Destination Earth

Our planet's digital twin

thomas.geenen@ecmwf.int

ECMWF

ECMWF's Forecasting Systems

Established in 1975, Intergovernmental Organisation
- 23 Member States | 12 Cooperating States
- 350+ staff

24/7 operational service
- Operational NWP – 4x HRES+ENS forecasts / day
- Supporting NWS (coupled models) and businesses

Research institution
- Experiments to continuously improve our models
- Reforecasts and Climate Reanalysis

Operate 2 EU Copernicus Services
- Climate Change Service (C3S)
- Atmosphere Monitoring Service (CAMS)
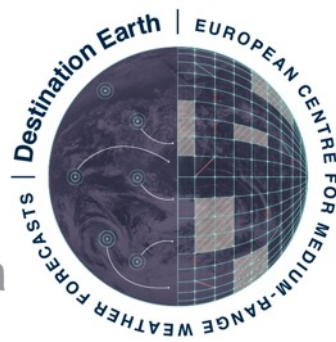- Support Copernicus Emergency Management Service CEMS

Destination Earth
- Deliver the DestinE Digital Twin Engine (DTE)
- Deliver and operate 2 Digital Twins

Reading, GB

Bonn, DE

Bologna, IT

Funded by the European Union

Destination Earth (DestinE) - ECMWF's role
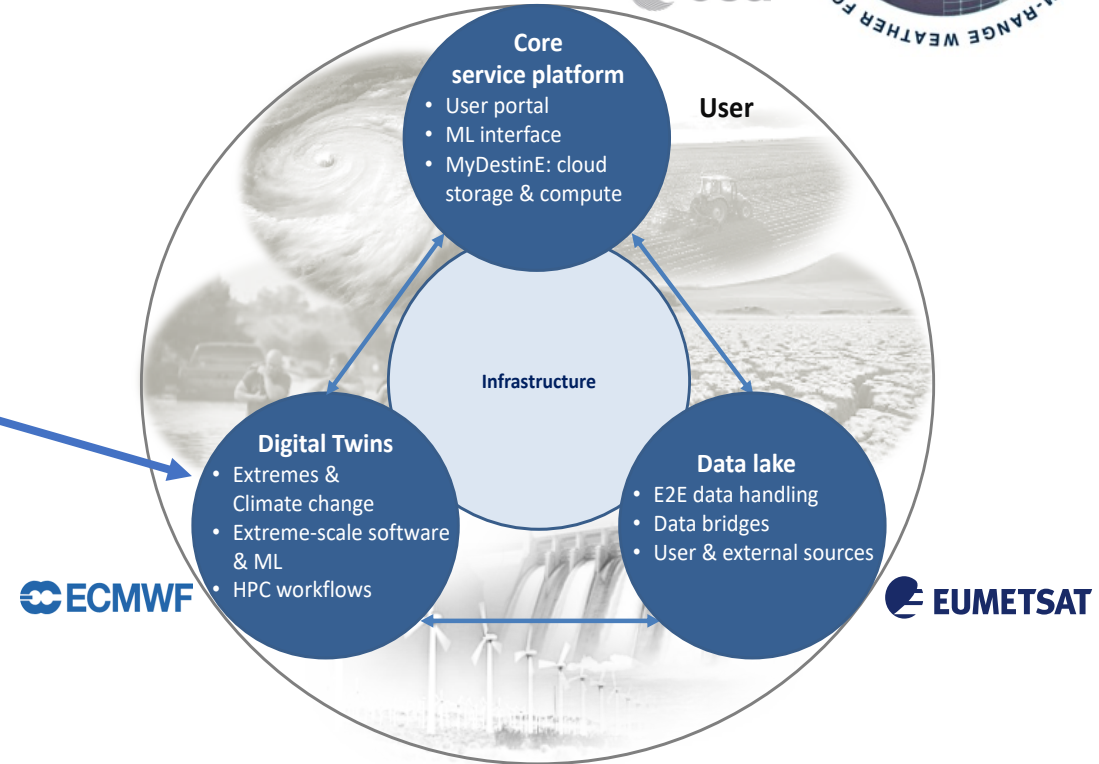
The DestinE **Digital Twin Engine** (DTE):
- common system approach to a unified orchestration of Earth-system simulations requiring large-scale HPC resources and the fusion of observations with models

Weather-induced and Geophysical **Extremes Digital Twin**:
- capabilities and services for the assessment and prediction of environmental extremes

Climate Change Digital **Twin**:
- capabilities and services support of climate change adaptation policy information scenario testing

Core service platform
- User portal
- ML interface
- MyDestinE: cloud storage & compute

User

Infrastructure

Digital Twins
- Extremes & Climate change
- Extreme-scale software & ML
- HPC workflows

Data lake
- E2E data handling
- Data bridges
- User & external sources

2021-2023
- Operational cloud-based platform
- First two digital twins

2023-2025
Platform integrates the next operational digital twins and offers services to public sector users

2025-2027+
Towards a full "digital twin of the Earth" through a convergence of multiple digital twins on the platform
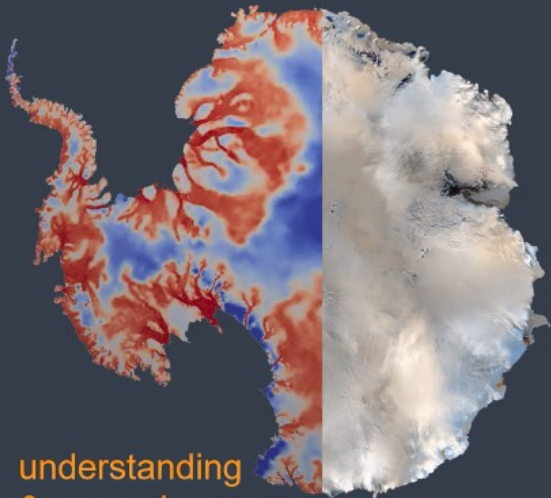
3

Funded by the European Union

What's a Digital Twin?

Digital twins have the potential to revolutionize decision-making across science, technology & society

ODEN

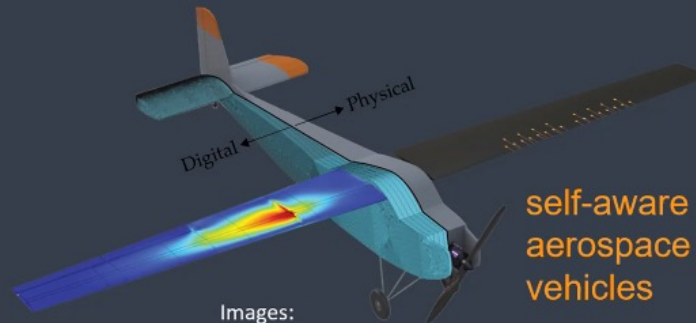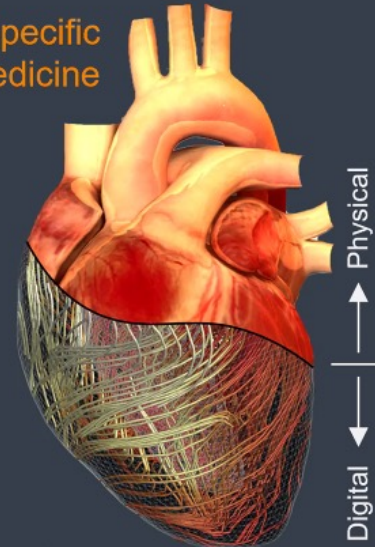Digital ← | → Physical

understanding & managing climate change

Images: O. Ghattas

Digital ← Physical

self-aware aerospace vehicles

Images: M. Kapteyn, K. Willcox

patient-specific medicine
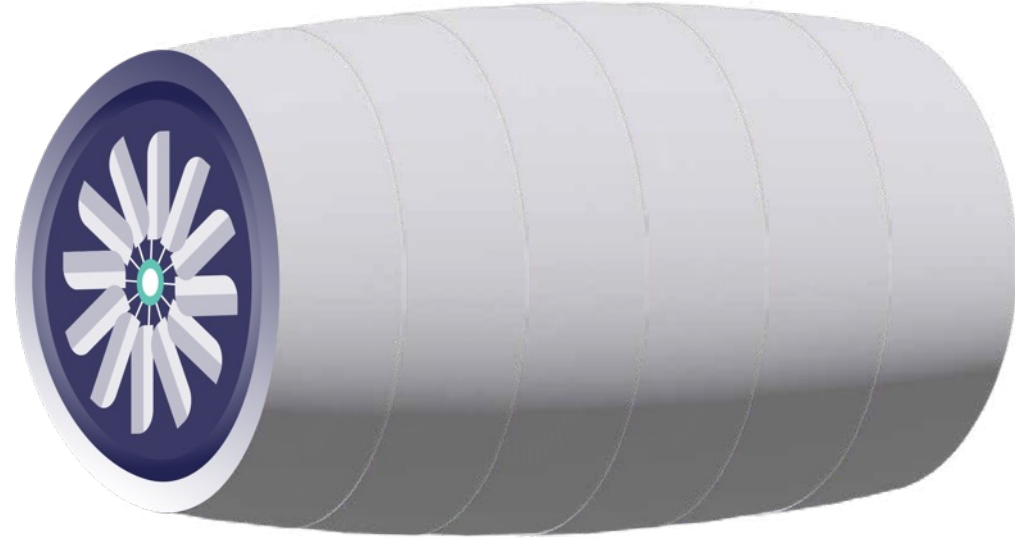
Physical / Digital

Images: G. Foss, H. Liu, M. Sacks

Some urban authorities are developing digital copies of cities, as portrayed in this artist's impression.

Make more digital twins

Virtual models boost smart manufacturing by simulating decisions and optimization, from design to operations, explain Fei Tao and Qinglin Qi.

