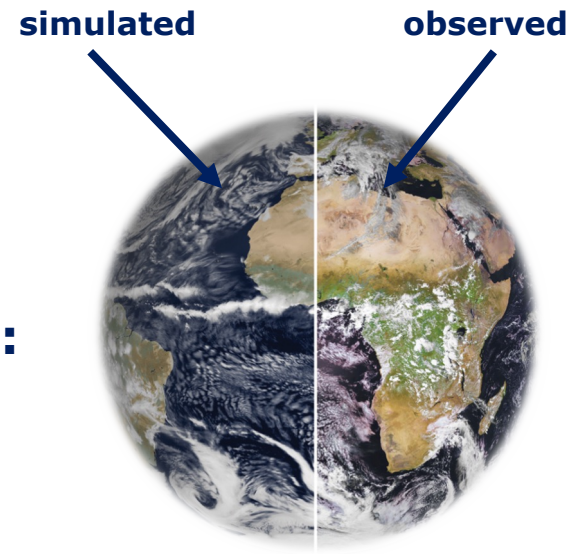
The European Destination Earth project and its potential for boosting machine learning

Peter Bauer, ECMWF

Weather & Climate: What do we really want?

1. Much **better simulations** based on **more realistic models**
2. Better ways of **combining all observed and simulated information** from entire (physical + human) Earth system
3. An information system that provides **convenient & interactive access to all data, models and workflows** so that we can:
   a. **understand and explain change, predict possible futures**
   b. **transform data → information → insights → decisions**

… which leads to Digital Twins
as a key tool to achieve all 3 objectives:



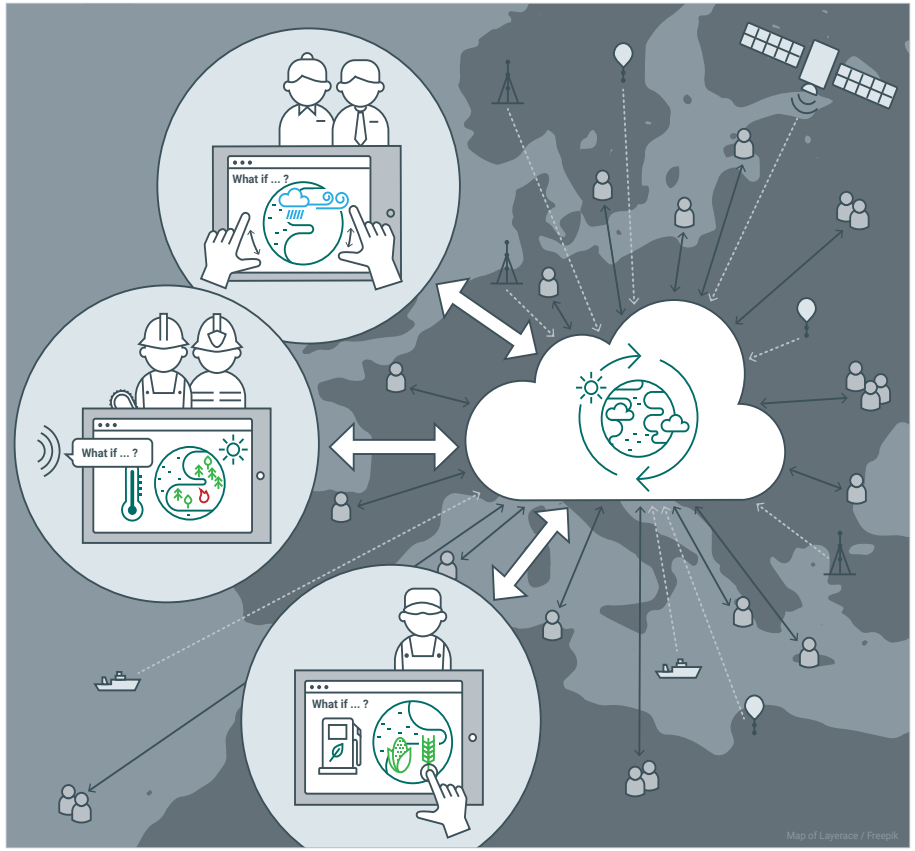
Interactive access to data, software and workflows

Check for updates | comment

A digital twin of Earth for the green transition

For its green transition, the EU plans to fund the development of digital twins of Earth. For these twins to be more than big data atlases, they must create a qualitatively new Earth system simulation and observation capability using a methodological framework responsible for exceptional advances in numerical weather prediction.

