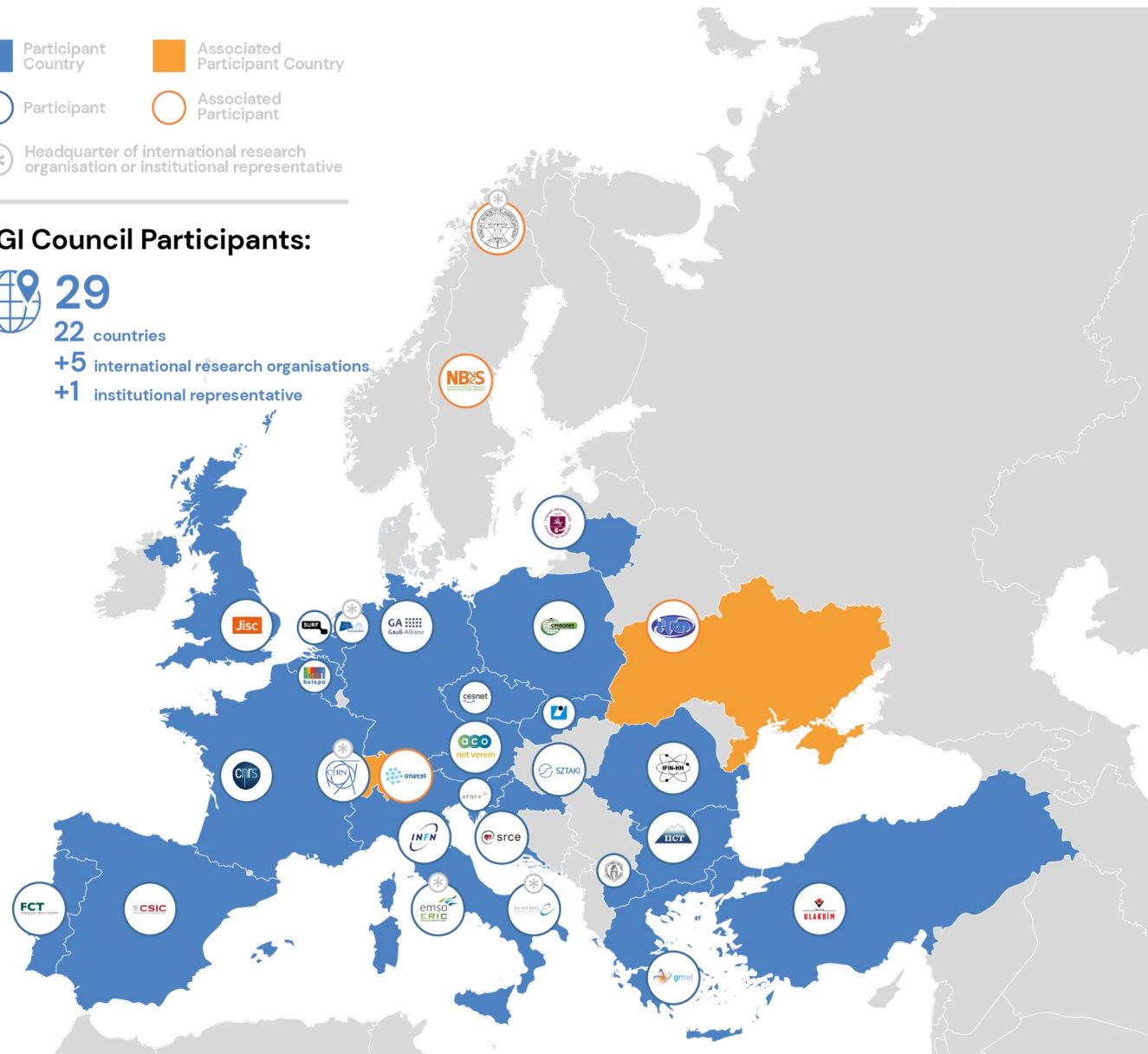
Mission of the EGI Foundation

Enable the EGI Federation
to serve international research and
innovation together, acting as it's
coordination body

- Participant Country
- Associated Participant Country
- Participant
- Associated Participant
- Headquarter of international research organisation or institutional representative

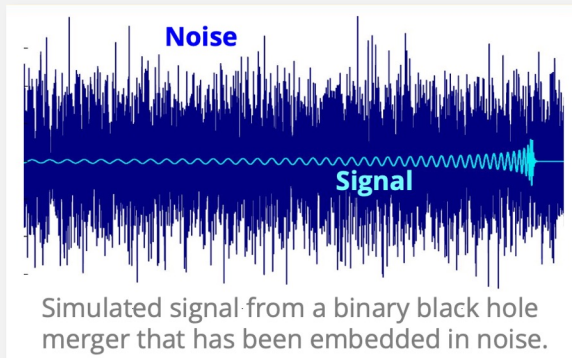
EGI Council Participants:

 **29**
22 countries
+5 international research organisations
+1 institutional representative



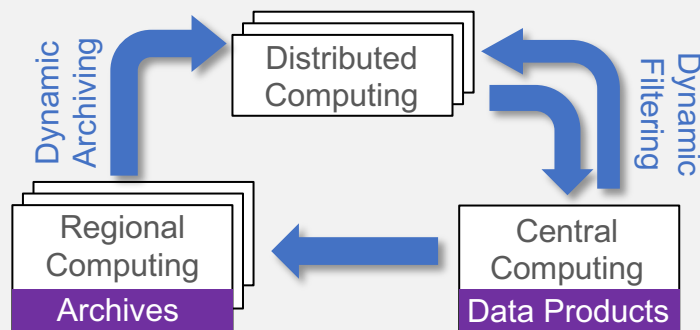


DT of the VIRGO Interferometer



It is meant to **realistically simulate** the noise in the detector, in order to study how it reacts to external disturbances and, in the perspective of the **Einstein Telescope**, to be able to detect noise “glitches” in **quasi-real time**, which is currently not possible. This will allow sending out **more reliable triggers** to observatories for multi-messenger astronomy.

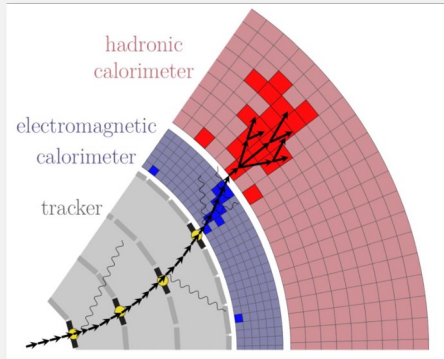
DT for noise simulation of next-generation radio telescopes



Meant to provide DTs to simulate the noise background of radio telescopes (**MeerKat**) will support the identification of rare astrophysical signals in (near-)real time. The result will contribute to a realisation of "**dynamic filtering**" (i.e. steering the control system of telescopes in real-time).

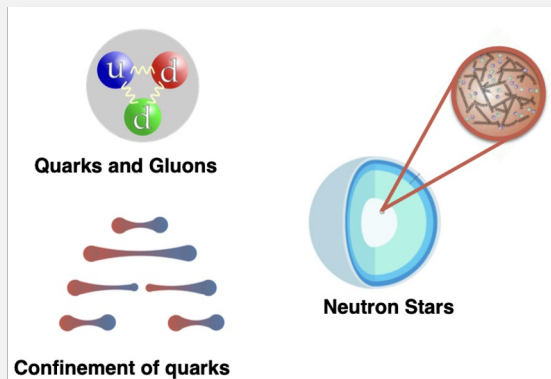


DT of Large Hadron Collider (LHC) detector components



Seeking for strategies to face the increase in the need for simulated data expected during the future High Luminosity LHC runs. The primary goal is to provide a fast simulation solution to complement the Monte Carlo approach. ***Faster and deeper cycles of optimisation of the experiment parameters*** in turn will enable breakthroughs in experimental design.

DT of the Standard Model in particle physics



Competitive results in Lattice QCD require the ***efficient handling of Petabytes of data***, therefore the implementation of advanced data management tools is mandatory. On the side of algorithmic advancement, ML algorithms have recently started to be applied in Lattice QCD. The goal is to ***systematize the inclusion of ML for large scale parallel simulations***.

EUROGEO WORKSHOP 2023



DT	Geographical region of Interest
Tropical Storms change in response to climate change	Indian and Pacific Ocean
Wildfires risk assessment in response to climate change	Europe
Flood Early Warning in coastal and inland regions	Selected European regions, Philippines
Alpine droughts early warning	European Alps
Extreme Rainfall events change in response to climate Change	Europe
Flood Climate impact in coastal and inland regions	Selected European regions, Mozambique





example: CMCC DTs Data needs (WildFires)

Repositories	ERA5	CMIP6	IBTrACS	Fire Danger Indices Historical Data
Data Types	Reanalysis - Gridded	Projections - Gridded	Observative records	Reanalysis - Gridded
Data Formats	NetCDF	NetCDF	CSV	NetCDF
Spatial Resolution/ Coverage	~0.25°x0.25°/global	~0.25°x0.25°/global	geographical coordinates/global	~0.25°x0.25°/global
Temporal Resolution/ Extent	Hourly up to daily/ 1979-2020 (past data)	6-hourly up to daily /2015-2100 (projections)	3-hourly/1979–2020 (past data)	Daily at 12:00 / 1979-2020 (past data)
Update Frequency	Daily with 5-day latency	Very rarely	Twice a week (at least a year to gather significant data)	Monthly
Usage	Model training	Model inference	Model training	Model training
APIs/Tools	Copernicus CDS	Synda	wget/curl	Copernicus CDS