➢ continuous **simulation & observation**
➢ **technical user interaction**
➢ scientific theory and adaptation **scenario testing**
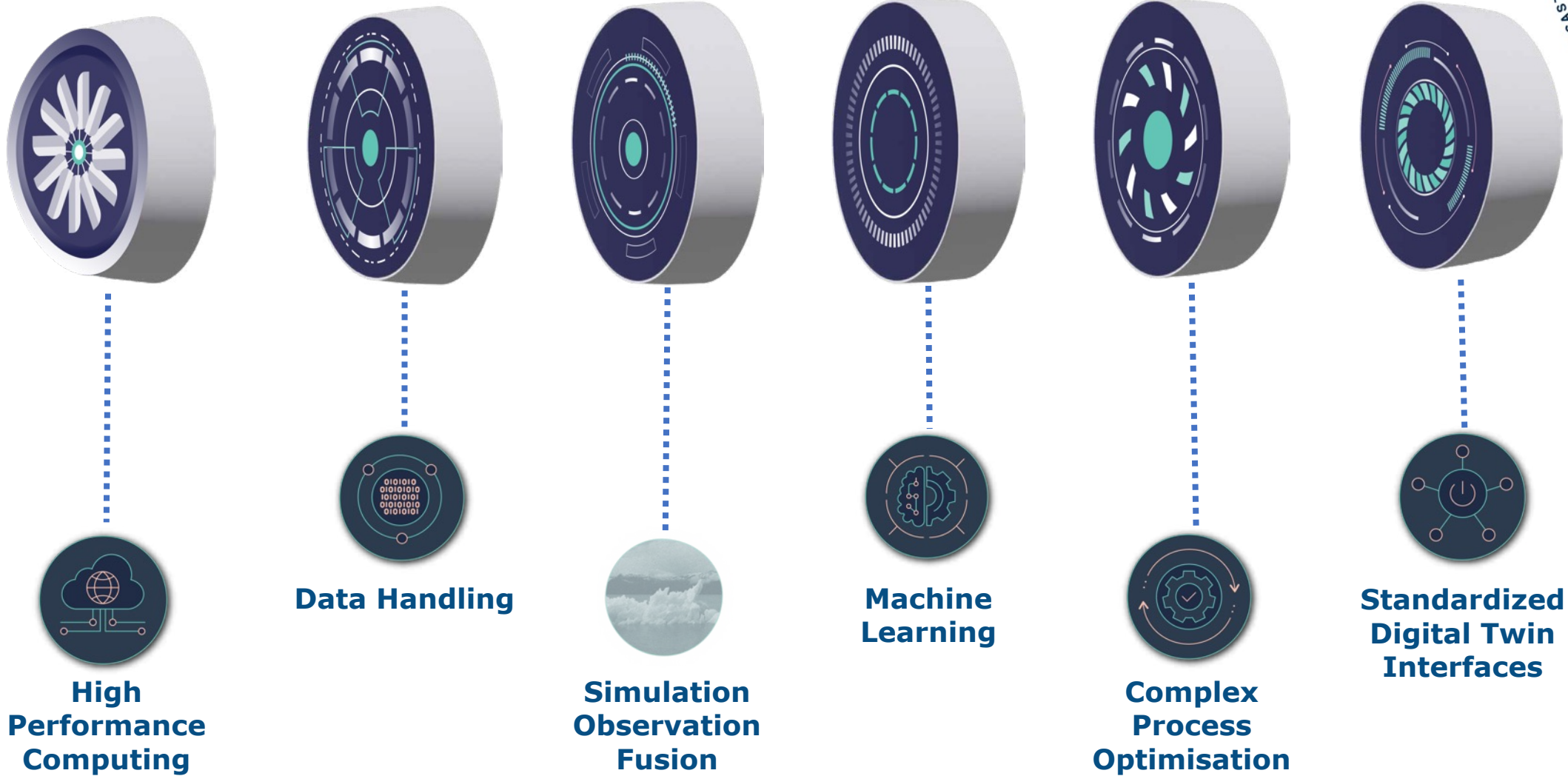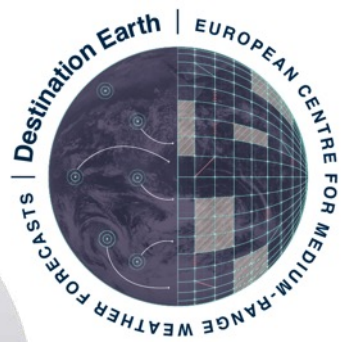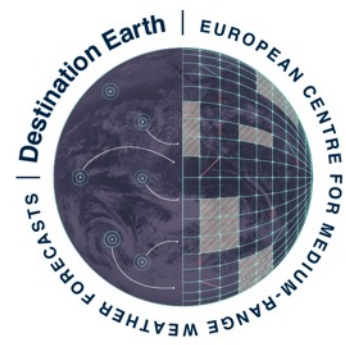
ECMWF

Funded by the European Union

Gaming Engine type framework

The Digital Twin Engine

High Performance Computing

Data Handling

Simulation Observation Fusion

Machine Learning

Complex Process Optimisation
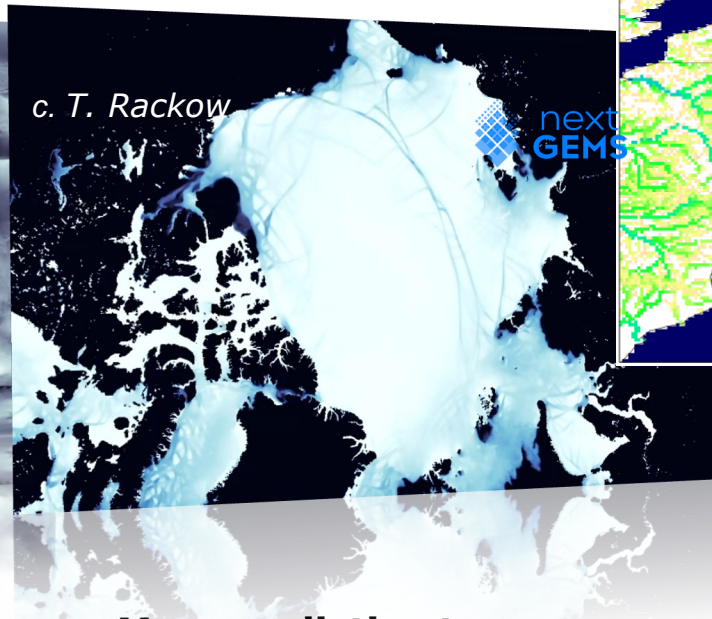
Standardized Digital Twin Interfaces

# DestinE's Digital Twins: Quality + Impacts + Interaction

1. **Better simulations** based on **more realistic models**

2. **Better ways of combining all observed and simulated information** from entire Earth system = physical + food/water/energy/health **supporting action scenarios**

3. **Interactive and configurable access to all data, models and workflows**

c. N Koldunov

c. T. Rackow
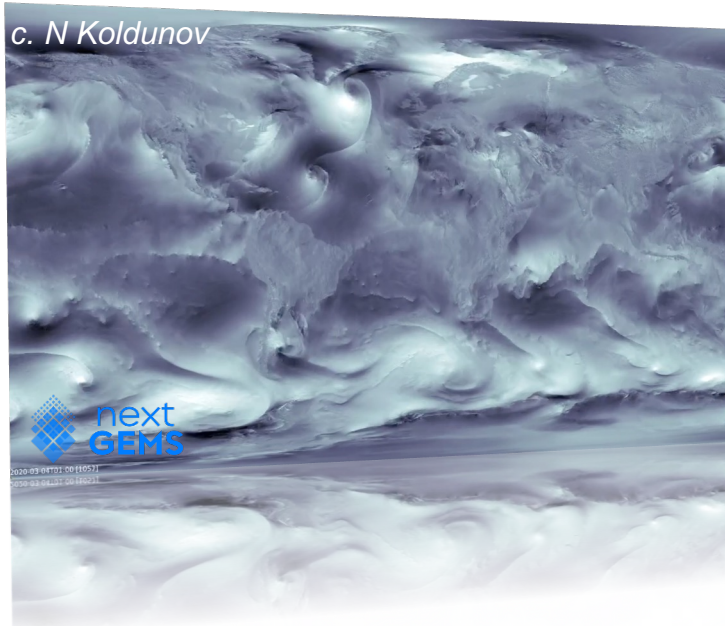
More realistic at *global* **scale**

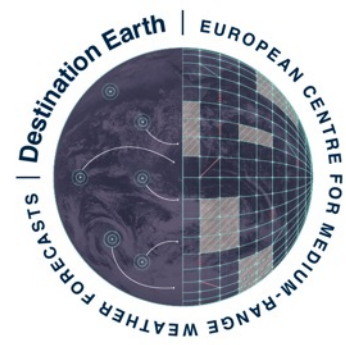**More realistic at** *local* **scale**
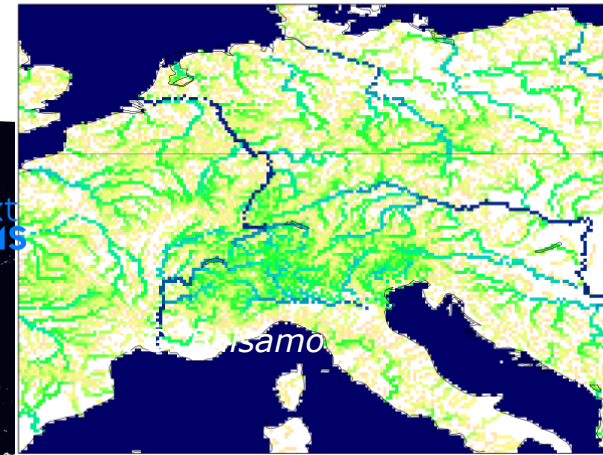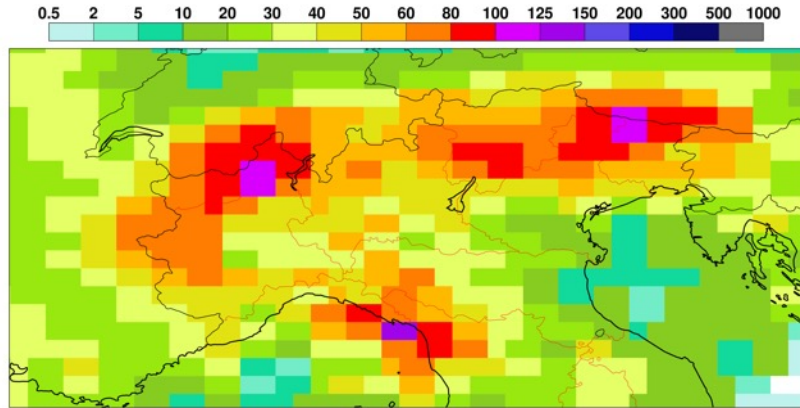
Include impacts where they matter

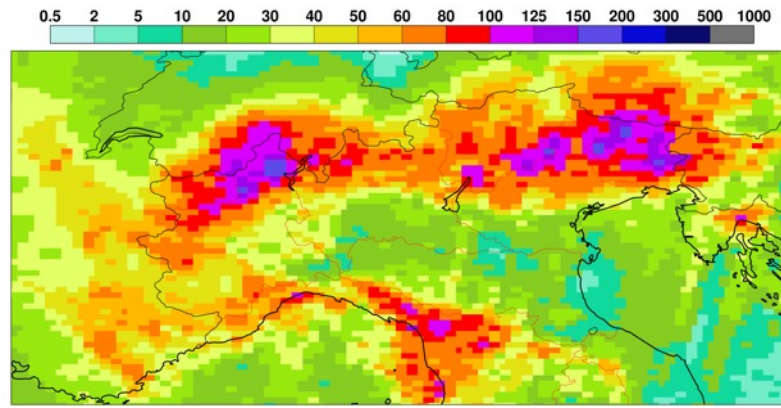**Trial different adaptation and mitigation scenarios**

