

# VRE and digital twins: the state of the art and challenges

Gordov E.P.

*Institute of Monitoring of Climatic and Ecological  
Systems SB RAS*

*Federal Research Center for Information and  
Computational Technologies*

**gordov@scert.ru**

# Outlines

- Intro: Digital revolution in Earth sciences
- Virtual research environment: definitions and examples
- Digital twins of systems and processes: definitions and examples
- Challenges and opportunities

# Earth science as a data intensive domain

**Big data (observations + modelling)**

**5V (Volume, Velocity, Variety, Veracity, Value)**

**Data usage in Earth science**

**Data**  **Information**  **Knowledge**  **Services**  
 **Control?**

First step : services for researchers

Virtual Research Environment as result



Next step: services for all users (researchers , decision makers, industry, population) Digital twins of systems/processes

# Definitions

**A virtual research environment (VRE) or virtual research laboratory is an online system helping researchers collaborate. Features usually include collaboration support (Web forums and wikis), document hosting, and some discipline-specific tools, such as data analysis, simulation management, visualisation, teaching tools and publication management.**

Synonyms: “information-computational infrastructure”, "collaborative e-research community", "collaboratory“, "virtual research community", scientific gateways/portals  
VREs - web-based, community-oriented, comprehensive, flexible, and secure working environments conceived to serve the needs of modern science.

**Virtual Research Environments (VREs) - a method of collaboration by researchers that is enhanced through the systematic use of information and networking technology.**

Many funding organisations have established programmes that pilot VREs in discipline-specific contexts over the past years

agenda is promoting the European Research Area; creating transnational VRE-structures helps create a real ERA.

# “A Europe-wide Interoperable Virtual Research Environment to Empower Multidisciplinary Research Communities and Accelerate Innovation and Collaboration” **VRE4EIC**

Funder: European Commission

- 1) understanding complex user requirements across domains;
- 2) improving the quality of VRE user experience by providing user centred, secure, privacy compliant, sustainable environments for accessing data, composing workflows and tracking data publications;
- 3) increasing VRE usage in multidisciplinary research domains;
- 4) improving the interoperability of heterogeneous discovery, contextual and detailed metadata; and
- 5) promoting the exploitation of VRE4EIC solutions to different research communities and commercially.

VRE4EIC covers all EU member states and EFTA countries, and will affect 70,000 researchers all over Europe.



## The EVER-EST Virtual Research Environment

A state of the art Research Lifecycle  
Management platform for Earth  
Science

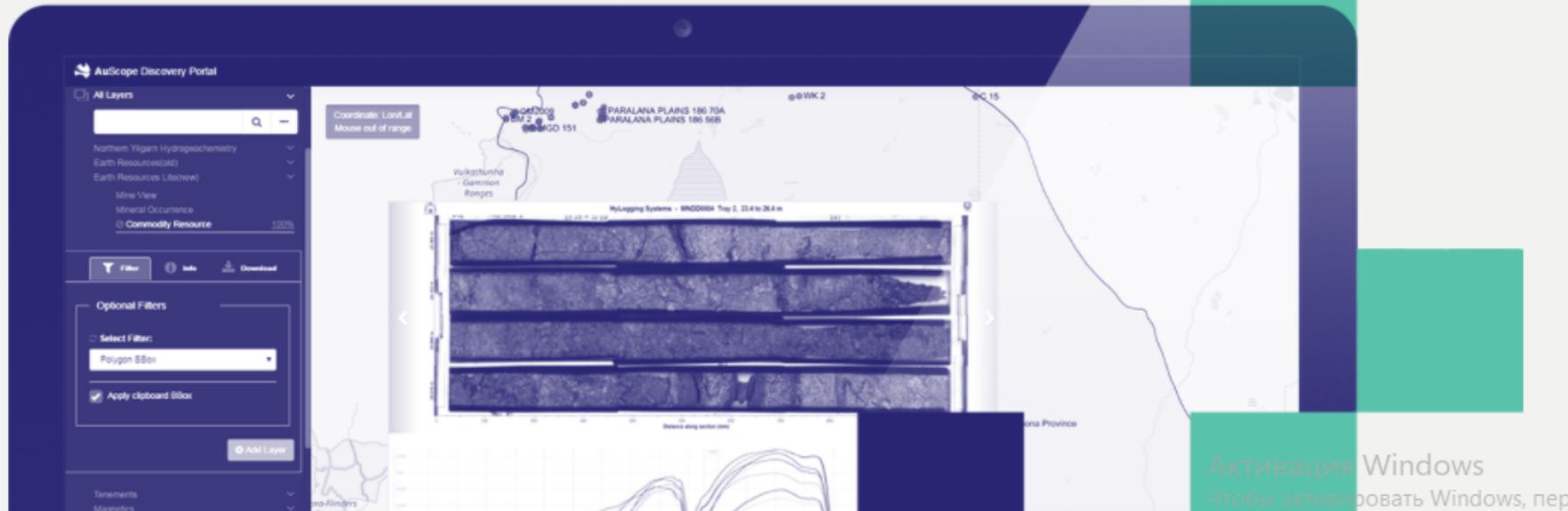
The EVER-EST VRE provides services to share Earth Science data and information allowing researchers to discover, access, and process heterogeneous data, algorithms, results and experiences

