NASA ESTO
Advanced Information Systems Technology (AIST)

ESTO
Earth Science Technology Office
AIST

Information Systems Technology for NASA Earth Systems Digital Twins (ESDT)

Jacqueline Le Moigne

June 28, 2023

Why Digital Twins in Earth Science?

ESTO
Earth Science Technology Office
AIST

- **Well documented data covering the entire Earth have now been collected continuously for more than 50 years**, not only from space but also from airplanes, balloons and in-situ sensors. With the addition of commercial remote sensing providers and many more Internet-of-Things sensors, these incredible amounts of data will soon be augmented by even larger amounts of diverse data and therefore will become more and more difficult to access, understand and utilize.

- At the same time, because of climate change and its impacts, the information produced by all of this data is becoming of **interest to many new non-traditional users** for analyzing and predicting various phenomena. As a consequence, the information derived from all of the data described above will need to be accessed and analyzed by multiple and diverse users for various uses and applications.

- Because of **advances in computational and visualization capabilities** and the parallel unprecedented development of **Artificial Intelligence technologies, especially Machine Learning (ML),** extracting relevant information from these large amounts of data and running complex models faster has become possible.

Digital Twins in Earth Science

ESTO
Earth Science Technology Office
AIST

- **2020/Loekken, Le Saux & Aparicio-ESA**
  A dynamic interactive replica of the past, present and future of our planet in the digital domain based on an effective integration of observations (satellite, in-situ, IoT and socioeconomic data), Earth-system science and simulations, the bridge to impact sectors science and simulations and artificial intelligence methodologies
- **2020/Bauer-ESA**
  A digital twin of Earth is an information system that exposes users to a digital replication of the state and temporal evolution of the Earth system constrained by available observations and the laws of physics.
- **2021/ECMWF-DestE Position**
  A digital twin is a digital replica of a living or non-living physical entity. The digital twins created in DestinE will give expert and non-expert users tailored access to high-quality information, services, models, scenarios, forecasts and visualisations. This includes models of the climate, weather forecasting, hurricane evolution and more. Digital twins rely on the integration of continuous observation, modelling and high-performance simulation, resulting in highly accurate predictions of future developments.
- **2022/**https://www.esa.int/Applications/Observing_the_Earth/ESA_moves_forward_with_Destination_Earth
  Constantly fed with Earth observation data, combined with *in situ* measurements and artificial intelligence, the digital twins provide us with highly accurate representation of the past, present and future changes of our world.

What is an Earth System Digital Twin (ESDT)?

ESTO
Earth Science Technology Office
AIST

An Earth System Digital Twin or ESDT is a dynamic and interactive information system that first provides a digital replica of the past and current states of the Earth or Earth system, as accurately and timely as possible, second provides the evolution of these systems by computing forecasts of future states under nominal assumptions and based on the current replica, and third offers the capability to investigate many hypothetical evolution scenarios under varying impact assumptions.

Build around relevant Earth system models and simulations and other relevant models (e.g., related to the world's infrastructure), the main characteristics of an Earth System Digital Twin are:

1. Two-way integration of continuous and timely (including near real time and direct readout) observations
2. Heavy use of Machine Learning in various ESDT components, e.g., allows for fast simulations and improve spatial and temporal resolutions
3. Various levels of users, ”from farmer to scientist and decision maker”

<u>**Key Attributes:**</u>

- Interactive, integrated, multidomain and multiscale
- Data and information from long-time records space, air, ground, over/underwater, Internet of Things (IoT), socioeconomic, etc.
- “Twin” as far as resolutions and uncertainties allow (similarly to traditional models)