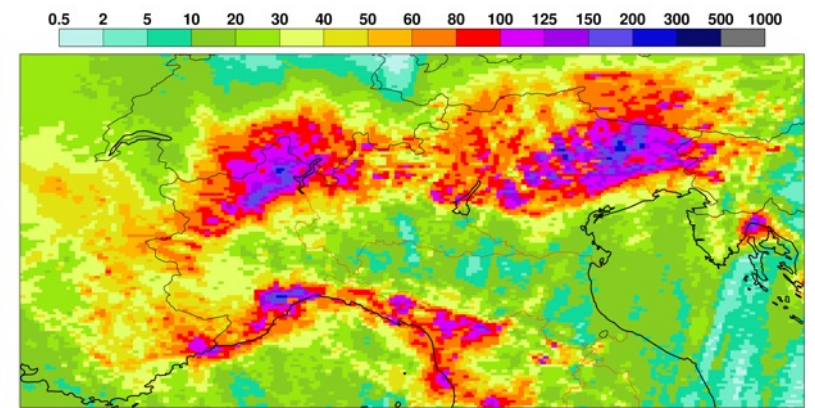# Continuous Extremes DT simulations : storm Adrian (ct 2018)
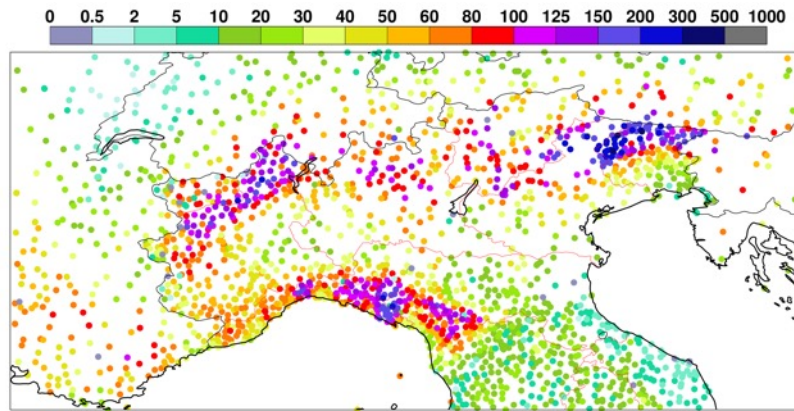
24h accumulated precipitation (T+54h - T+78h)

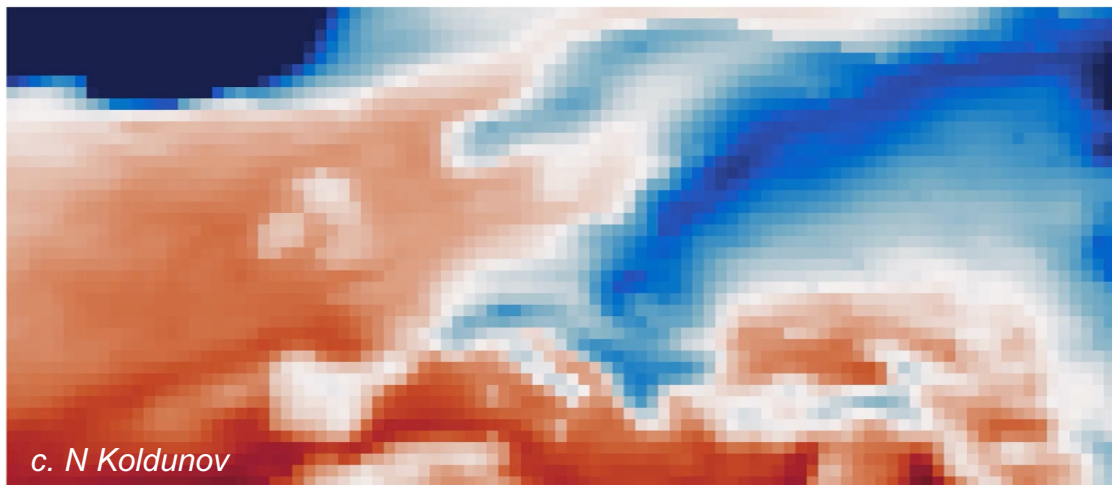IFS 48r1 29 km

IFS 48r1 9 km

IFS 48r1 4.5 km

Observations

E. Gascon

Climate DT

multi-decadal, global, storm/eddy-resolving numerical Earth-system simulation capability with the timely delivery of climate **information** for policy adaptation; observation based assessment framework; use cases for impact-sectors such as water, energy, food or health

Storm & eddy resolving simulations

c. N Koldunov

Today's global climate models

IPCC AR6 (2021)

c. N Koldunov
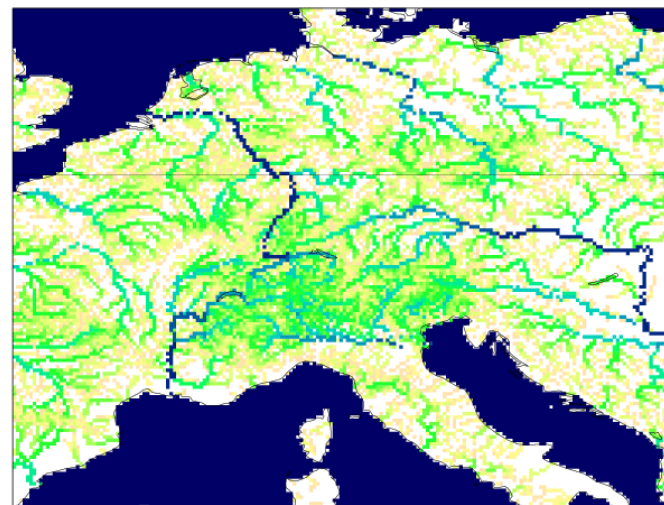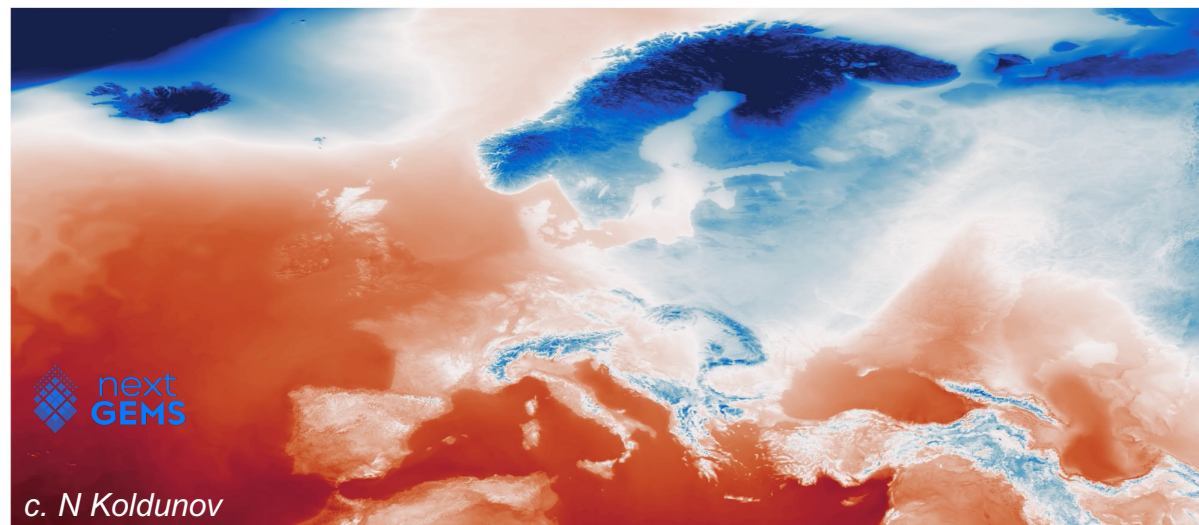
Collocated weather, climate and impact-sector information on scales where impacts of climate change and extreme events are felt

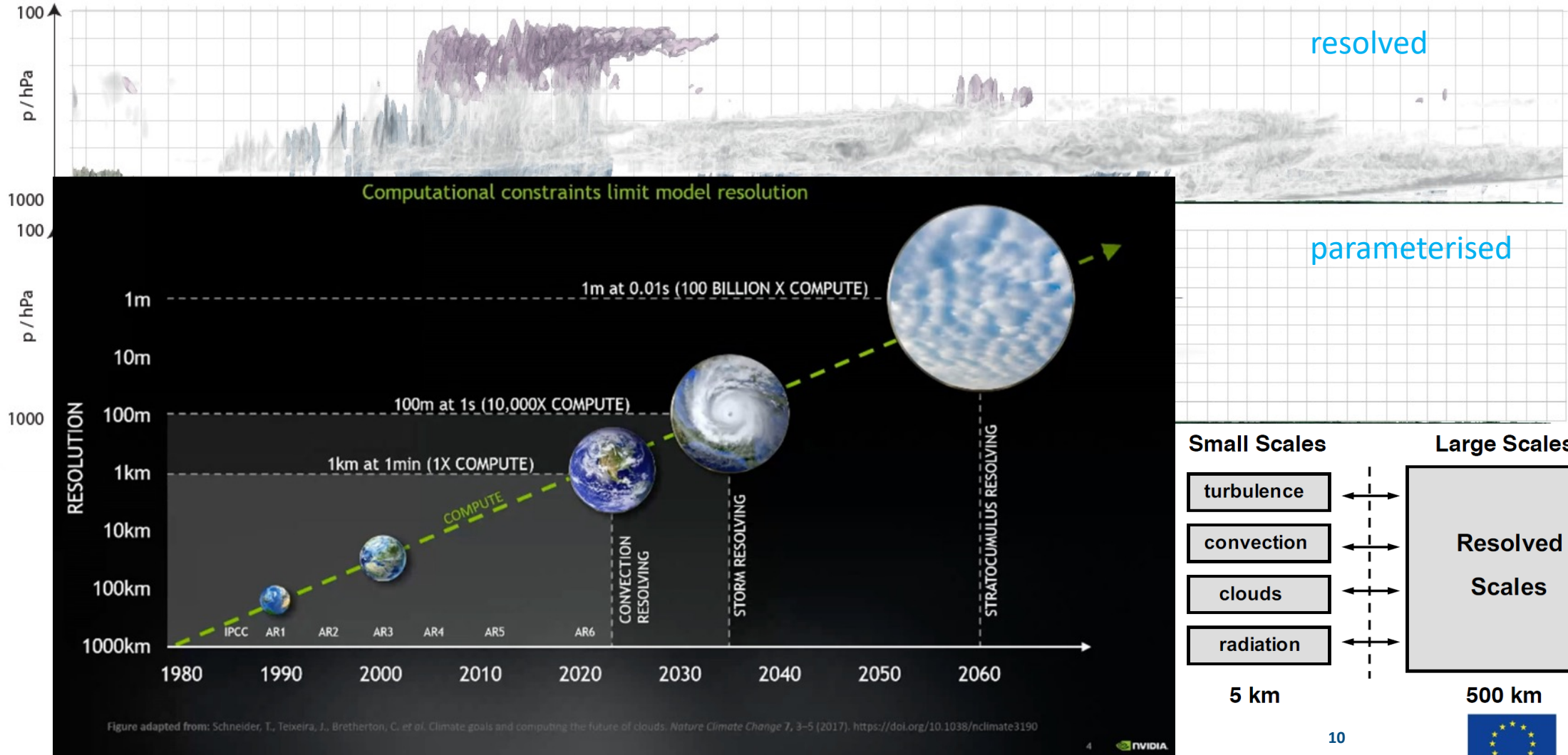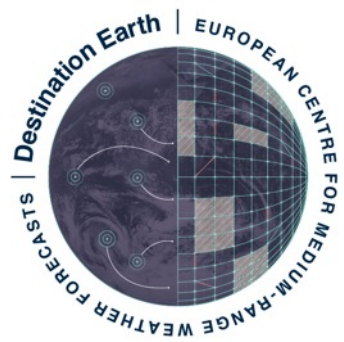ECMWF

High resolution simulation is essential but why?