## AuScope Virtual Research Environment (AVRE)

### Data, Visualisation & Analytics

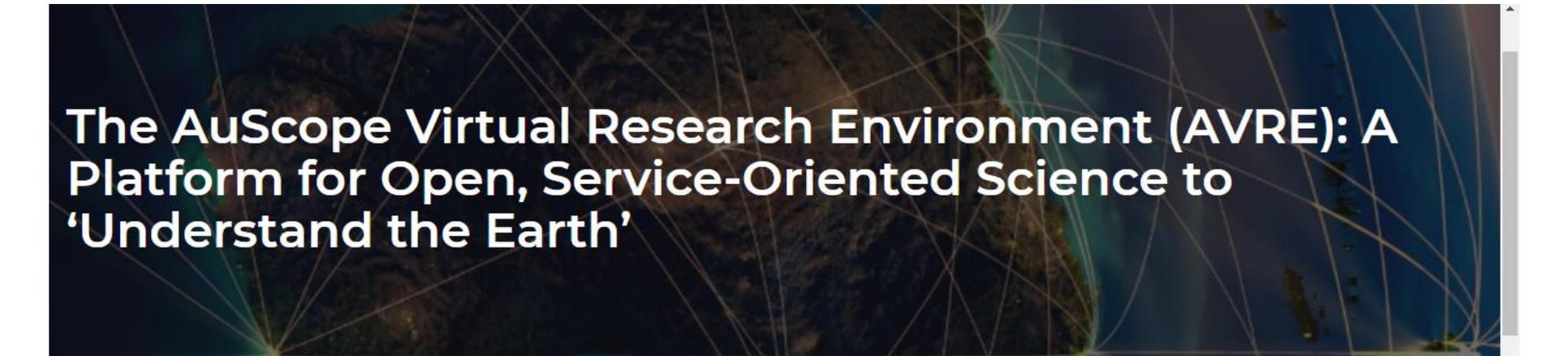
[Visit the AVRE Store](#)

Click here to get straight to data, visualisation & analytics.



The screenshot displays the AuScope Discovery Portal interface. On the left, there is a sidebar with 'All Layers' and 'Optional Filters'. The main area shows a map with various data points and labels, including 'Wulkathurra Gossion Ranges' and 'PARALANA PLAINS 186 70A'. Below the map, there is a 3D geological model showing different layers and structures. The interface is dark-themed with blue and green accents.

Активация Windows  
Чтобы активировать Windows, пер



## The AuScope Virtual Research Environment (AVRE): A Platform for Open, Service-Oriented Science to 'Understand the Earth'

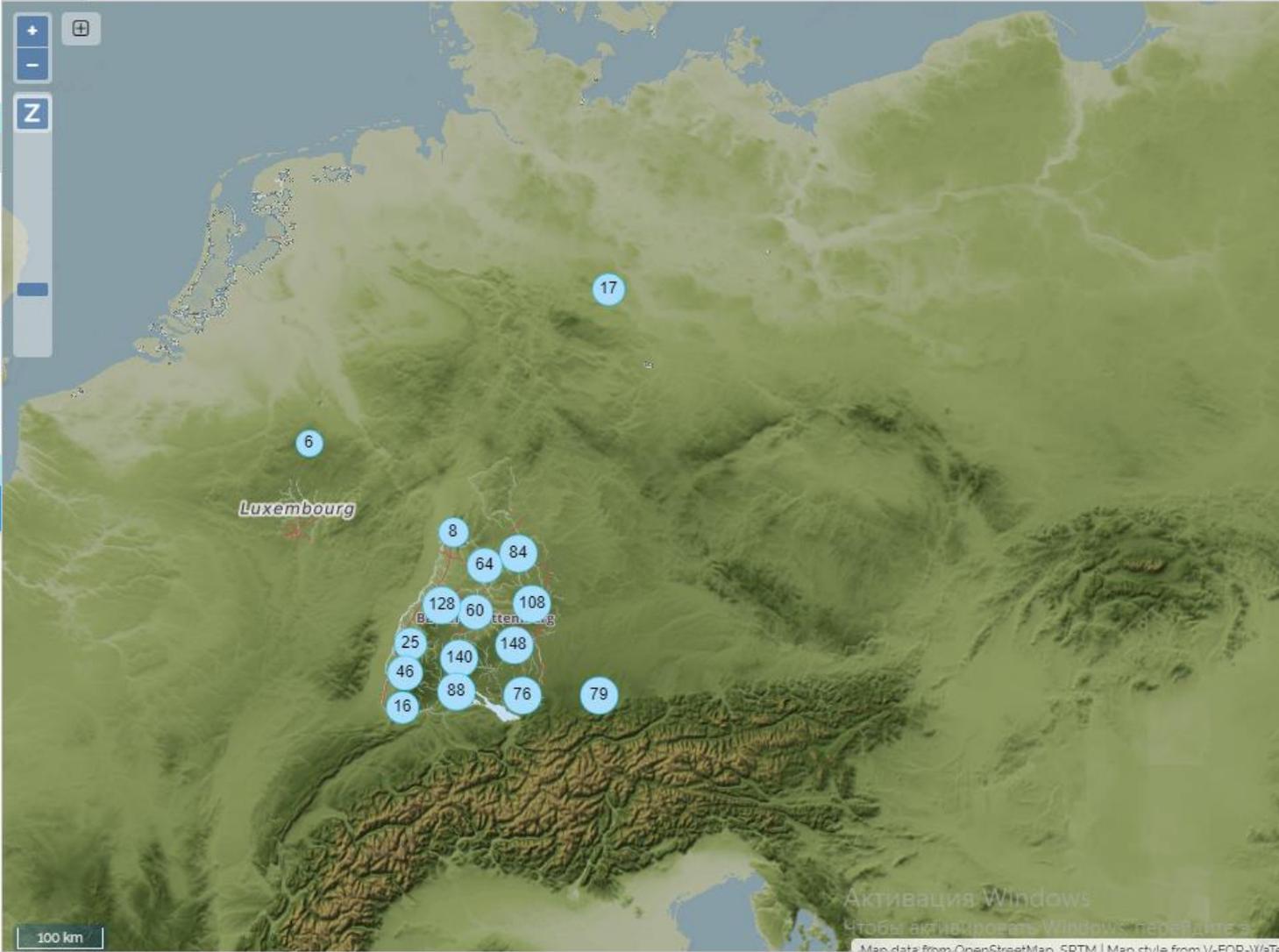
Since 2006, NCRIS projects (AuScope, NCI, ANDS, Nectar, RDS), and Government have collaborated on building a suite of data portals, tools, software and virtual laboratories to support a diverse community of Earth scientists operating on a range of computational facilities including HPC, cloud, on-premise servers and desktops.

The aspiration was to realise a service-oriented science platform that will empower data assimilation and modelling across three networks: geophysics, geochemistry and geology.