A Few ESDT Science Use Cases/Scenarios

ESTO
Earth Science Technology Office
AIST

ESDT Use Case	SCOPE
Wildfires	A digital twin of Earth systems involved in wildfires to represent and understand the origins and evolution of wildfires and their impacts on ecosystem, infrastructure, and related human systems.
Ocean Carbon	An Earth system digital twin of ocean, land, atmospheric Earth systems to understand ocean carbon processes such as carbon export and ocean-atmosphere processes and coupling; land-ocean continuum and interactions with human systems; coastal ecological changes and impacts to ecosystem services; feedback processes (e.g., storm intensification and sea level rise) and impacts on coastal communities and the blue economy; assessing feasibility and impacts of various Carbon Dioxide Removal (CDR) approaches as a strategy to remove and sequester atmospheric carbon.
Water Cycle	A local or regional digital twin to understand all the complexities of the Water Cycle, how it is affected by various Earth Systems at multiple temporal and spatial scales, and how it is impacted by decision making and human influence. It would provide capabilities *such as* zooming out in time and space; helping understand water availability and origin for agriculture; how events such as floods and droughts affects life, property and infrastructure; and more generally how the effects of weather and climate variability can be mitigated under various scenarios.
Central Africa Carbon Corridors	An Earth System digital twin of “Carbon Corridors” (i.e., connected regions of protected forests/vegetation. They store carbon and maintain habitat connectivity for biodiversity) in Central Africa to: understand the current conditions; assess their ability to store carbon and promote biodiversity; forecast future conditions; conduct what-if scenarios to assess the impact of policy decisions and potential climate conditions.
Atmospheric Boundary Layer	An Earth system digital twin of the atmospheric boundary layer to provide a digital replica of the lowest portions of the atmosphere and of their processes and interactions with other systems – land, ocean, and ice surfaces – and how these interactions control exchanges with materials such as trace gases, aerosols; coupled atmospheric systems to understand underlying processes and their relationship to climate and air quality, the role of these interactions on the global weather and climate system; atmospheric systems related to greenhouse gasses (GHG), sources of pollution, and their transport in the atmosphere to understand air quality and human health impacts at multiple scales from hyper local to long term global climate projections; proper characterization of the Planetary Boundary Layer (PBL) is also critically important for modeling nighttime minimum temperatures for agricultural applications, and for prediction of wildland fire risk.
Coastal Zone Digital Twin	An Earth System digital twin of local and regional coastal zones that considers both natural and human systems to understand changes in coastal flooding severity, land and marine morphology, nutrients and water quality, ecological makeup, sea level, and the short and long-term risks to climate change adaptation, sustainable development, disaster management, tourism and recreation, quality of life, ecosystem management, and coastal infrastructure management.

Earth System Digital Twins Components

. . .

Digital Replica . . . What now?

An integrated picture of the past and current states of Earth systems.

Forecasting . . . What next?

An integrated picture of how Earth systems will evolve in the future from the current state.

Impact Assessment . . . What if?

An integrated picture of how Earth systems could evolve under different hypothetical what-if scenarios.