Courtesy Bjorn Stevens

Systems are becoming more modular and more cloud like

Funded by the European Union
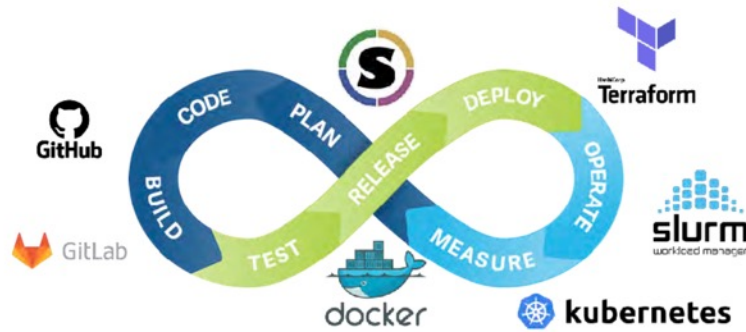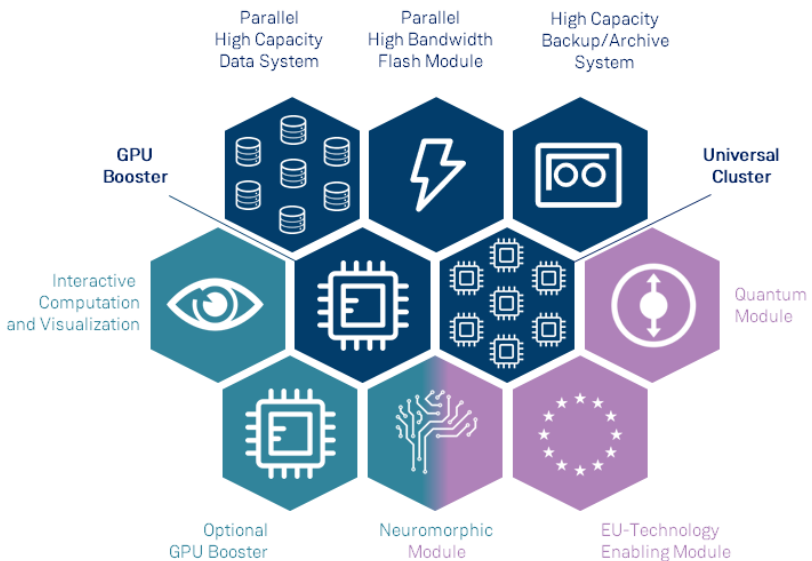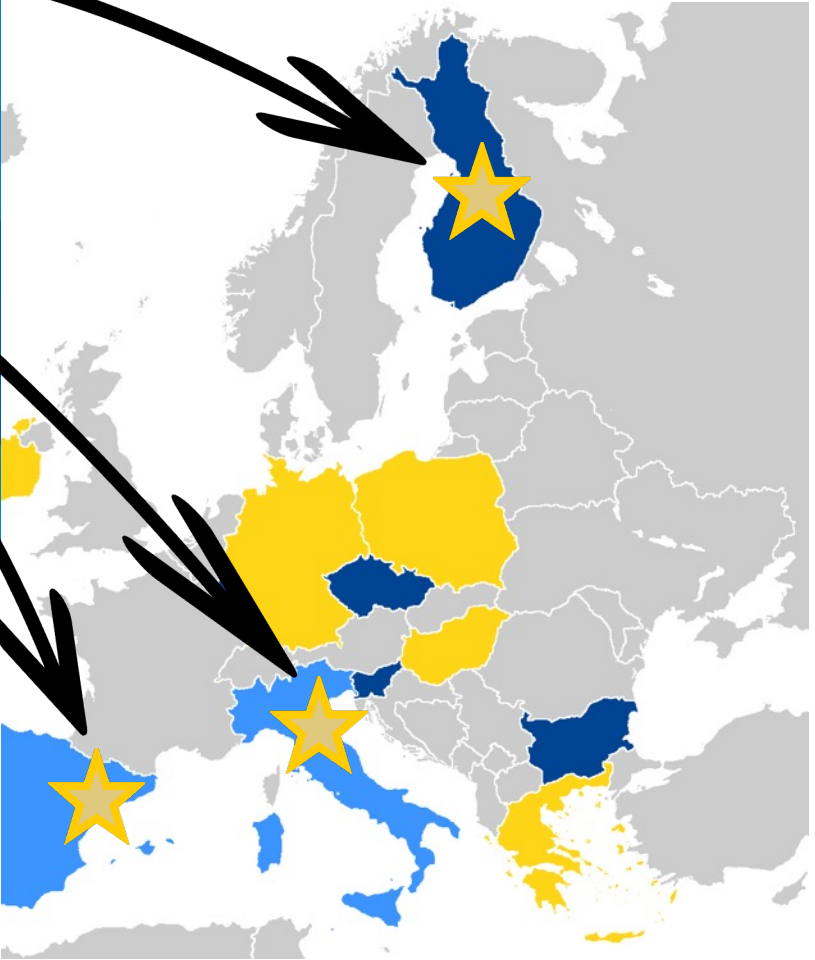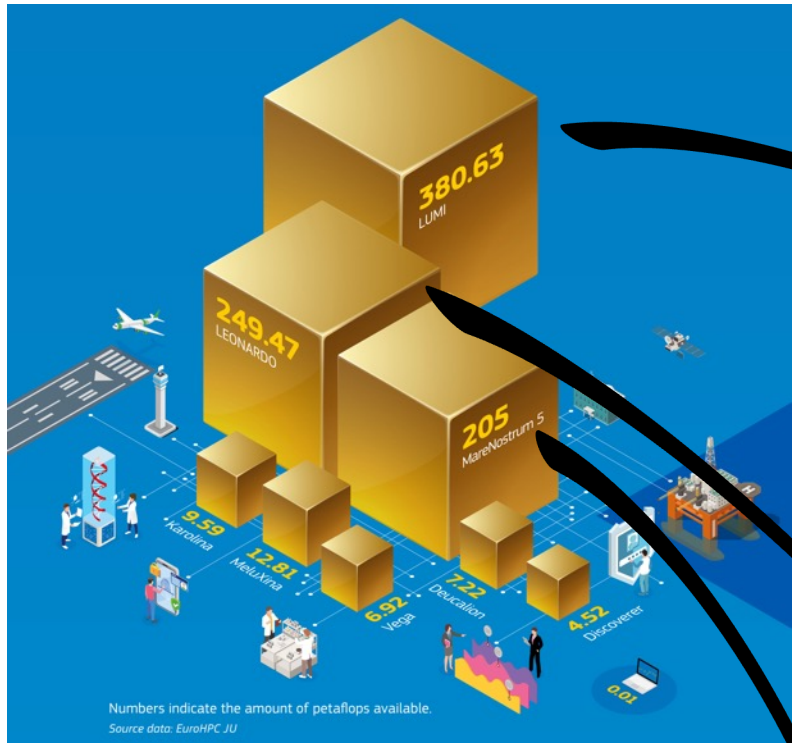
port your models: an example (The ECMWF way)

A portable approach

Atmosphere
- Spectral transform
- Grid point dynamics
- Semi-implicit solver
- Physical parameterisations