Peter Bauer, Bjorn Stevens and Wilco Hazeleger



The predictive modeling challenges

Complex physical phenomena
multiscale, multiphysics

Cyber-physical interactions
software, hardware, sensors, automation

Complex lifecycle
multiple stages, multiple stakeholders

Limited data
observations are noisy, indirect & expensive/intrusive to acquire

Evolving asset state
degradation, damage, maintenance, upgrades

[Courtesy Karen Willcox, U TX]

Make full use of all data to enable decisions

Data → **Information** → **Insights** → **Decisions**

10^{20+} → **10^{6}** → **10^{1}** → **10^{0}** **degrees of freedom**

Machine learning will be essential to *make* this reduction!

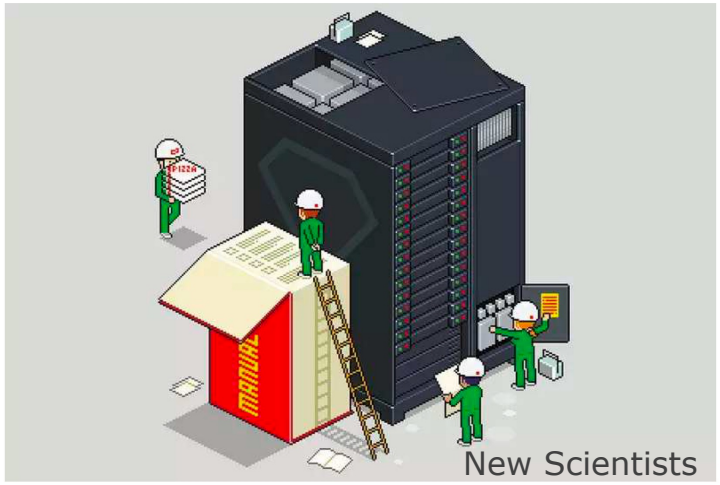
Uncertainty quantification will be essential to *trust* this reduction!

No silver bullet

FEATURE 10 October 2018

Could the world's mightiest computers be too complicated to use?

China, Japan and the US are racing to build the first exascale computer – but devising programmes clever enough to run on them is a different story