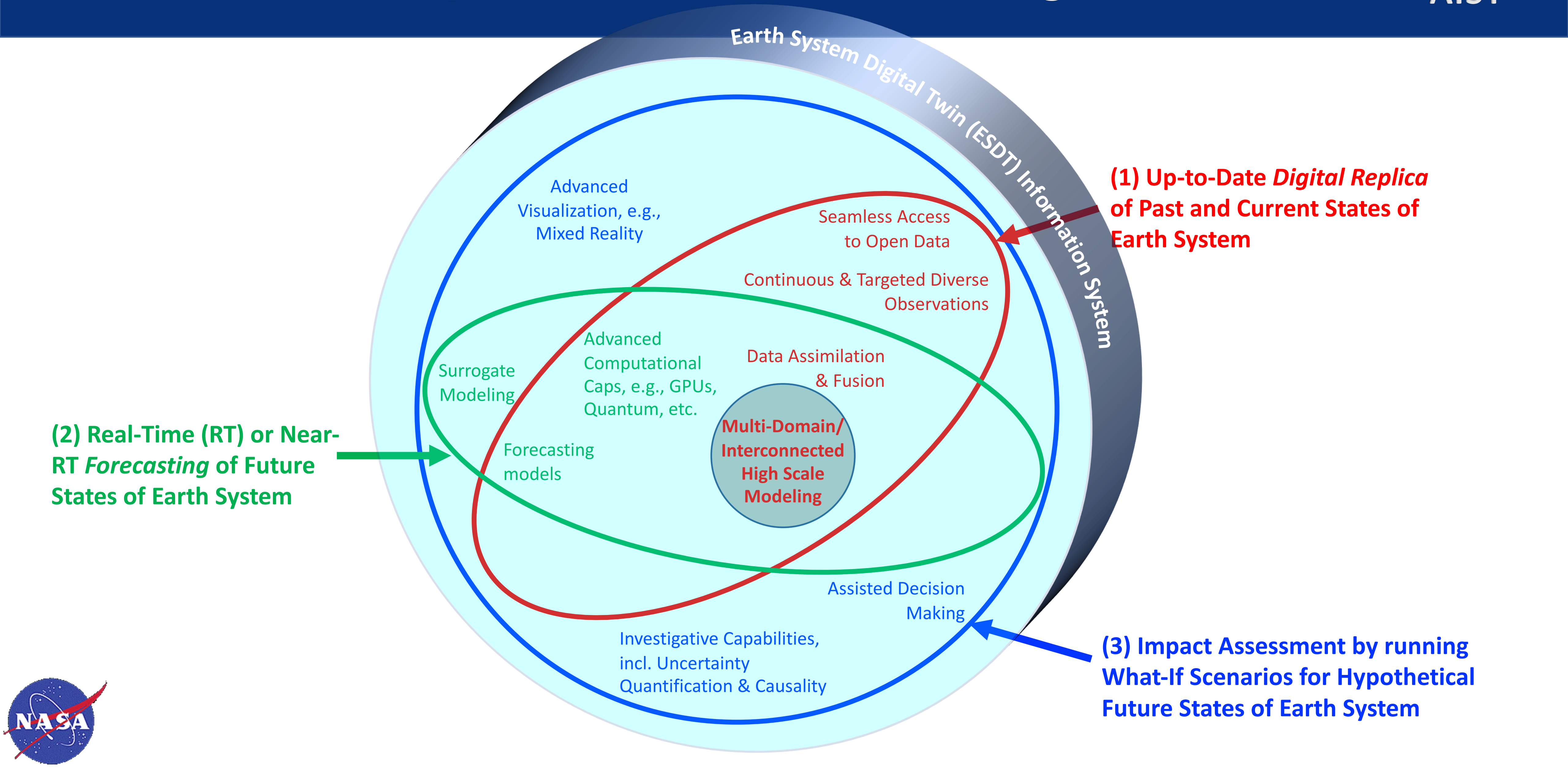
- **Continuous observations** of interacting Earth systems and human systems
- From many **disparate sources**
- Driving **inter-connected models**
- At many **physical and temporal scales**
- With fast, powerful and integrated **prediction, analysis** and **visualization** capabilities
- Using **Machine Learning, causality and uncertainty quantification**
- Running at **scale** in order to improve our **science** understanding of those systems, their **interactions and** their **applications**

General ESDT Challenges

- Local vs. Regional vs. Global vs. Thematic Digital Twins?
- Future “web” of Digital Twins co-existing in a hierarchy or in a network, and capable of being connected or federated depending on the needs. How will we federate future ESDT?

  ➔ Interoperability/Standards and Protocols: Syntactic, semantic, legal and organizational levels
- Which sustainable digital twin governance model should be adopted to address software configuration changes, security and full life cycle management?
- hat are the main architecture components of an ESDT? What could various ESDT architectures look like?
- How to build a Digital Replica? Which data, Analysis ready Data, Information should be included? Datacubes, data lakes, indexing, on demand, etc.
- How will various data, models, ESDT interoperate? Which basic interfaces and standards will be required?
- How do we trust/validate ESDT (e.g., using historical data, etc.)? How to quantify uncertainty?
- Which computational resources will be required? Cloud, GPU’s, Quantum, Neuromorphic, etc.?
- How to visualize a Digital Twin? User Interface? Interactivity? Refresh Rate?
- Which sustainable digital twin governance model should be adopted to address software configuration changes, security and full life cycle management?