Surface (Land, Lakes) | Ocean waves | Sea ice

Ocean
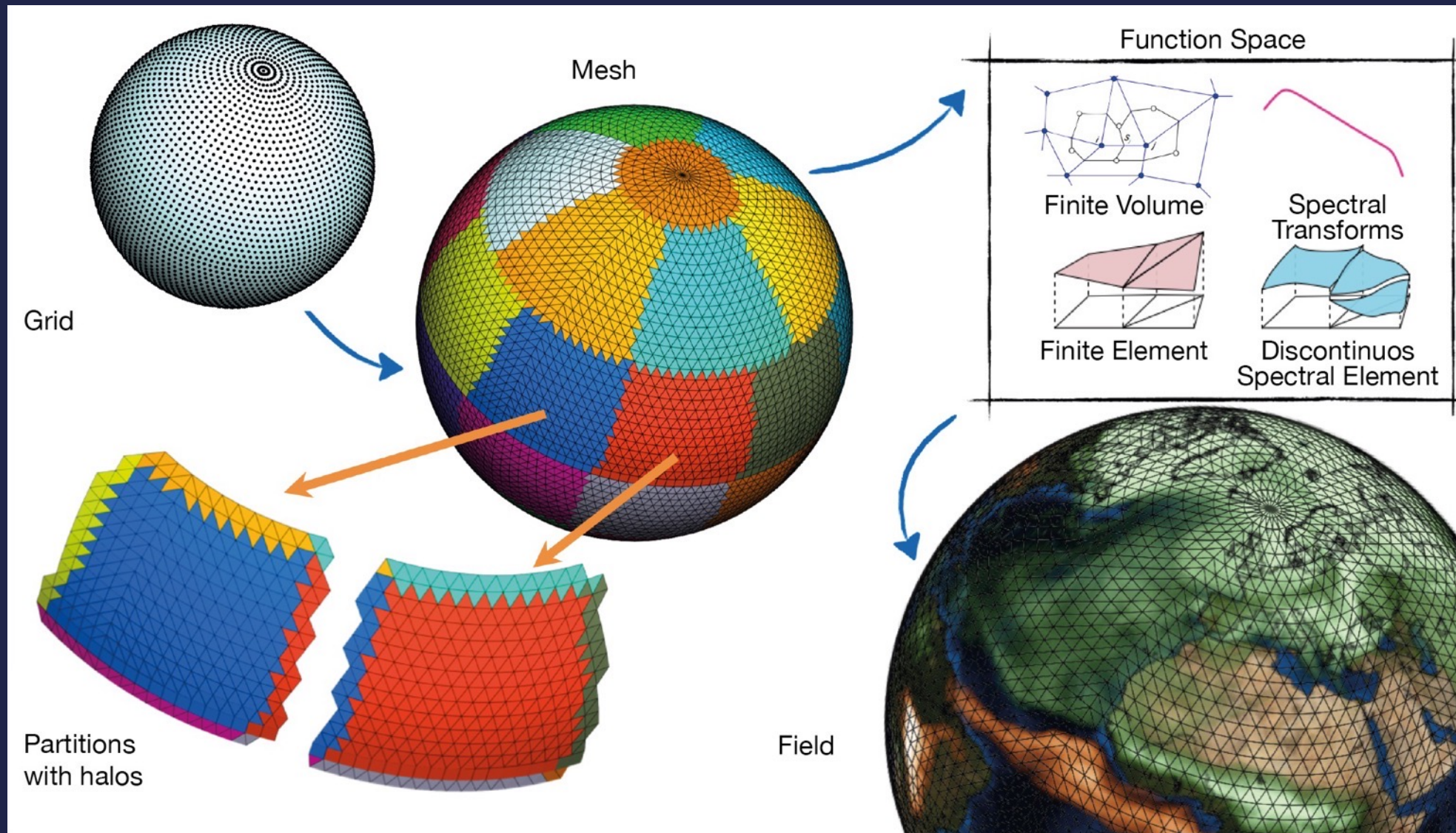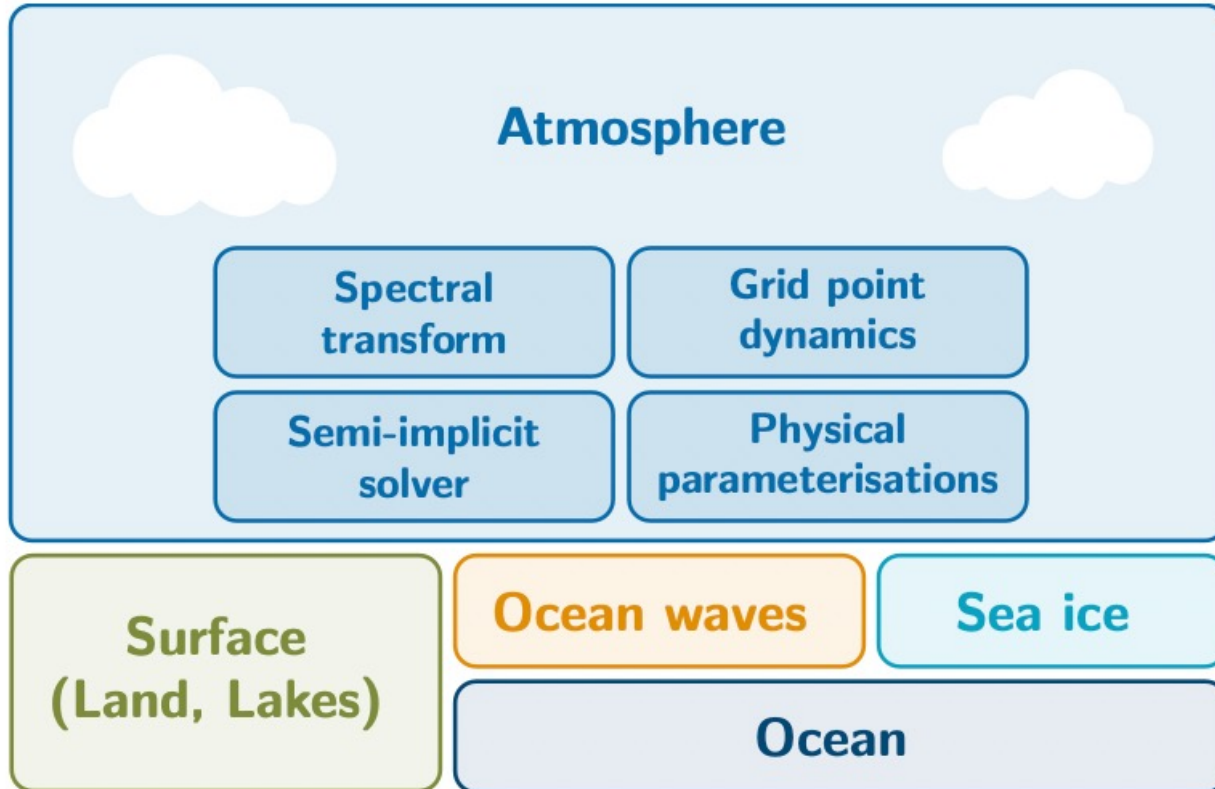
Runtime shares at 9km horizontal resolution (operational HRES)

- Grid point dynamics: 26%
- Physics + radiation: 30%
- Semi-implicit solver: 2%
- Spectral transform: 24%
- Ocean model: 9%
- Wave model: 9%
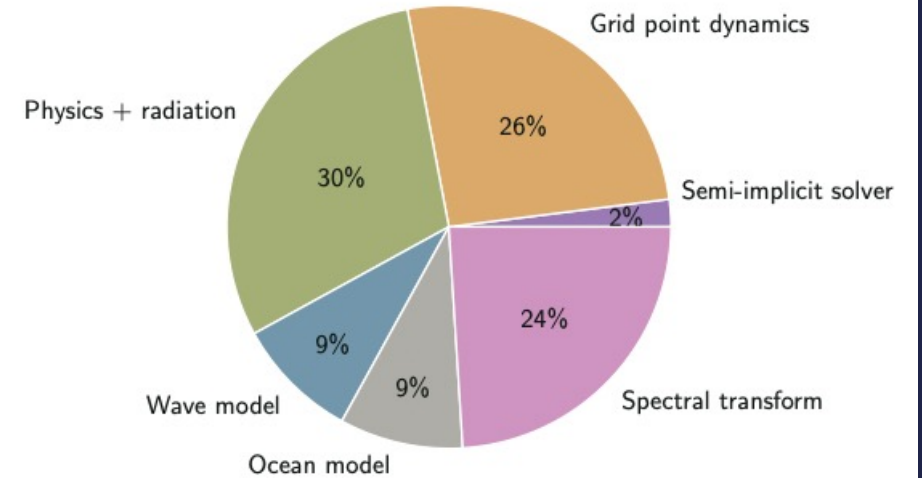
- Similar runtime shares for different model components
- Some parts of the code are frequently changed

ECMWF approach to porting

... hardware adaptation ...

Extract model dwarfs ...

... explore alternative numerical algorithms...

... reassemble model
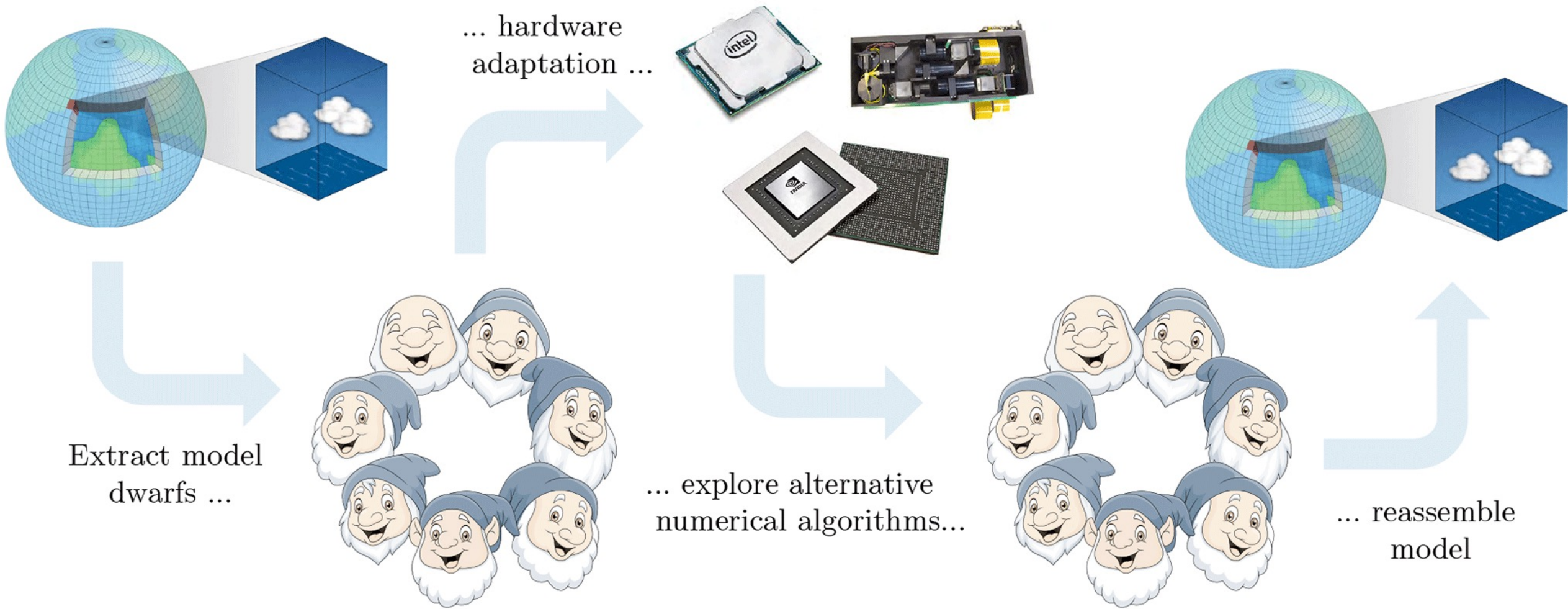
Earth illustration: used under license from GraphicsRF/Shutterstock.com.
Dwarf illustrations: used under license from Teguh Mujiono/Shutterstock.com

Separation of concerns: Gridtools example

Domain scientist:

- Controls grid, resolution, …
- Maintains single source code!
- No hardware specifics!
- No parallelisation specifics!
- No memory layout concerns

DSL Toolchain

- Provides performance portability across a variety of hardware
- Provides parallelisation
- Memory layout
- **Introspection**

https://github.com/GridTools/gridtools

Domain science

Physics

$$\rho\dot{\boldsymbol{u}} = -\nabla p + \rho g - 2\Omega \times (\rho\boldsymbol{v}) + \boldsymbol{f}$$
$$\dot{p} = -\left(\frac{c_{pd}}{c_{vd}}\right)p\nabla\cdot\boldsymbol{u} + \left(\frac{c_{pd}}{c_{vd}}-1\right)Q_h$$
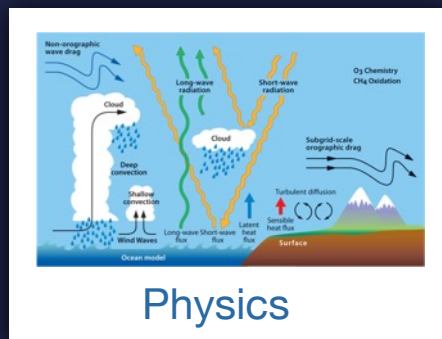$$\rho c_{pd}\,\dot{T} = \dot{p} + Q_h$$