The end goal is to ensure Australian Earth science data and analytics can play a leadership role in next-generation transdisciplinary research to more comprehensively 'Understand the Earth'.



Map Table



Filter menu

Quick Draw

Variables

Licenses

Entries

Pass datasets to data store

Reset filter menu

Data store

Remove all datasets

The virtual research environment V-FOR-WaTer aims at making this data abundance available in an easy-to-use [web portal](#).

Environmental scientists get access to data from different sources, e.g. state offices or university projects, and can share their own data through the portal.

Tools help to easily pre-process and scale the data and make them available in a consistent format. Tools for more complex scientific analyses and uncertainty propagation will be included.

The possibility to store workflows together with the tools and respective data ensures reproducible data analysis.

Additionally, interfaces with existing data repositories enable easy publication of the scientists' data directly from the portal.

V-FOR-WaTer addresses the needs of researchers of hydrology and environmental science to not only find and access datasets but also conduct efficient data-based learning with standardised tools and reproducible workflows.

# M-VRE: The MOSAiC - Virtual Research Environment

The M-VRE project is funded by the German BMBF initiative „MARE:N – Polarforschung/MOSAiC“ from July 2021 to June 2024. This project is a collaboration of AWI Bremerhaven, DKRZ Hamburg and DLR Jena.

The main goal of the M-VRE project is to provide novel methods and software tools for the [MOSAiC](#) consortium, the global climate community, and the general public for efficiently exploring, analysing, and visualizing [MOSAiC data](#) in an online and user-friendly manner. This enables scientists to exploit large, complex, and heterogeneous data sets to answer fundamental research questions of the MOSAiC expedition.



The VRE offers online data analysis, exploration (data cubes), and visualization (maps, scatter plots, time series, section plots and more) of MOSAiC data as well as automated data quality control through Deep Learning methods.

The following Infrastructure Gateways serve a number of Research Communities, Initiatives, and Projects, each of which having a specific scientific domain and dedicated Virtual Research Environments Virtual Research Environments for Data-Driven Research



## AGINFRAPlus Gateway, 14 VREs / VLabs

AGINFRA+ aims to exploit core e-infrastructures such as D4Science, EGI.eu, OpenAIRE and EUDAT, towards the evolution of the AGINFRA data infrastructure, so as to provide a sustainable channel addressing adjacent but not fully connected user communities around Agriculture and Food. AGINFRAPlus Gateway helps the project to develop and provide the necessary specifications and components for allowing the rapid and intuitive development of variegating data analysis workflows, where the functionalities for data storage and indexing, algorithm execution, results visualization and deployment are provided by specialized services utilizing cloud based infrastructure(s).



## ARIADNEPlus Gateway, 6 VREs / VLabs

The ARIADNEplus Project builds on the ARIADNE results, extending and supporting the research community that the previous project created and further developing the relationships with key stakeholders such as the most important European archaeological associations, researchers, heritage professionals, national heritage agencies and so on. The ARIADNEPlus Gateway helps the project to develop a data infrastructure embedded in a cloud that offer the availability of Virtual Research Environments where data-based archaeological research may be carried out. The project will furthermore develop a Linked Data approach to data discovery, making available to users innovative services, such as visualization, annotation, text mining and geo-temporal data management. Innovative pilots will be developed to test and demonstrate the innovation potential of the ARIADNEplus approach.



## Blue-Cloud Gateway, 10 VREs / VLabs

Blue-Cloud, a thematic EOSC cloud to better understand and manage the ocean sustainability, through a set of five compelling pilot Blue-Cloud demonstrators. It develops and deploys, the Pilot Blue Cloud as a cyber platform bringing together and providing access to: 1) multidisciplinary data from observations and models, 2) analytical tools, 3) computing facilities essential for key blue science use cases.



## D4Science.org Gateway, 42 VREs / VLabs

D4Science is a Data Infrastructure connecting +7000 scientists in 44 countries, integrating +50 heterogeneous data providers, executing +13.000 models & algorithms/month, providing access to over a billion quality records in repositories worldwide with 99.7% service



## CLIMATE

A specialized geoportal for analyzing regional climate and environmental changes provides access to resources to four categories of users.

- The portal provides full access to tools for supporting the modeling and monitoring of regional climate change based on spatial data services in the "**Climate Analysis**" section for specialists in related fields.
- In the "**Educational Resources**" section, undergraduate and graduate students have access to a course and practical tasks, the main purpose of which is a comprehensive analysis of current climate changes and their possible consequences.
- The section "**General Public**" contains a basic information course on climate change. The course outlines and popularly illustrates the basic concepts and problems of modern climate changes and their possible consequences.
- The section "**Climate Characteristics Database**" for decision makers offers ready-made data packages on key characteristics and extreme climate manifestations in various formats for downloading and further working with them in desktop GIS.

Limited functionality is provided by the portal by default. Full functionality is available for authorized users only. Authorization is under the heading labeled with a special icon  in the upper right corner of the screen. To register please contact portal administrator [webmaster@scert.ru](mailto:webmaster@scert.ru).



Climatic analysis ▾

Educational resources ▾

General public

Изменения климата и адаптация ▾

Climate glossary

Links

Climate characteristics archive



RU EN

Catalogue / Task  



↑ Up one level

WEB MAPPING SYSTEM 4.0A

Web mapping system 3.0a

Web mapping system 3.0 demo

Web mapping system 4.0a

CLIMATE - Information-computational system for analysis of climate and ecological change

File ▾ Edit ▾ Navigation ▾ Layer ▾ Map ▾ Help ▾



Location 



Layers

- All layers
- User vector data
- Google Hybrid
- Google Satellite
- Google Physical
- Bing Aerial
- Bing Aerial with Labels
- Bing Road

Legend 



Активация Windows

Чтобы активировать Windows, перейдите в раздел "Параметры"

New challenge - necessity of adaptation to climate change

Climatic services

Additional users: decision makers, population, agriculture, industry, etc.