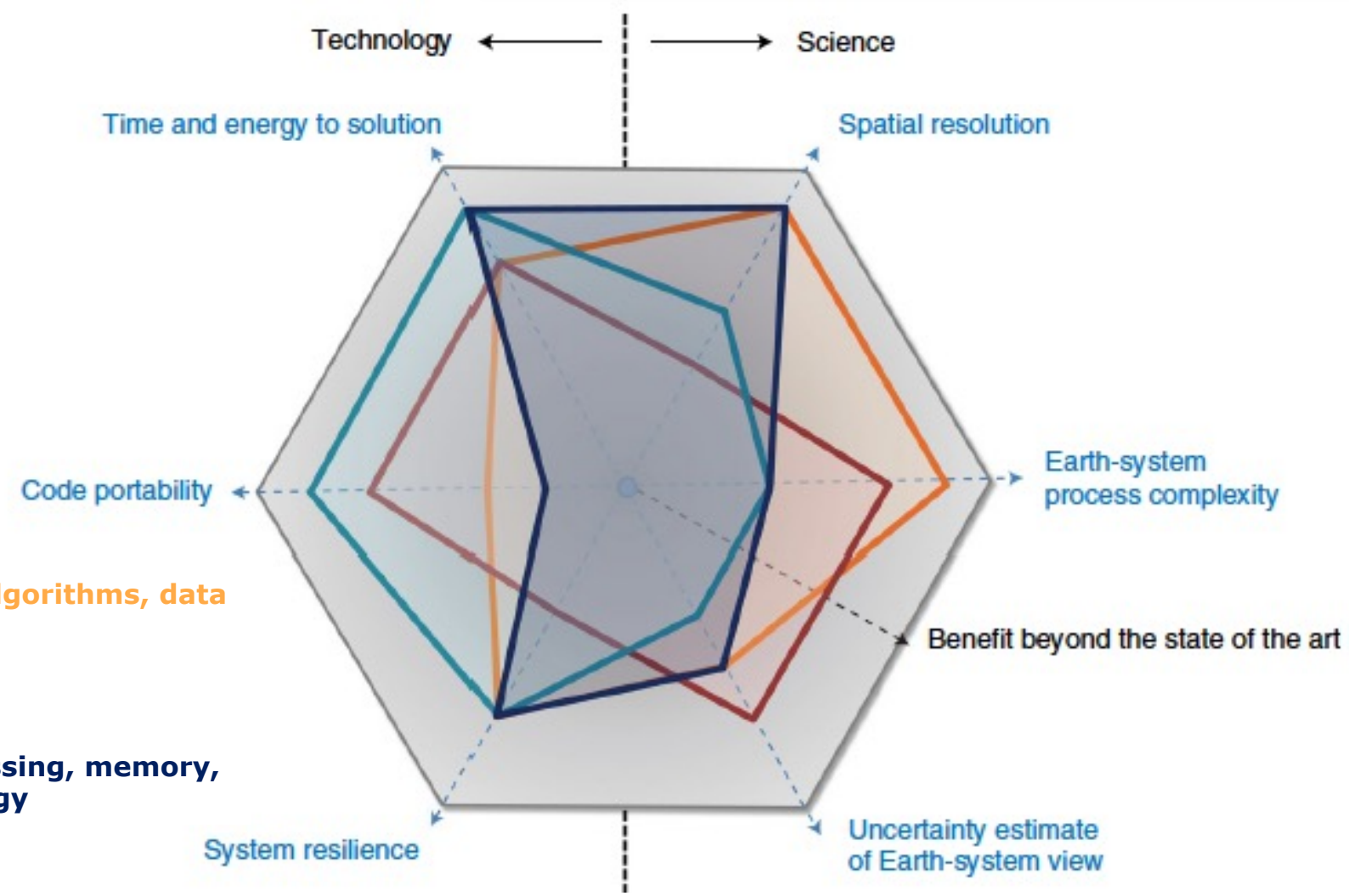


Totto Renna

PERSPECTIVE
https://doi.org/10.1038/s43588-021-00023-0
nature computational science
Check for updates

The digital revolution of Earth-system science

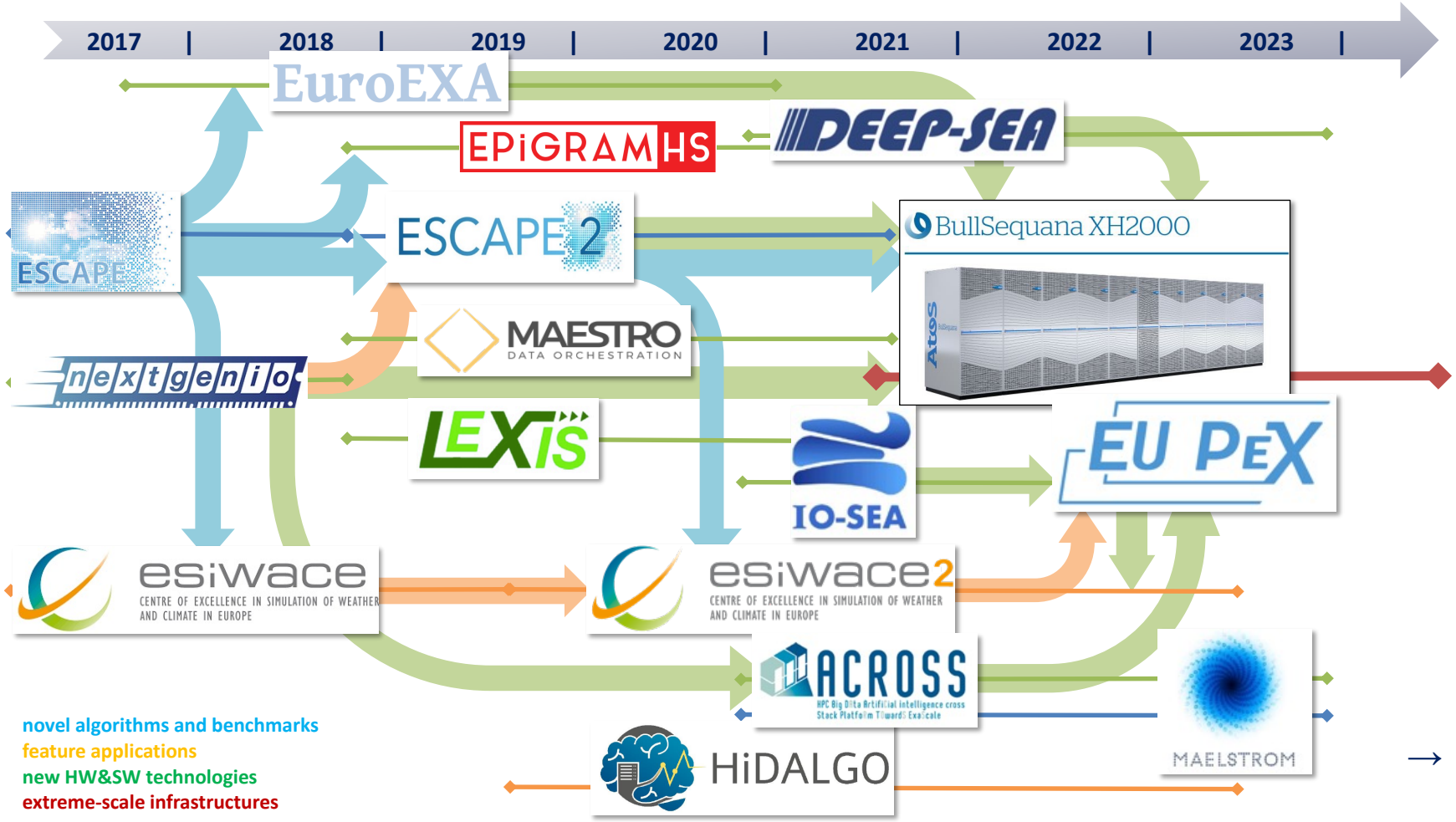
Peter Bauer[1]✉, Peter D. Dueben[1], Torsten Hoefler[2], Tiago Quintino[3], Thomas C. Schulthess[4] and Nils P. Wedi[1]