Mathematical description
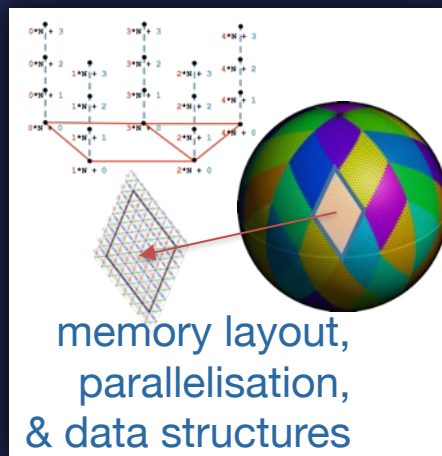
$$\nabla\cdot\mathbf{v} := \frac{1}{A}\sum_{k\in\mathcal{E}}\mathbf{v}_k\cdot\mathbf{l}_k$$

Algorithm development

```
on_edges( sum_reduction, v() * l() ) / A()
```

Domain specific language

Multidisciplinary Abstractions

memory layout, parallelisation, & data structures

Programming models & libraries

Hardware specific instructions

Apply recipes to lots of physics

Application (Fortran)

Driver (Fortran)

Spectral Element
DSL kernel

DESTINATION EARTH

SOURCE-TO-SOURCE TRANSLATION

At build time, transform large sub-trees of complex code base under continuous scientific development

Encode transformations as "complex preprocessor"

- Integration alongside scientific development
- Restrict changes to defined sub-trees
- In-house developed Python tool "Loki"

Bulk processing of IFS physics

- Use bespoke knowledge of IFS code
- Apply transformation recipes tailored to GPU adaptation

https://github.com/ecmwf/atlas
https://github.com/ecmwf/loki (soon)

10

Deploying your ML models

model emulation/acceleration
data fusion
user application models
uncertainty quantification

machine learning models are applied on the fly

data compression

Earth-system observations

observation operators
uncertainty quantification

Data stream

Data selection & pre-processing
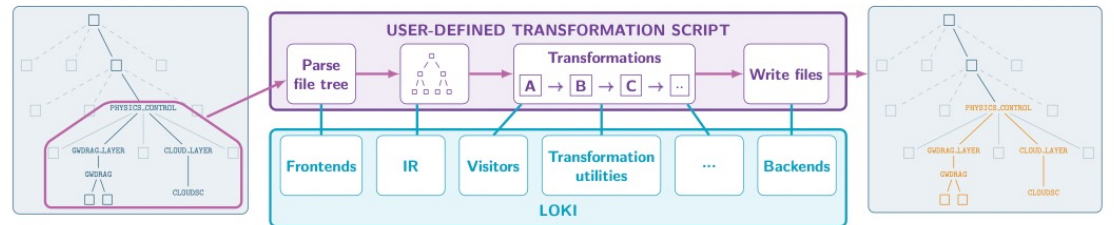
Input Tensor

Interface
• C/C++
• Fortran
• Python

Infero

Output Tensor

Trained model

offline training
user application models

Product generation
Impact modelling

Novel observations

https://github.com/ecmwf-projects/infero

https://github.com/ecmwf-projects/plume (generic plugin system, SOON)

ECMWF

Funded by the European Union

Data handling pipelines for the twins
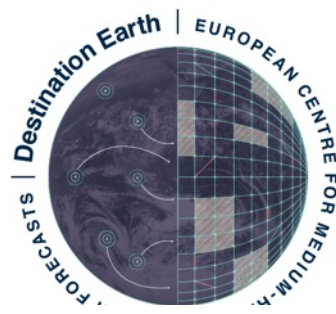
Edge computing: "a part of a distributed computing topology where information processing is located close to the edge, where things and people produce or consume that information"

Streaming data
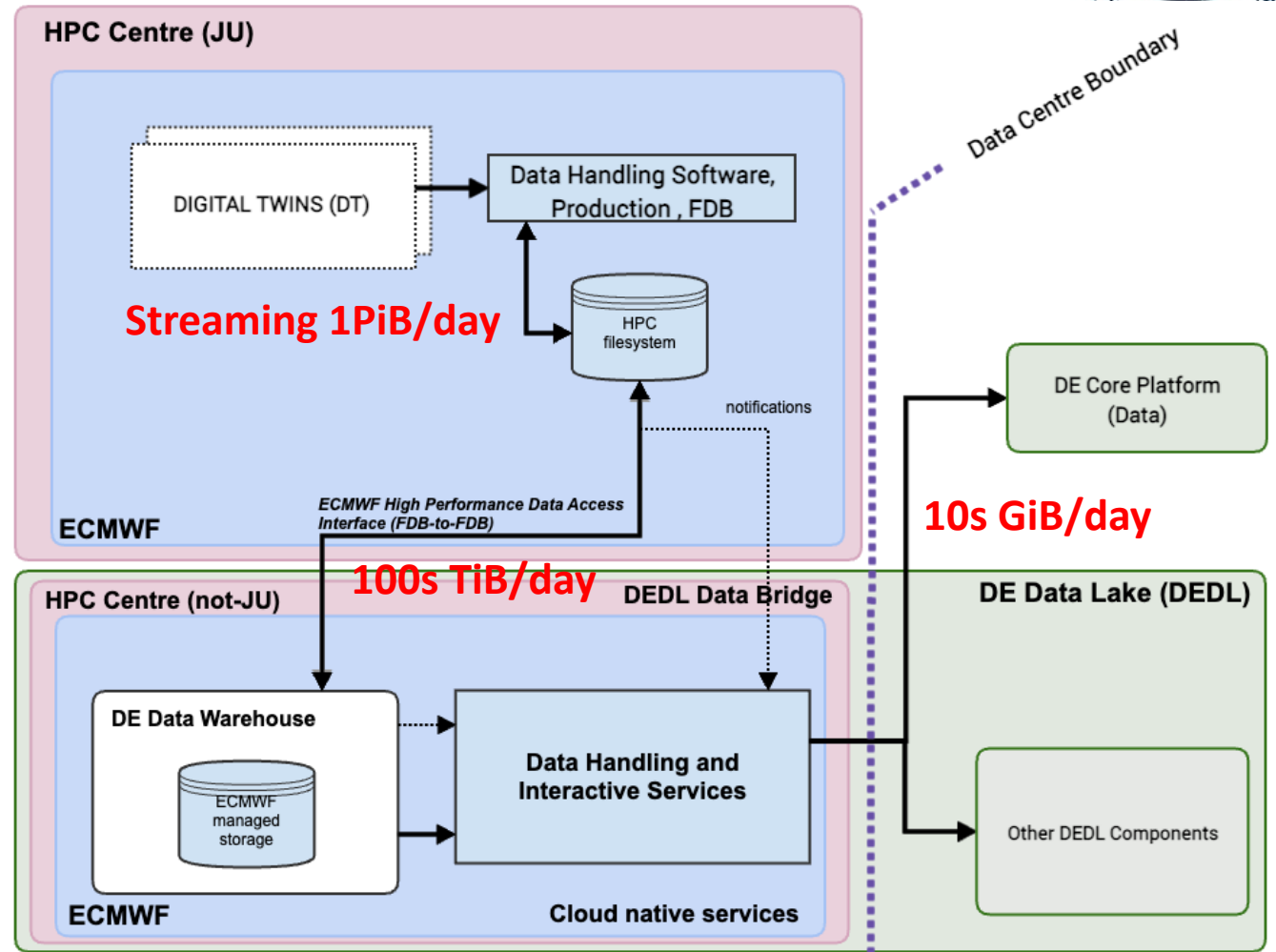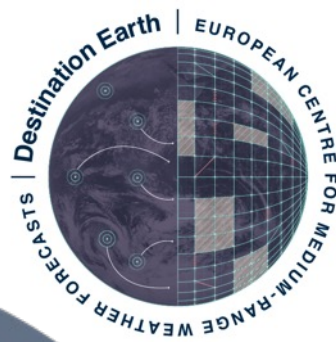On the fly consumption (AI/ML)
Store model state not full 4D-resolution

Storage volume is **capped!**
Data reconstructed NOT stored

HPC Centre (JU)

DIGITAL TWINS (DT) → Data Handling Software, Production , FDB

Streaming 1PiB/day

HPC filesystem

notifications

ECMWF

ECMWF High Performance Data Access Interface (FDB-to-FDB)

100s TiB/day

Data Centre Boundary

DE Core Platform (Data)

10s GiB/day

HPC Centre (not-JU)   DEDL Data Bridge   DE Data Lake (DEDL)

DE Data Warehouse

ECMWF managed storage

Data Handling and Interactive Services

Other DEDL Components

ECMWF   Cloud native services
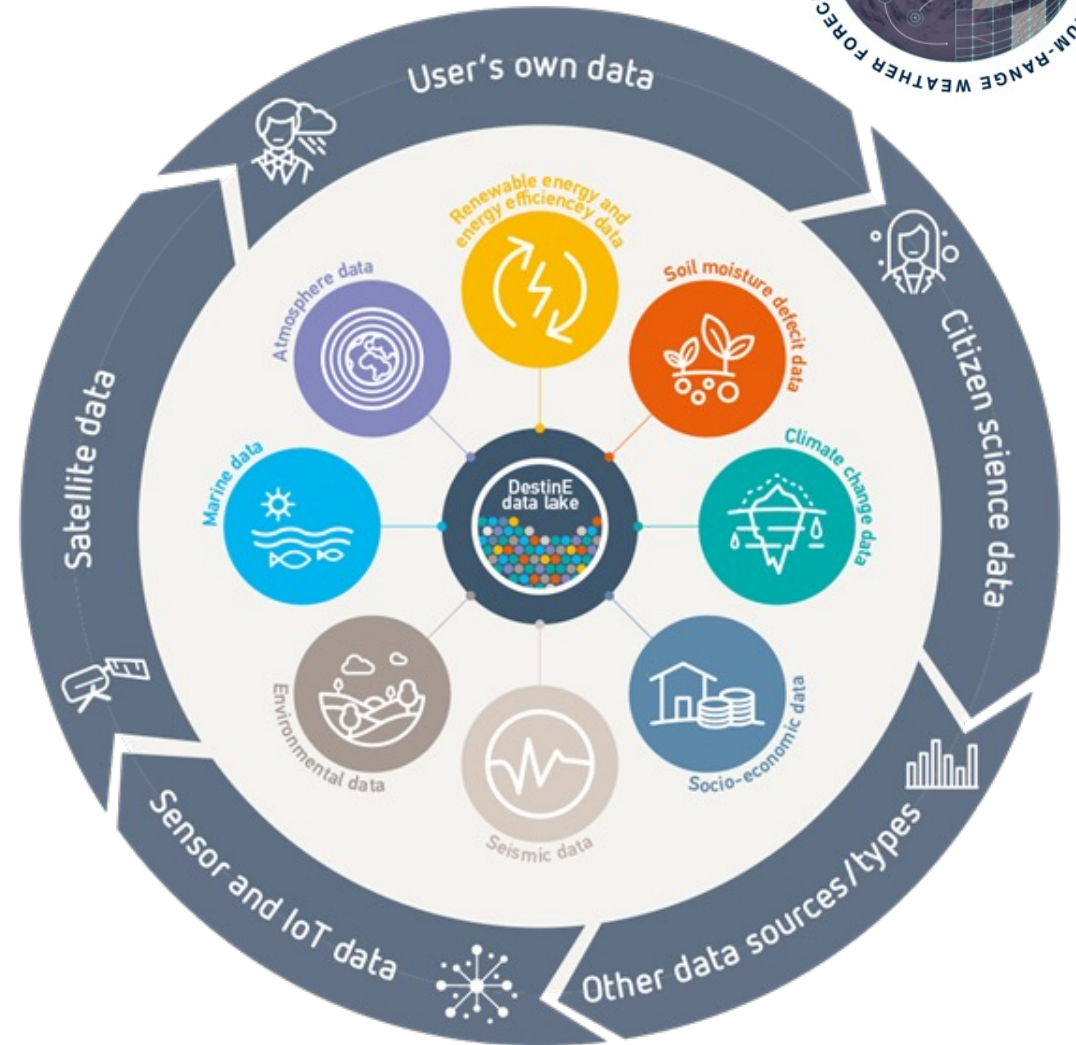
The datalake: harmonized data access