AIST ESDT Capabilities and Technologies

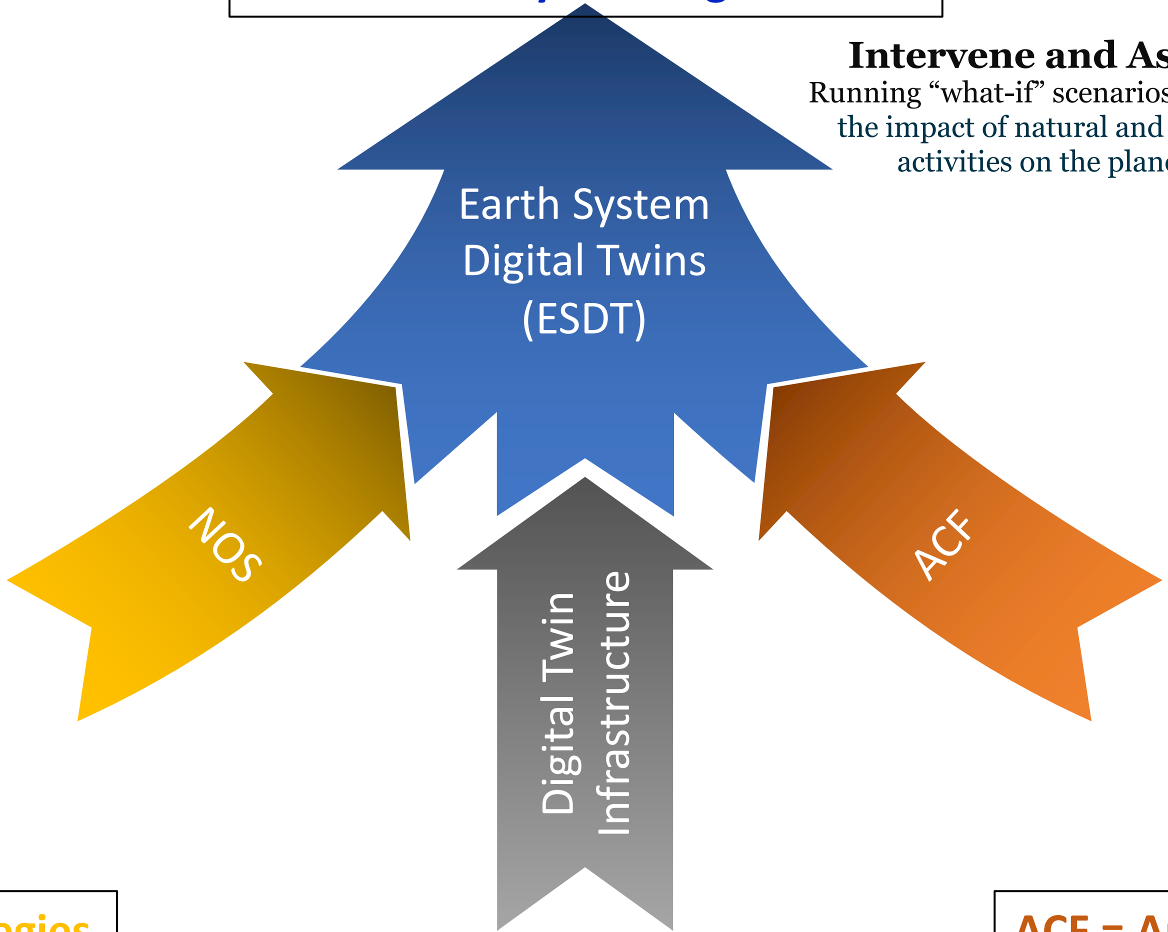


Three Advanced Information Systems Thrusts

ESTO Earth Science Technology Office AIST

ESDT = Earth System Digital Twins

Intervene and Assess
Running "what-if" scenarios to assess the impact of natural and human activities on the planet.