- **Numerical methods, algorithms, data structures**
- **Machine learning**
- **Programming models**
- **Heterogeneous processing, memory, interconnect technology**

Scalability Programme projects aiming to cover it all

Projects with ECMWF lead/partner roles supported by DG CNECT's FET-HPC & EuroHPC R&I actions:



→ **Programming**

→ **Algorithms**

→ **Architecture**

→ **Data handling & machine learning**

→ **Centres of Excellence**

What is Destination Earth (DestinE)?

Shaping Europe's digital future

Home | Policies | Activities | News | Library | Funding | Calendar | Consultations

Home > Policies > Destination Earth

Destination Earth

Destination Earth (DestinE) and its development of digital earth twins are key to predicting the effects and building resilience to climate-change.

PAGE CONTENTS

Destination Earth (DestinE) aims to develop – on a global scale - a highly accurate digital model of the Earth to monitor and predict the interaction between natural phenomena and human activities. As part of the European Commission's Green Deal and Digital Strategy, DestinE will contribute to achieving the objectives of the twin transition, green and digital.

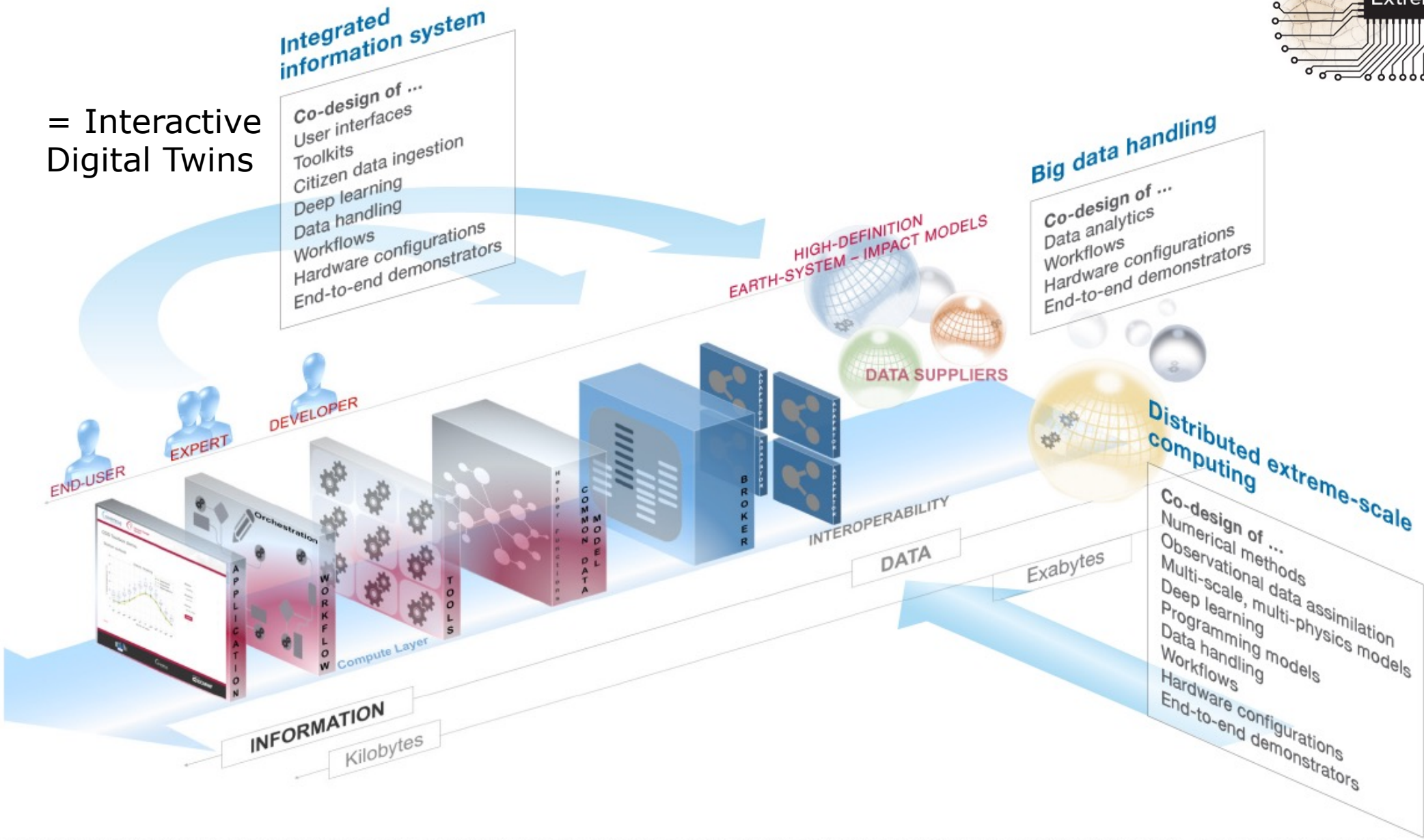
Implementation phase 1: 2021-2024; overall 7-10 programme envisaged