First responses

# World Climate Services Programme

The World Climate Services Programme (WCSP) contributes to improving the availability of, and access to, reliable climate data, monitoring and forecasts. The Programme promotes the development of appropriate institutional mechanisms and operational infrastructure to generate, exchange and disseminate quality information on climate **and provides guidance on related applications and services.**

The WCSP spans four inter-related areas:

- Climate data and analysis;
- Climate monitoring, watch and prediction;
- Climate system operation and infrastructure; and
- Climate adaptation and risk management.**



X close

We provide authoritative information about the past, present and future climate, as well as tools to enable climate change mitigation and adaptation strategies by policy makers and businesses.

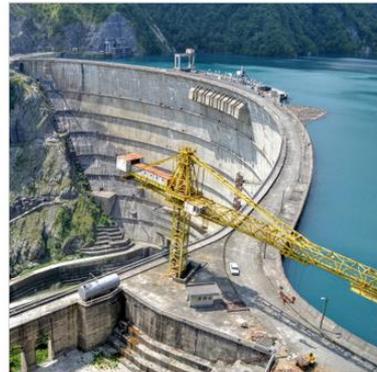
Key products and services



Climate bulletins



Climate Data Store



Data in action



In focus

The European State of the Climate 2018, an essential snapshot of the region and a useful benchmark for future assessments of the environment.

Read More >







## DATA LIBRARY OF CLIMATE CHARACTERISTICS INCL. EXTREMES

## TO CALCULATE THE CHARACTERISTICS OF THE MODERN PERIOD ARE USED:

1. High spatial resolution 0.25x0.25 (~30 km) reanalysis data **Era5** (Fifth generation of ECMWF atmospheric reanalyses of the global climate, <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>) from 1979 to 2020 years.

## TO CALCULATE THE CHARACTERISTICS OF THE FUTURE PERIOD ARE USED:

1. High spatial resolution data T359 (~50 km) **CMIP6** project data (Six phase of Coupled Model Intercomparison Project, <https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6>) of subproject **HighResMIP** (High Resolution Model Intercomparison Project) according to the global model **CNRM-CM6-1-HR** (Centre National de Recherches Météorologiques (CNRM) and Cerfacs) according to the RCP 8.5 scenario for the period from 2015 no 2050 years.
2. High spatial resolution data 0.22x0.22 (~25 km) from the **CORDEX** (The Coordinated Regional climate Downscaling Experiment, <https://www.cordex.org>) project for Central Asia (CAS-22) according to the global model of the Earth system **MPI-ESM-LR** (model of the Max Planck Institute for Meteorology), the calculation results of which were used as the basis for downscaling using the regional model **REMO2015** of the GERICS center (Climate Service Center Germany); according to the RCP 8.5 scenario for the period from 2006 no 2100 years.
3. Spatial resolution data T127 (~100 km) **CMIP6** project (Six phase of Coupled Model Intercomparison Project, <https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6>) data according to the global model of the Earth system **MPI-ESM-HR2** (model of the Max Planck Institute for Meteorology) according to the SSP585 scenario for the period from 2015 to 2100 years.

>> [Anomalies of meteorological characteristics](#)

>> [Climate Indices](#)

Активация Windows

Чтобы активировать Windows, перейдите в раздел "Параметры".

# Digital twins as a response to the challenge

A **digital twin (DT)** is

- a virtual representation that serves as the real-time digital counterpart of a physical object or process.
- a digital representation of a physical objects or process
- *also referred to as digital shadow, digital replica or digital mirror - is a digital representation of a physical asset.*

**DT** optimally combines simulations and near-real-time observations to monitor the evolution of a system or process. *Linked to each other, the physical asset and digital twin regularly exchange data.*

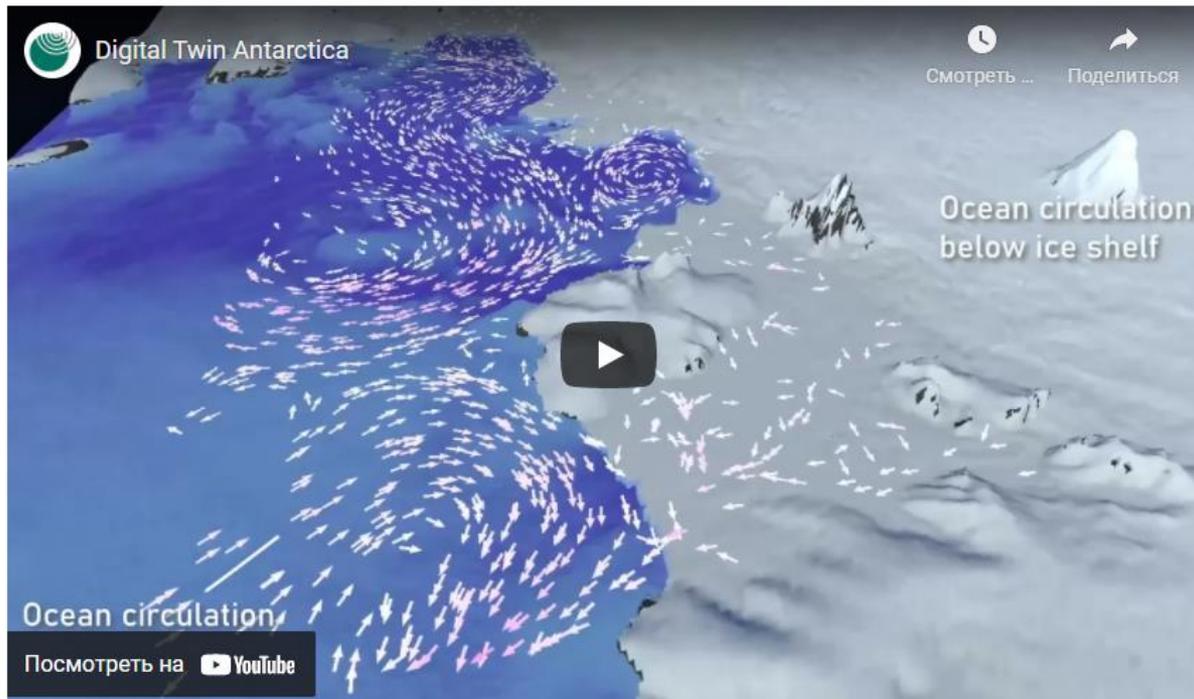
**DT** strongly relies on geospatial data, and the processing and analytics thereof. Cloud environments provides the flexibility and scalability needed to cope with the potential enormous geospatial datasets.