Observe, Target and Coordinate
Edge and on-the-ground intelligent planning, evaluating, coordinating and operating collections of diverse and distributed observing assets

Fuse, Analyze, Share and Collaborate
Simplify access to diverse and large amounts of data, analytics & modeling tools and advanced computational resources for collaborative science

NOS = Novel Observing Strategies

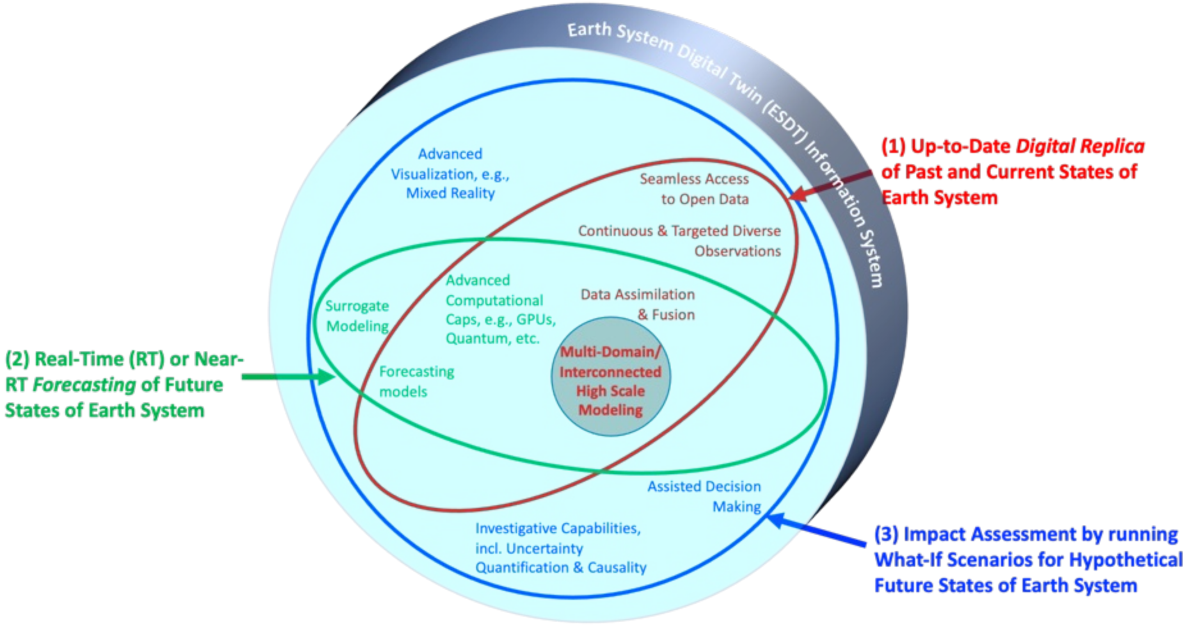
ACF = Analytic Collaborative Frameworks

Interrogate, Simulate, Trade and Visualize
Robust tools for interrogating, assessing uncertainties & causality, and for visualization, leveraging diverse data, models and products

Machine Learning for Earth System Digital Twins

ESTO Earth Science Technology Office
AIST

- Improving data content and information extraction
- Improving data fusion and data assimilation
- Improving models spatial and temporal accuracy
- Enabling model interconnection
- Accurate and trusted surrogate modeling
- Providing full explainability
- Integrating or fully relating to physics models
- Speeding up What-If simulations
- Enabling causal analysis and impact assessment
- Enabling straightforward, dynamic, and interactive user interfaces



ESDT Technologies Requested in AIST-21

ESTO Earth Science Technology Office AIST

- **AIST-21 Solicitation, first US government Solicitation requesting Digital Twins Technology for Earth Science:**
  - **New ESDT thrust** building on and advancing previously AIST-funded technology for Analytic Frameworks, Machine Learning (ML), and Seamless Integration of Multi-Source and Timely Observations
  - **14 AIST-funded ESDT projects (2020 – 2023)** focusing on developing:
    - Underlying analytic capabilities and frameworks to build Digital Replicas
    - Novel ESDT infrastructure technologies (containers, VR, uncertainty quantification, etc.)
    - Surrogate modeling and ML emulators
    - Preliminary prototypes including interconnected modeling