Road to …


May 2008
World Modelling Summit for Climate Prediction:
• Weather & climate prediction synergy
• New Climate Prediction Project
• Dedicated computing
January 2014
ECMWF Scalability Programme launch 1st phase:
• Observation pre-processing
• Data assimilation
• Model reformulation
• Output post-processing
• New architectures
January 2016
Chicheley Hall Meeting:
• Km-scale models
• Dedicated exascale computing
• Computational climate science
April 2016
European Programme on Extreme Computing and Climate (EPECC):
• Km-scale models
• Dedicated exascale computing
• Computational climate science
February-May 2016
EC FET Flagship Consultation
September 2018
ExtremeEarth:
• Km-scale models fluid&solid Earth
• Dedicated exascale computing
• Computational science
• ExtremeEarth Science Cloud
October 2019
ECMWF Scalability Programme launch 2nd phase:
• Data centric workflows
• Portable/performant codes
• Computational weather science
• New architectures
October 2017-September 2018
EC FET Flagship Preparatory Action
May 2019
EC FET Flagship Programme terminated
December 2021
DestinE:
• Digital twins weather&climate
• Core platform & Data lake
• EuroHPC HPC allocations
• Partnership programme
December 2021
EC DestinE action launched

The 2018 ExtremeEarth concept still holds

= Interactive Digital Twins



Partnership triangle of Destination Earth (DestinE)

The DestinE **Digital Twin Engine** (DTE):

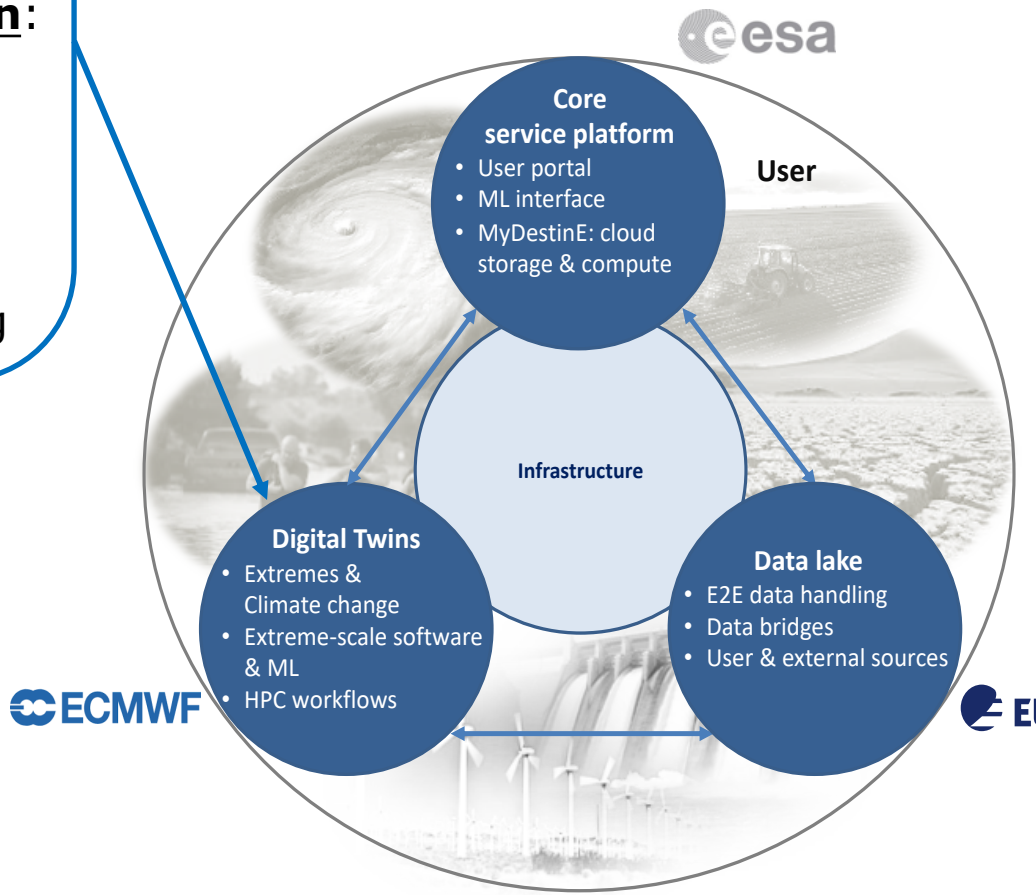
- common system approach to a unified orchestration of Earth-system simulations requiring **large-scale HPC** resources and the fusion of observations with models