*Technology like AI, machine learning, sensors and IoT allow for dynamic data gathering and right-time data exchange to take place*

Digital twins enables analysis of current state, future predictions and simulations of possibilities of objects in cyberspace. Based on observation and simulation results DT also enables feedback to real world to enable an intelligent control the real world object.

**DT** of the environment can help reaching sustainability goals and tackling climate change related issues.

The first practical definition of digital twin originated from NASA in an attempt to improve physical model simulation of spacecraft in 2010.



# Digital Twin Antarctica

“By harnessing satellite observations, numerical simulations, and Artificial Intelligence, we have built a twin of the Antarctic ice sheet system, its hydrology, surrounding ocean, atmosphere, and biosphere. We have used the Antarctic twin to track the whereabouts of melt water on and under the ice sheet, and to explore how fringing ice shelves melt under various hydrology scenarios.” (Noel Gourmelen, University of Edinburgh)

## Digital Twin Hydrology

Luca Brocca (NRC, Italy) “In the ESA Digital Twin Earth Hydrology project, we have developed a 4D reconstruction of dynamic hydrology at unprecedented resolution through the integration of Earth observation and an advanced modelling system. The DTE Hydrology Prototype has been used for water resources management and for identifying locations and times of risk for landslides and flooding in the Po River Basin, in northern Italy.”



## **Forest DTEP**

Project aims to prototype a digital twin of forests using the existing scientific knowledge and capabilities.

The project consortium will work with Earth observation data to determine the structural and biological properties of forests.

The consortium will use this information to initialize a set of forest ecosystem models that include processes above and below the ground.

Further, the consortium will develop new interfaces and visualization tools, making the cloud platform easily accessible to end users. Goal of the system developed in the project is to bring together forest data and users interested in forests.

# Digital Twin Earth

Dynamic interactive replica of past, present and future of our planet in the digital domain based on effective integration of observations (satellite, in-situ, IoT and socioeconomic data), Earth system science and simulations to bridge impact sectors science and simulations on the base of AI methodologies.

Outcomes

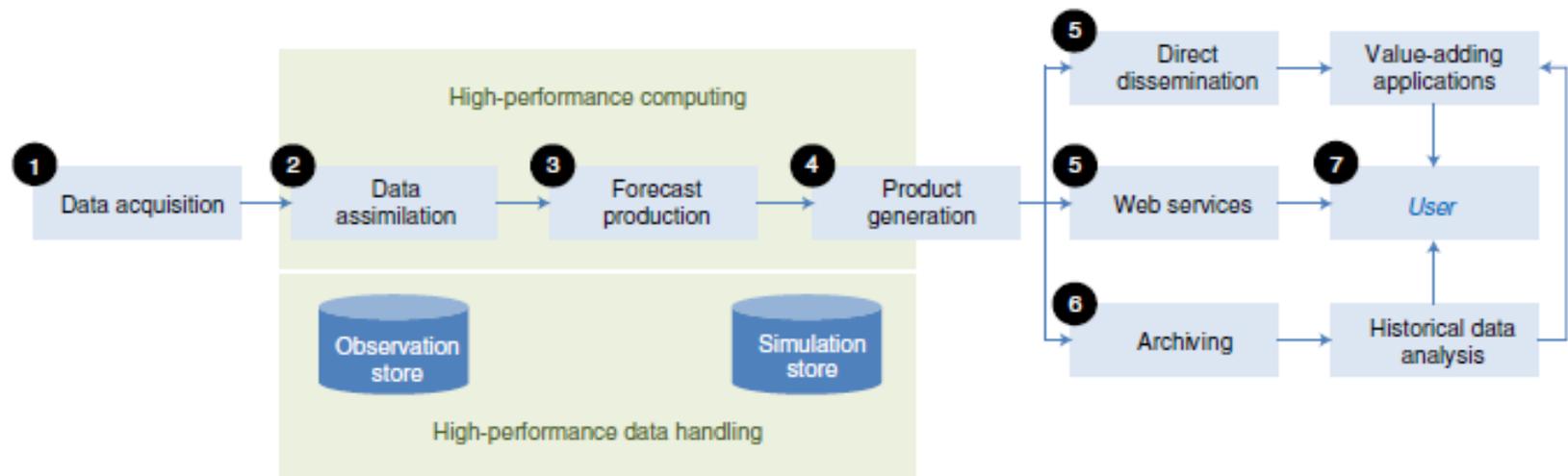
- 1. Enhanced understanding of complex processes involving different components of the Earth system** across different scientific domains.
- 2. Enhanced the scientific understanding and characterisation of the role of human activities as an integral part of the Earth system.**
- 3. Enhanced understanding and characterisation of multi-scale connections** of complex processes from local scales to regional and global scale dynamics.



# The digital revolution of Earth-system science

Peter Bauer <sup>1</sup>✉, Peter D. Dueben<sup>1</sup>, Torsten Hoefler<sup>2</sup>, Tiago Quintino <sup>3</sup>, Thomas C. Schulthess<sup>4</sup> and Nils P. Wedi<sup>1</sup>

Computational science is crucial for delivering reliable weather and climate predictions. However, despite decades of high-performance computing experience, there is serious concern about the sustainability of this application in the post-Moore/Dennard era. Here, we discuss the present limitations in the field and propose the design of a novel infrastructure that is scalable and more adaptable to future, yet unknown computing architectures.



**Fig. 1 | Typical production workflow in operational numerical weather prediction.** (1) High-volume and high-speed observational data acquisition and pre-processing; (2) data assimilation into models to produce initial conditions for forecasts; (3) forecast production by Earth-system simulation models; (4) generation of output products tailored to the portfolio of weather and climate information users; (5) direct dissemination of raw output and web-products; (6) long-term archiving for reuse in statistical analyses and performance diagnostics; (7) user-specific applications and data-driven analytics.