- **AIST-CNES Collaborative Development of Flood ESDT Prototype**

  => Federated System IDEAS (NASA) – FloodDAM DT (CNES)

AIST-21 ACF for ESDT Awards

ESTO
Earth Science Technology Office
AIST

- **ACF Towards ESDT**

PI's Name	Organization	Title	Synopsis
Thomas Allen	Old Dominion University	Pixels for Public Health: Analytic Collaborative Framework to Enhance Coastal Resiliency of Vulnerable Populations in Hampton Roads, Virginia (VA)	Proposes to design and operationally demonstrate a system linking the VA Open Data Cube, a socio-spatial-health information “Digital Neighborhood” (Hampton Roads Biomedical Research Cons.), hydrodynamic models, and in-situ flood sensor network. Will connect observational and physical environmental domains with human vulnerability.
Arlindo Da Silva	NASA Goddard Space Flight Center (GSFC)	An Analytic Collaborative Framework for the Earth System Observatory (ESO) Designated Observables	Will develop an Analytic Collaborative Framework for the Earth System Observatory (ESO) missions, based on realistic, science-based observing system simulations and the Program of Record (PoR), tied together in a cloud-based cyberinfrastructure. Create a 3D, holistic view of Earth with all ESO unique satellites.
Thomas Huang	NASA Jet Propulsion Laboratory (JPL)	Fire Alarm: Science Data Platform for Wildfire and Air Quality	Proposes to advance AIST's Air Quality Analytics Collaborative Framework (AQACF) to establish a wildfire and air quality ACF, Fire Alarm, focusing on the prediction and analysis of wildfire, burned area and the air quality as an integrated platform to guide decision-makers, science researchers, and first-responders.

AIST-21 ESDT Awards

AIST

- **ESDT Infrastructure**

PI's Name	Organization	Title	Synopsis
Thomas Clune	NASA Goddard Space Flight Center (GSFC)	A Framework for Global Cloud Resolving OSSEs	Will enable global, cloud-resolving Observing System Simulation Experiments (OSSEs) by addressing key computational challenges to enable existing technologies to scale to the spatial resolutions needed by the end of decade, e.g., extending parallel I/O capabilities, adopting a 2-phase Nature Run approach and a flexible API for customization.
Thomas Grubb	NASA Goddard Space Flight Center (GSFC)	Goddard Earth Observing System (GEOS) Visualization And Lagrangian dynamics Immersive eXtended Reality Tool (VALIXR) for Scientific Discovery	Proposes to develop a scientific exploration and analysis mixed augmented and virtual reality tool with integrated Lagrangian Dynamics (LD) to help scientists identify, track, and understand the evolution of Earth Science phenomena in the NASA GEOS model. It will provide both a scientific discovery tool and a model analysis and improvement tool.
Matthias Katzfuss	Texas A&M University (TAMU)	A scalable probabilistic emulation and uncertainty quantification tool for Earth-system models	Proposes to develop a fully automated toolbox for uncertainty quantification in Earth-system models, to provide insight into the largest and most critical information gaps and identify where potential future observations would be most valuable. It would allow interpolation between observed covariate values and running extensive what-if scenarios.
Tanu Malik	De Paul University	Reproducible Containers for Advancing Process-oriented Collaborative Analytics	Aims to establish reproducible scientific containers that are easy-to-use and are lightweight. Reproducible containers will transparently encapsulate complex, data-intensive, process-oriented model analytics, will be easy and efficient to share between collaborators, and will enable reproducibility in heterogeneous environments.

AIST-21 ESDT Awards (cont.)