Weather-induced and Geophysical **Extremes Digital Twin**:

- capabilities and services for the assessment and prediction of **environmental extremes**

Climate Change Adaptation **Digital Twin**:

- capabilities and services in support of climate change **adaptation policies and mitigation scenario** testing



Main tasks

HPC resource provision, in kind

ECMWF

- ***Digital Twin Engine***
- ***Digital Twin Extremes continuous***
- ***Use cases***

Extremes
(on-demand)
Procured*

- ***Digital Twin***
- ***Use cases***
- ***Digital Twin Engine support***

Climate
(both)
Procured*

- ***Digital Twin***
- ***Use cases***
- ***Digital Twin Engine support***

Visualisation
Procured*

Use cases
Procured*

*https://www.ecmwf.int/en/about/suppliers/destine-procurement/update-destine-itts

Where DestinE needs machine learning

Training datasets:

- Climate DT & Extremes DT will create unprecedented datasets for training
- PB/day output rates require new ways of data lifecycle management

Acceleration:

- Available HPC capacities may not suffice to run simulations/data fusion at full scale

Information extraction/Applications:

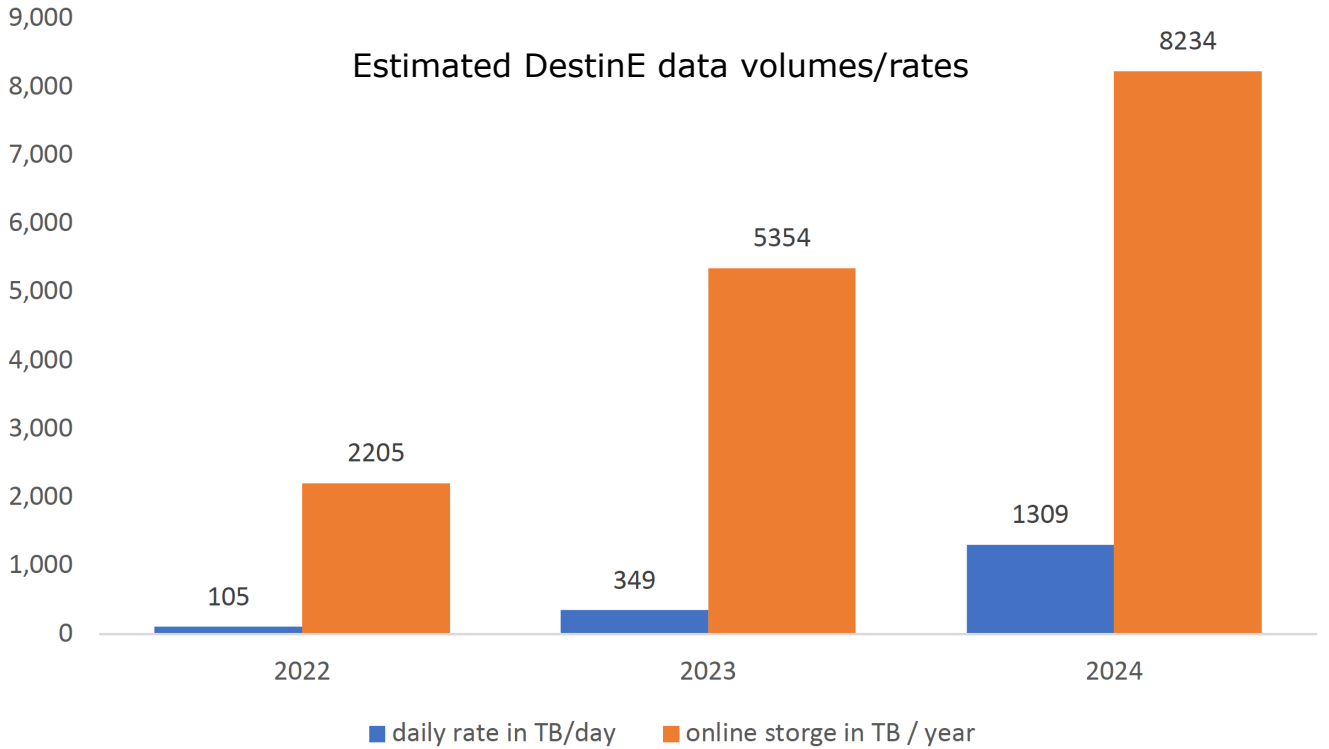
- DT design is with applications in food/ water/energy/health sectors in mind
- Sectors add models/data