A **data space** is considered an open ecosystem of

distributed, federated actors that share data, applications, and infrastructure amongst each other.

Annex 1: Common European Data Spaces – Timeline

The blue **colour** represents legislative and political initiatives. The green **colour** represents funding initiatives of the Commission. The brown **colour** describes other actions.

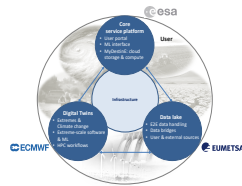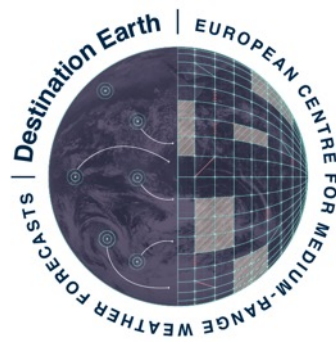| Data Spaces	2021	2022	2023
Manufacturing	DIGITAL CSA to bring stakeholders together, define ref. architecture and building blocks	DIGITAL Grant to deploy and test, using procured building blocks and middleware solutions	Put into practice a federated infrastructure
		Guidelines cross-sectoral data sharing conditions	
Green Deal	DIGITAL CSA to bring stakeholders together, define ref. architecture and building blocks		
	DIGITAL CSA Smart Communities		
	"GreenData4all" evaluation	"GreenData4all" proposal for revising the governance legal framework	
	DIGITAL Destination Earth		
	Review ITS Directive		

Interfaces for the Digital twin Engine

A **data space** is considered an open ecosystem of

distributed, federated actors that share data, applications, and infrastructure amongst each other.
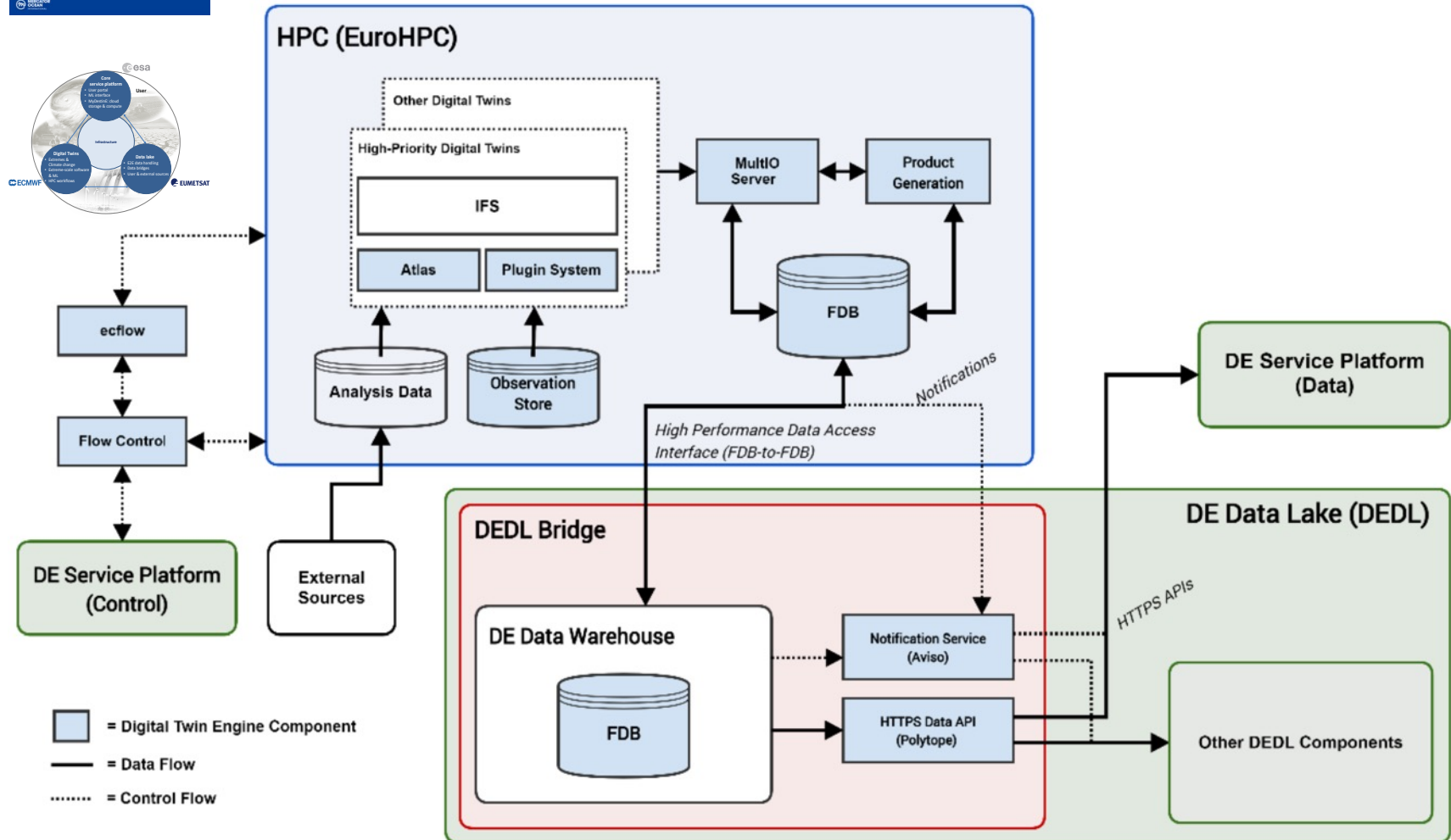