ESTO Earth Science Technology Office AIST

- **AI-Surrogate Modeling for ESDT**

PI's Name	Organization	Title	Synopsis
Allison Gray	Univ. of Washington, Seattle	A prototype Digital Twin of Air-Sea Interactions	Proposes to develop hybrid physics-informed AI model that ingests several existing flux estimates and observation data products and train against simultaneous ocean-atmosphere data from Saildrones. This will ascertain uncertainty of existing flux measurements and optimize combination of near-real-time existing flux data and observational data => This represents the first step towards a Digital Twin for the Planetary Boundary Layer.
Christopher Keller	Morgan State University (MSU)	Development of a next-generation ensemble prediction system for atmospheric composition	Proposes to develop a next-generation modeling framework for the real-time simulation of reactive gases and aerosols in the atmosphere. Will deploy computationally efficient parameterizations of atmospheric chemistry and transport and will develop generative models based on machine learning (ML) to predict model uncertainties.
Jouni Susiluoto	NASA Jet Propulsion Laboratory (JPL)	Kernel Flows: emulating complex models for massive data sets	Proposes a general-purpose, versatile emulation tool to provide fast, accurate emulation with little tuning, to scale up to very large training sets, and to provide uncertainties associated with outputs. This tool set will facilitate large-scale implementation of forward modeling and retrievals, and of UQ at production scales. To be applied to SBG radiative transfer emulation & convective storm nowcasting.

NASA

AIST-21 ESDT Awards (cont.)