Uncertainty quantification:

- Available HPC capacities may not suffice to run fully ensemble based methods
- UQ in application sectors requires new approaches

Visualisation/virtualization:

- Multi-variate (space/time/state vector incl. Applications, UQ) interactivity/rendering requires powerful tools



How DestinE meets Earth-2

future

Earth-2:
- Dedicated = hosts all production and R&D
- 100% allocation = true exascale
- Climate Digital Twin = hybrid ML – physics based modelling

DestinE

EuroHPC LUMI:
- General purpose
- 10% allocation (DestinE)
- Climate prediction DT demonstration

CH Alps:
- General purpose
- tbd
- tbd

EuroHPC MareNostrum5:
- General purpose
- tbd
- tbd

EuroHPC Leonardo:
- General purpose
- 10% allocation (DestinE)
- Extremes prediction DT demonstration

Global Climate Computing Federation

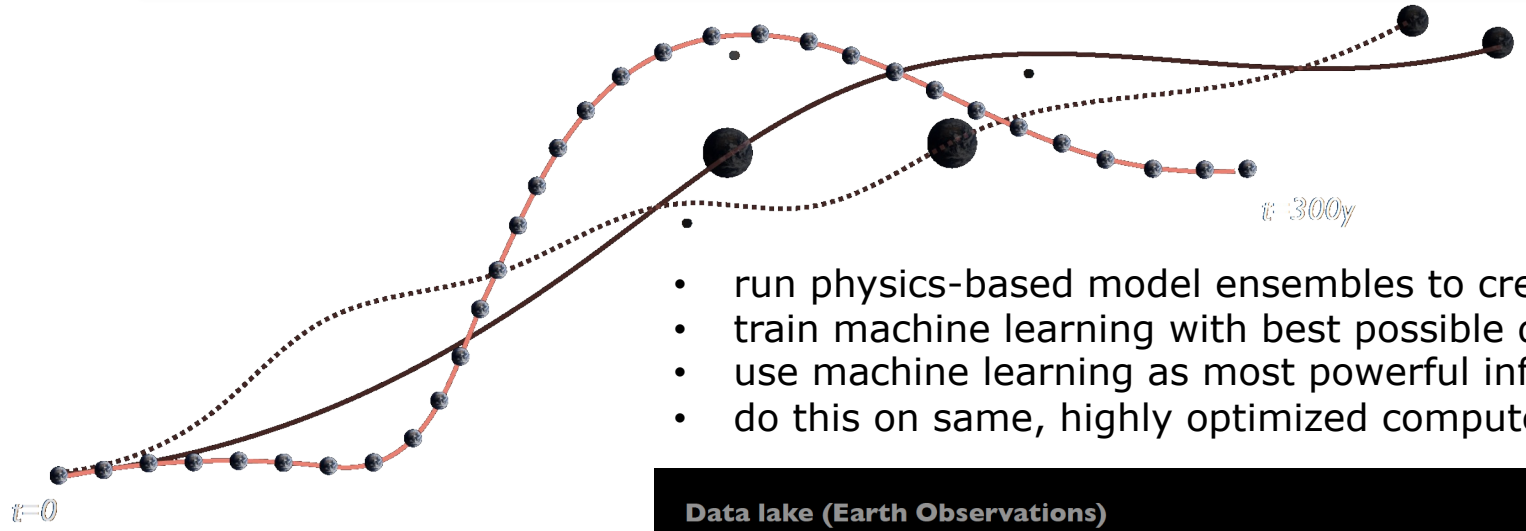
past/present

DE JUWELS:
- General purpose
- Competitive % access (PRACE)
- Benchmarking

US DoE:
- General purpose
- Competitive % access (INCITE)
- Benchmarking

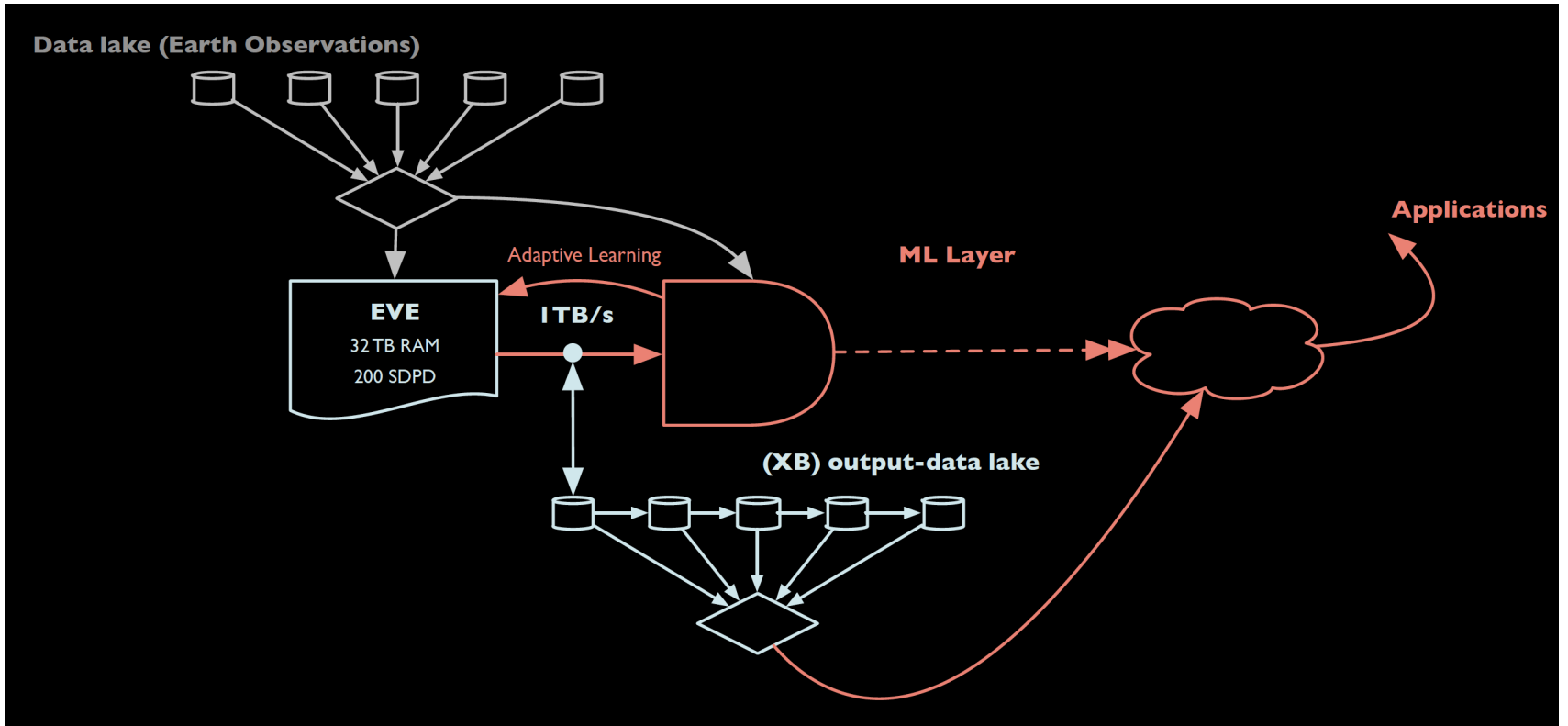
JP Fugaku:
- General purpose
- Competitive % access (TAP)
- Benchmarking

The future is hybrid



- run physics-based model ensembles to create most realistic trajectories
- train machine learning with best possible data; refresh frequently
- use machine learning as most powerful information tool in sliding, full state vector windows
- do this on same, highly optimized compute – data hardware infrastructure

B. Stevens, GTC 2022



Machine learning will be one of the key enablers to make DestinE a success!

Digital Twins:

- *creating more realistic models & better combination of simulations + observations*
- *integrating policy sectors (energy, food, water, health …) in workflows*
- *creating open configuration and information-extraction platforms & access to all data*

Challenges*:*

- *federating resource management across Destination Earth and existing infrastructures*
- *creating synergy across science, technology and service programmes*

Partnerships*:*

- *Destination Earth needs international partnerships – weather, climate, policy sectors*
- *There will be many additional funding opportunities at national & European level*