Microsoft builds the **Planetary Computer** which will help to understand where and how we can best mitigate the impact of humans on the environment, as well as to support decision makers for climate adaptation measures.

The Planetary Computer consist of four major components that also can facilitate to build Environmental **Digital Twins**:

- Planetary Computer Data Catalog — to search and find open datasets
- Planetary Computer API's — to access data and retrieve data
- Planetary Computer Hub — for data processing
- Planetary Computer Applications — third party open and closed applications build on Planetary Computer infrastructure

Organizations can build their Environmental Digital Twins partly based upon the Planetary Computer.

The [Earth Archive Initiative](#) is an unprecedented scientific effort to create a digital twin of the entire surface of the Earth – and everything on it. By scanning the planet’s land surface with very high-resolution lidar, the Earth Archive will create a true three-dimensional digital twin of our world — an open source, digital record of the Earth that will reflect the landscape exactly as it was at the time of scanning. The geospatial data captured will serve as the baseline for understanding and exploring our world.

[Home](#)[The Congress](#)[Networking](#)[Support EA](#)[Sponsors](#)[About](#)

# The Earth Archive Virtual Congress

An unprecedented scientific effort to create a digital twin of the entire surface of the Earth and everything on it before it's too late.

Chapter I : The Amazon  
June 15th - 16th 2021

- ✓ Home
- ✓ Technologies
- ✓ Sectors
- ✓ COVID-19
- ✓ AI/ML/DL
- ✓ Exascale
- ✓ Specials
- ✓ Resource Library
- ✓ Podcast
- ✓ Events
- ✓ Job Bank
- ✓ About
- ✓ Our Authors
- ✓ Solution Channels
- ✓ Subscribe



## Nvidia's Digital Twin Moonshot – Saving Planet Earth

By Todd R. Weiss

November 11, 2021

Nvidia is continuing to expand enterprise digital twin capabilities and features in its Nvidia Omniverse 3D virtual world design platform, but the company's nascent plans for its grandest digital twin yet might have even bigger impacts after its announcement at this week's Nvidia GTC21 virtual conference.

The GPU-maker wants to accelerate its digital twin dreams from helping enterprises model products, factories, assembly lines and more – and rev it up to create the grandest digital twin so far – of the planet Earth itself that will continuously model, predict and track climate change in real-time so scientists can seek ways to reverse or stop the destructive effects of its spread.

“We will build a digital twin to simulate and predict climate change,” said Nvidia CEO Jensen Huang [during his Nov. 9 keynote](#) at the virtual event. The work will be done using a powerful new supercomputer that the company is now building, called Earth-2, which will run AI physics created by the new [Nvidia Modulus AI framework](#) at million-X speeds on the [Nvidia Omniverse platform](#), he said.

“All the technologies we have invented up to this moment are needed to make Earth-2 possible,” said Huang. “I cannot imagine a greater and more important



### Leading Solution Providers



[nature](#) > [nature climate change](#) > [comment](#) > article

Comment | [Published: 01 February 2021](#)

## A digital twin of Earth for the green transition

[Peter Bauer](#) , [Bjorn Stevens](#) & [Wilco Hazeleger](#)

[Nature Climate Change](#) **11**, 80–83 (2021) | [Cite this article](#)

**5032** Accesses | **23** Citations | **260** Altmetric | [Metrics](#)

**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.**

A key element of the Green Deal is its dependence on the ‘digital transformation’ – an openly accessible and interoperable European dataspace as a central hub for informed decision making. The EU identified two landmark actions to support the necessary information systems: GreenData4All<sup>2</sup> and Destination Earth<sup>3</sup>. Whereas GreenData4All will develop the European approach to discover, manage and exploit geospatial information, Destination Earth aims to construct highly accurate models, or ‘digital twins’, of the Earth

# EU Megaproject Destination earth (Destine)

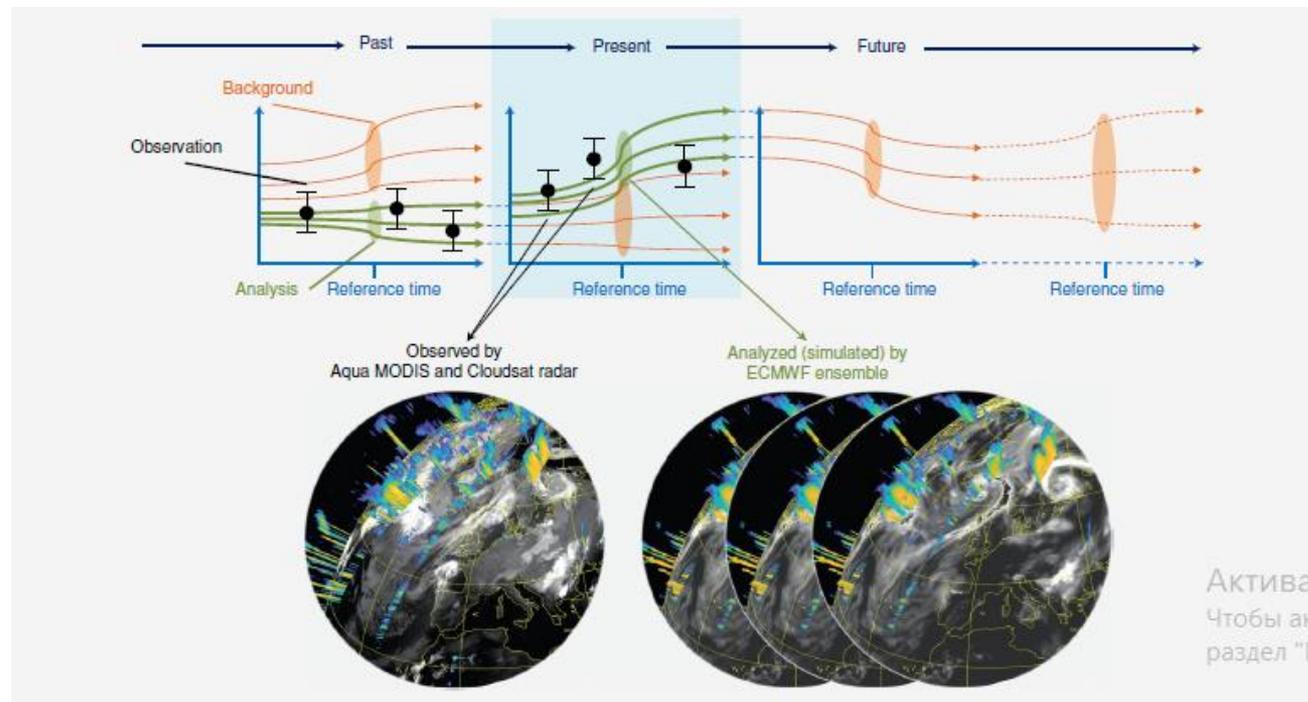
“The objective of the Destination Earth initiative is to develop a very high precision digital model of the Earth to monitor and simulate natural and human activity, and to develop and test scenarios that would enable more sustainable development and support European environmental policies.