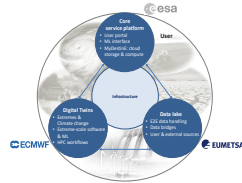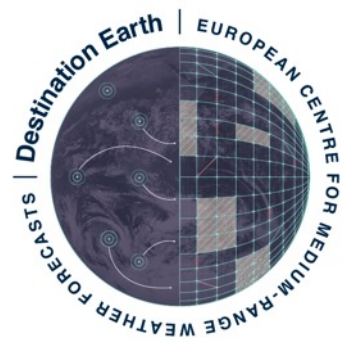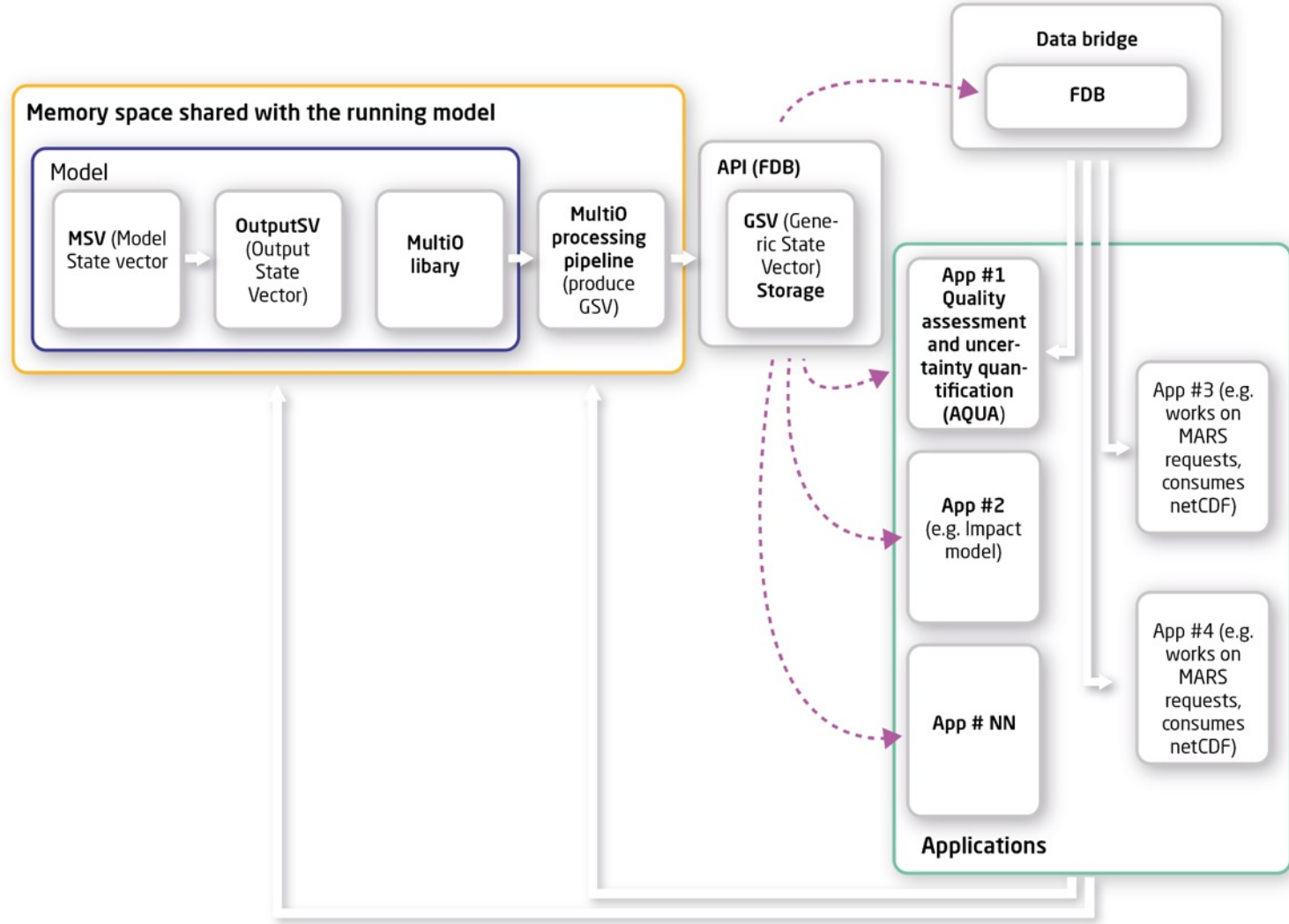https://digital-twin-engine.readthedocs.io
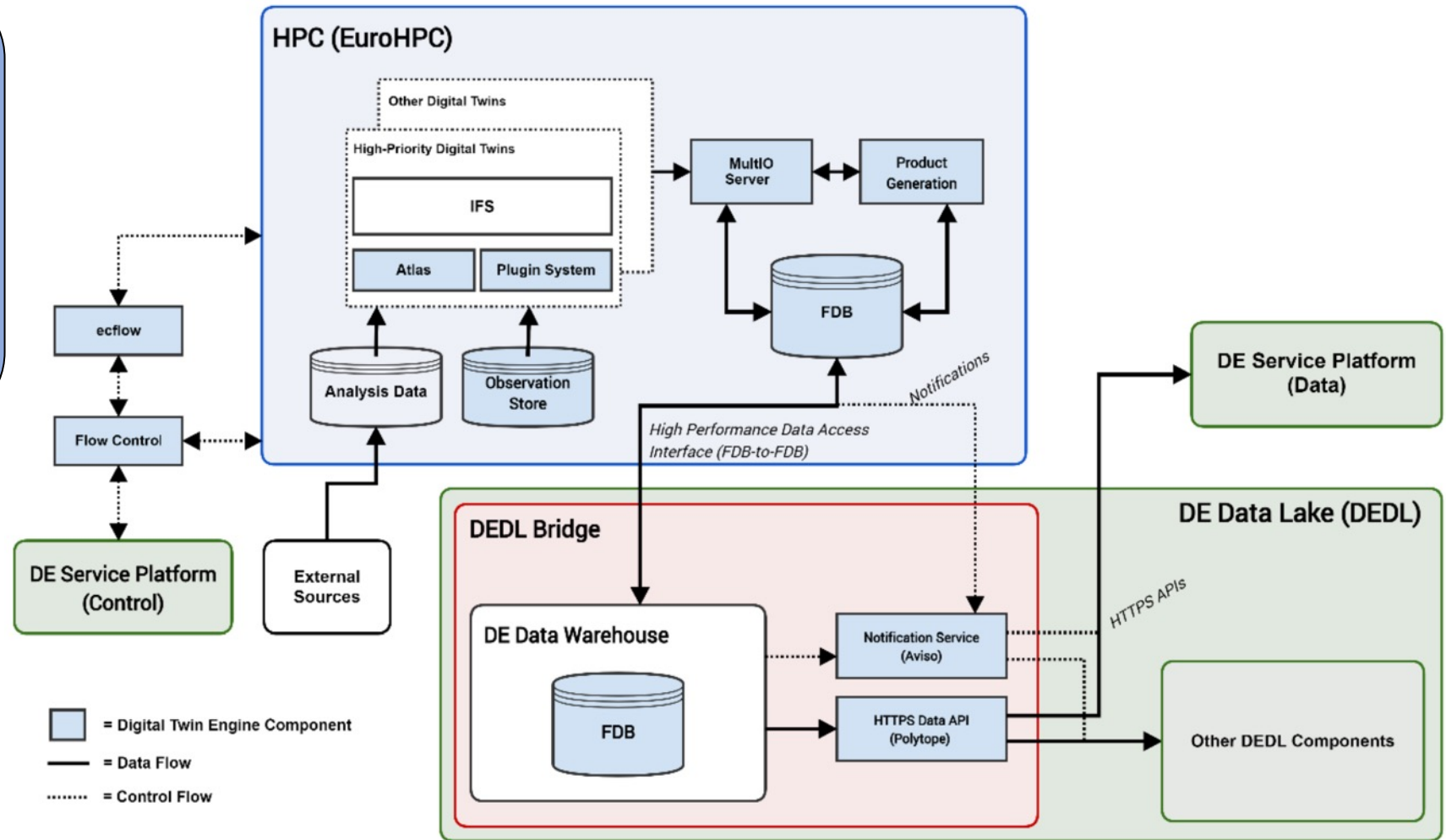
Connecting digital twins

HORIZON-INFRA-2021-TECH-01-01

Connecting digital twins

HORIZON-INFRA-2021-TECH-01-01

Memory space shared with the running model

Model

MSV (Model State vector)

OutputSV (Output State Vector)

MultiO libary

MultiO processing pipeline (produce GSV)

API (FDB)

GSV (Generic State Vector) Storage

Data bridge

FDB

App #1 Quality assessment and uncertainty quantification (AQUA)

App #3 (e.g. works on MARS requests, consumes netCDF)

App #2 (e.g. Impact model)

App # NN

App #4 (e.g. works on MARS requests, consumes netCDF)

Applications

Process optimization leveraging digital technology

Capabilities

Digital Continuum

Impact sectors

Extreme-scale capabilities

- Data assimilation & fusion
- Earth-system models
- Earth observation
- Impact models
- Quality and uncertainty tracing
- Interactive workflows

- Physical systems
- HPC & Cloud
- Smart sensors
- Data analytics & cognitive computing
- Internet of things
- Big data

Digital continuum

Artificial Intelligence Methods

Digital Twin

Sectors of socio-economic impacts

- Economics & insurance
- Water resources
- Energy
- Health
- Food & agriculture
- Biodiversity
- Transport & tourism

continuous **simulation & observation**
technical user interaction
scientific theory and adaptation **scenario testing**

Funded by the European Union
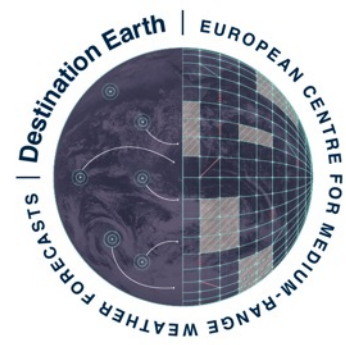
Community engagement

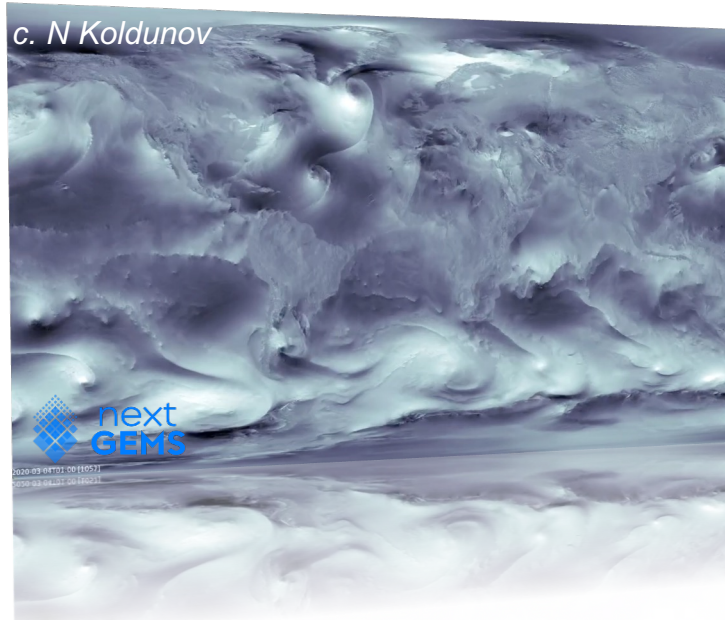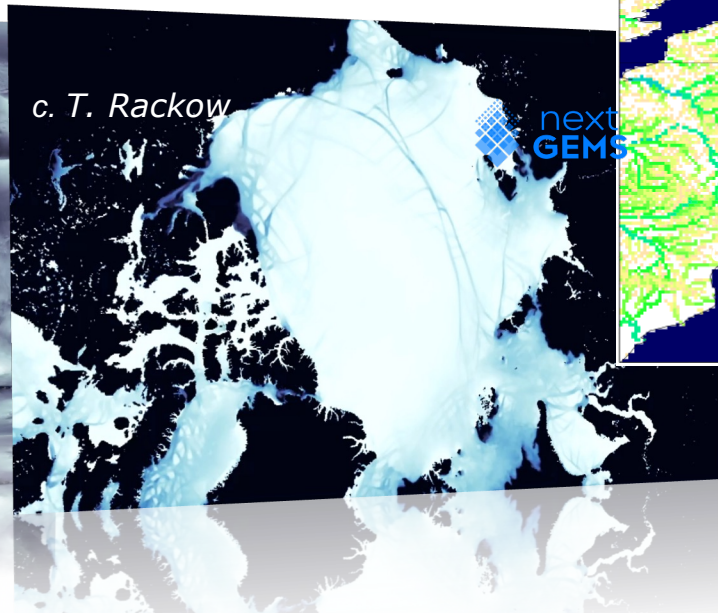https://nikal.eventsair.com/1st-destination-earth-23

# DestinE's Digital Twins: Quality + Impacts + Interaction

1. **Better simulations** based on **more realistic models**

2. **Better ways of combining all observed and simulated information** from entire Earth system = physical + food/water/energy/health **supporting action scenarios**

3. **Interactive and configurable access to all data, models and workflows**

c. N Koldunov

c. T. Rackow

More realistic at *global* scale

More realistic at *local* scale
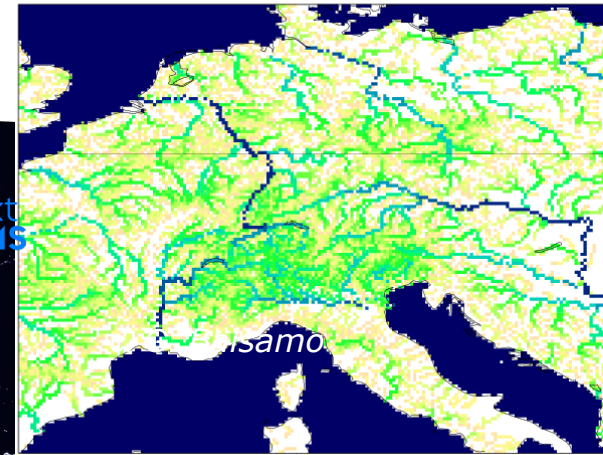
Include impacts where they matter

Trial different adaptation and mitigation scenarios

Thank you

thomas.geenen@ecmwf.int

ECMWF