At the heart of Destination Earth will be a federated cloud-based modelling and simulation platform, providing access to data, advanced computing infrastructure (including high performance computing), software, AI applications and analytics. The digital twins created in DestinE will give users access to high-quality information, services, models, scenarios, forecasts and visualizations. Digital twins are based on the integration of continuous observation, modelling and high-performance simulation, resulting in highly accurate predictions of future developments.”\*



Implementation: 2021-2030

Optimally combined simulations and near-real-time observations to monitor the evolution of the Earth system. For each cycle, the simulation generates a background forecast ensemble (orange arrows) of the Earth system, which is compared to observations (black dots) throughout a time window and eventually corrected to an analysis ensemble (green arrows), which fits the observations better than the background. Uncertainties of forecasts (ellipses) and observations (error bars) are fully taken into account from ensembles, which are multiple, perturbed realizations of both model and observations.



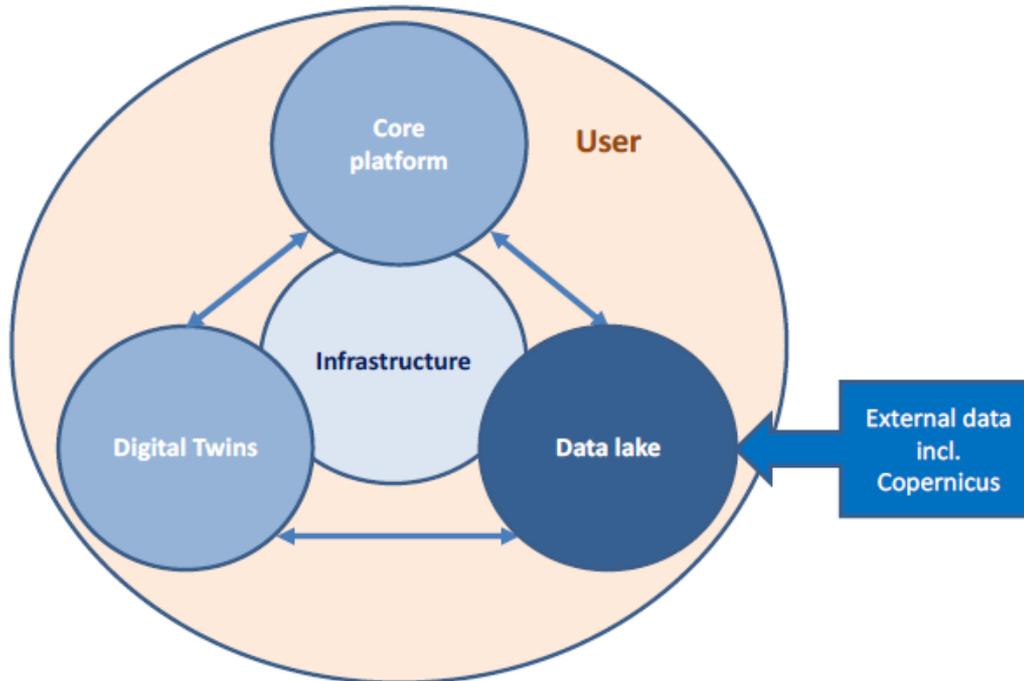
# Digital Twins in DestinE

*Peter Bauer*

*ECMWF*

## Digital Twins

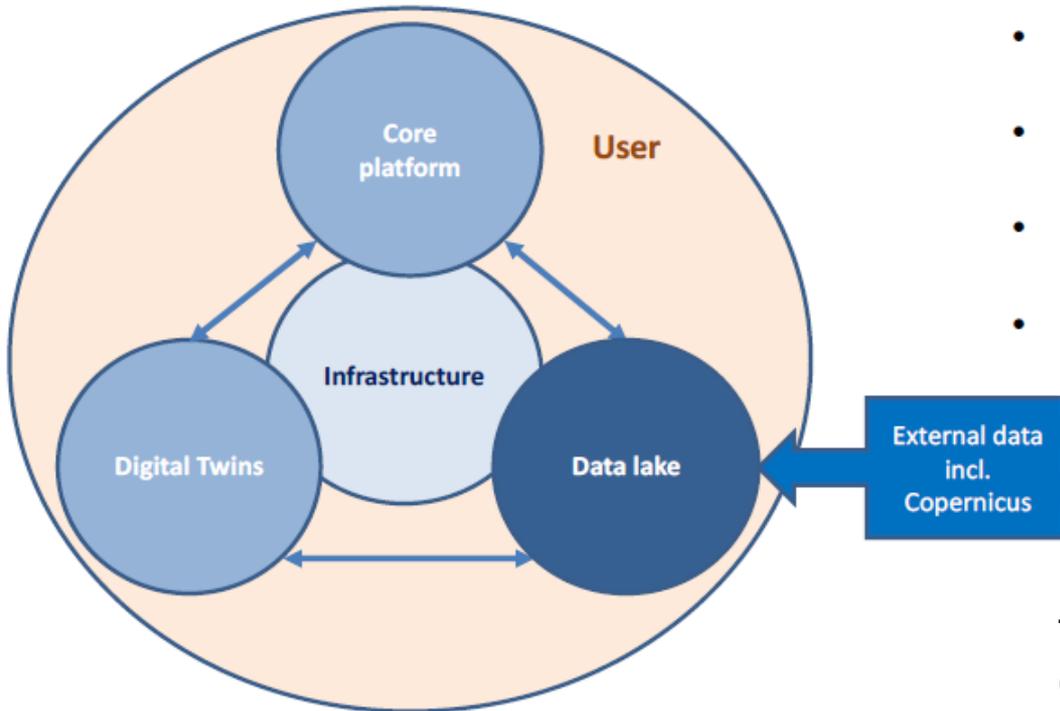
---



### 2 high-priority Digital Twins:

- 1. Weather-induced and Geophysical Extremes:**  
Environmental extremes at very high spatial resolution and close to real-time decision-making support at continental, country, coastline, catchment and city scales in response to meteorological, hydrological and air quality extremes.
- 2. Climate Change Adaptation:**  
Climate change adaptation policies and mitigation scenario testing at decadal timescales aiming at a real breakthrough at the level of reliability at regional and national levels, for understanding the causes and explaining the feedback mechanisms of change, and predicting possible evolution trajectories

## Core Platform



- Portal to Users
- User Hosting (user space: MyDestinE)
- Catalogue and data access services
- PaaS & SaaS for analytics and modelling
- Community management (market place, annotations, open source)
- Advanced interface for interactive modelling (system dynamic, 3D,...)
- Operate the User Data API for Data Lake access and User DTs interface for on-demand production
- **PaaS** (Platform as a Service) — модель предоставления облачных вычислений, при которой потребитель получает доступ к использованию информационно-технологических платформ: ОС, СУБД, ПО, и т.д.

**SaaS** (*software as a service* — одна из форм облачных вычислений, модель обслуживания, при которой подписчикам предоставляется готовое прикладное программное обеспечение, полностью обслуживаемое провайдером.

# Role of users!

## Core Platform

---

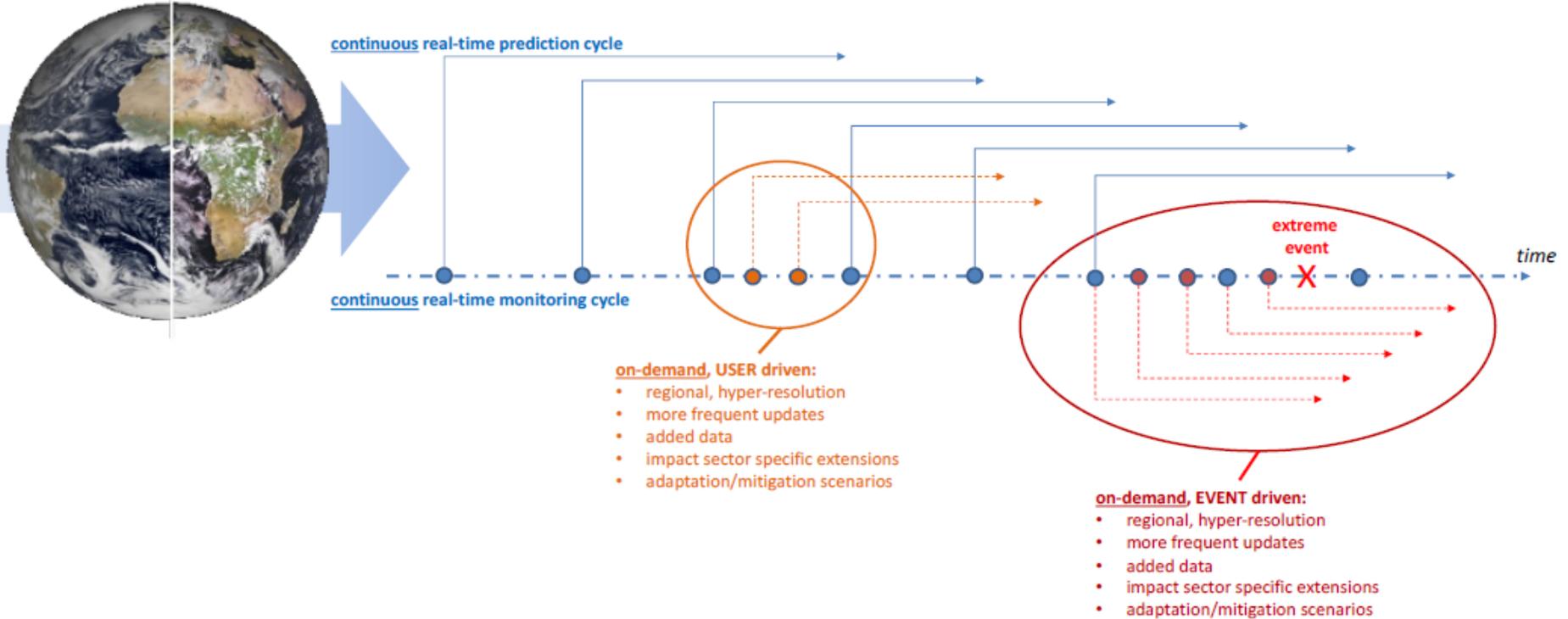
**Users make use of the DestinE Core Service Platform (DESP) to produce answers to their problems\***

Platform combines generic & user specific services ...      ... in a user customizable environment to :

- Unified access to data generated or collected/referenced by DestinE
- Applications to support development / modelling / analysis / visualization
- On-demand DT simulation / data retrieval / data transformation
- Develop & operate own applications / services
- Share results / data / applications / libraries

# Digital Twin production

simulations observations



## Current projects in Russia:

“Digital field” project (I.V. Oseledets, Scoltech)

Federal project “Digital Ob’-Irtysh catchment” (DT of river catchment - V.P. Potapov, FRC for Information and computational Technologies)

## **Possible activities**

DTs of Northern Eurasia ecosystems for (bogs, forest, permafrost, etc.) as an instrument to “green” economic transition (should be used in Carbon polygons)!



**Thanks for attention!**