ESTO Earth Science Technology Office AIST

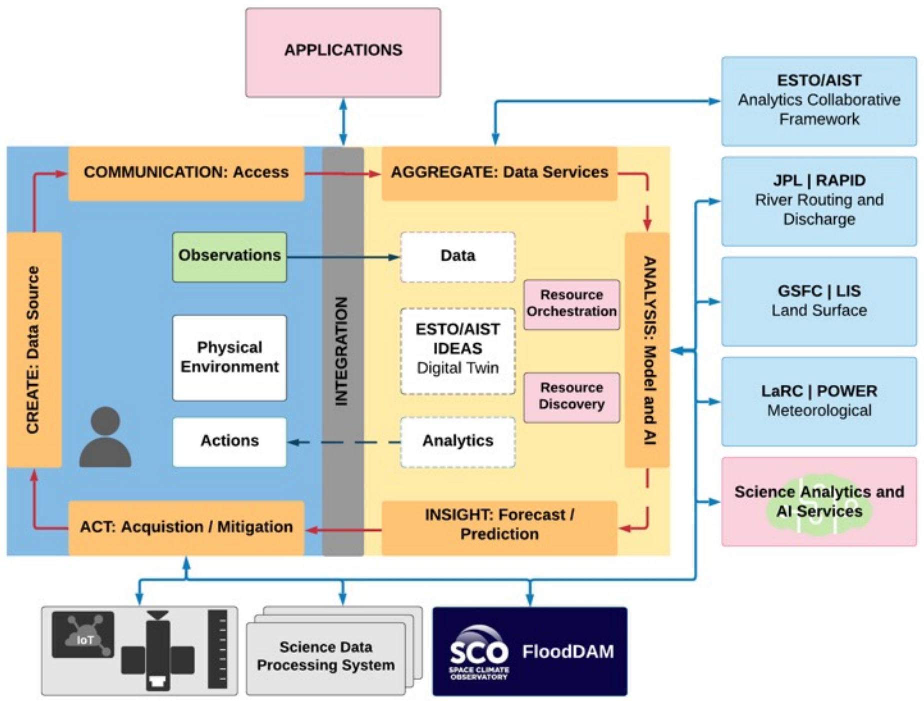
- **ESDT Prototypes**

PI's Name	Organization	Title	Synopsis
Rajat Bindlish	NASA Goddard Space Flight Center (GSFC)	Digital Twin Infrastructure Model for Agricultural Applications	Will develop an agriculture productivity modeling system over Continental United States as an example of incorporating representations of infrastructure-oriented process, for the understanding, prediction, and mitigation/response of Earth system process variability, with application to crop growth, yield, and agricultural production information, critical to commodity market, food security, economic stability, and government policy formulation.
Milton Halem	*University of Maryland, Baltimore County (UMBC)*	*Towards a NU-WRF based Mega Wildfire Digital Twin: Smoke Transport Impact Scenarios on Air Quality, Cardiopulmonary Disease and Regional Deforestation*	*Will develop and implement a Regional Wildfire Digital Twin (WDT) model with a sub-km resolution to enable the conduct of mega wildfire smoke impact scenarios at various spatial scales and arbitrary locations over N. America. WDT will provide a valuable planning tool for impact scenarios by season, location, intensity, and atmospheric state.*
Craig Pelissier	Science Systems and Applications, Inc. (SSAI)	Terrestrial Environmental Rapid-Replicating Assimilation Hydrometeorology (TERRAHydro) System: A machine-learning coupled water, energy, and vegetation terrestrial Earth System Digital Twin	Proposes to develop a terrestrial Earth System Digital Twin (TESDT) that couples state-of-the-art ML with NASA (and other) EO data. It will combine the best ML hydrology models with capabilities for uncertainty quantification and data assimilation to provide ensemble & probabilistic forecasting, sensitivity analyses, and counterfactual “what if" experiments.

QRS-21/Huang (NASA JPL, GSFC, LaRC with CNES) –
ESTO/AIST Integrated Digital Earth Analysis System (IDEAS)

ESTO
Earth Science Technology Office
AIST

***IDEAS is a NASA ESTO/AIST Earth System Digital Twin** project that bridges the physical environment and its virtual representation by continuously assimilating new observations to improve forecast and prediction for integrated science and decision support.*



IDEAS – Digital Twin for Water Cycle and Flood Detection and Monitoring

Jet Propulsion Laboratory
California Institute of Technology
Goddard Space Flight Center
Langley Research Center

- Using water cycle and flood analysis as the prototype application to integrate NASA, CNES, and Space Climate Observatory (SCO) data and science
- Multi-Agency and Multi-Center partnership
- Advanced numerical models and analysis
  - **JPL's RAPID**: Routing Application for Parallel computation of Discharge
  - **GSFC's LIS**: Land Information System
  - **LaRC's POWER**: Prediction of Worldwide Energy Resources
  - **CNES and SCO's FloodDAM:** Automated service to reliably detect, monitor and assess flood events globally
  - Integration with **NASA IPCC Sea Level Prediction** data for coastal flooding
  - Joint developed and trained flood detection and prediction machine learning algorithms
  - Promote and advance interoperable standards
- Improve the Machine Learning flood prediction model (Huang *et al.* 2020) from the JPL-CNES Joint Data Science pilot
- Scenario-based prediction for infrastructure and population impacts
- Ongoing formulation and planning with CNES and SCO's FloodDAM
- Next interchange will be on October 6, 2021 and project is expected to start soon after

What's Next for ESDT?

ESTO
Earth Science Technology Office
AIST

Additional Use Cases – Related to "Earth Science to Action" Strategy

Overall Capabilities:

- (Semi-)Autonomy – Two-Way Information Exchange between Observing Systems and Model Interactions
- Data Mining and Discovery
- Fusion of Disparate including Unstructured Data Sources, esp. from Human and Social Systems
- Causality Assessment between diverse Earth systems and with Human Activity
- Validation and Uncertainty Quantification of AI Systems and ML Emulators
- Explainability of AI Systems/True Integration of Physics Models and ML Models => Build trust in ML and ESDT
- Real- or Near-RealTime Visualization and interactivity capabilities are needed for data analysis and exploration

End-to-End Earth System Digital Twins:

- Digital Twin Architecture(s) and Interoperability/Federation
- Integration of NOS capabilities for continuous ingest of real-time and/or timely data and information into ESDT
- Design of ACF System as ESDT "Core Engine"
- User-friendly interfaces and visualizations with dynamic and interactive capabilities

ESDT Workshop
October 26-28, 2022

ESTO
Earth Science Technology Office
AIST

Available mid-July 2023 on the AIST Website:
https://esto.nasa.gov/aist/

Advanced Information Systems Technology (AIST)
Earth Systems Digital Twin (ESDT)
Workshop Report

Jacqueline Le Moigne - NASA Earth Science Technology Office
Benjamin Smith – NASA Earth Science Technology Office

DRAFT

Workshop Co-Organized with Earth Science Information Partners (ESIP)
Report Edited by ESDT Workshop Participants

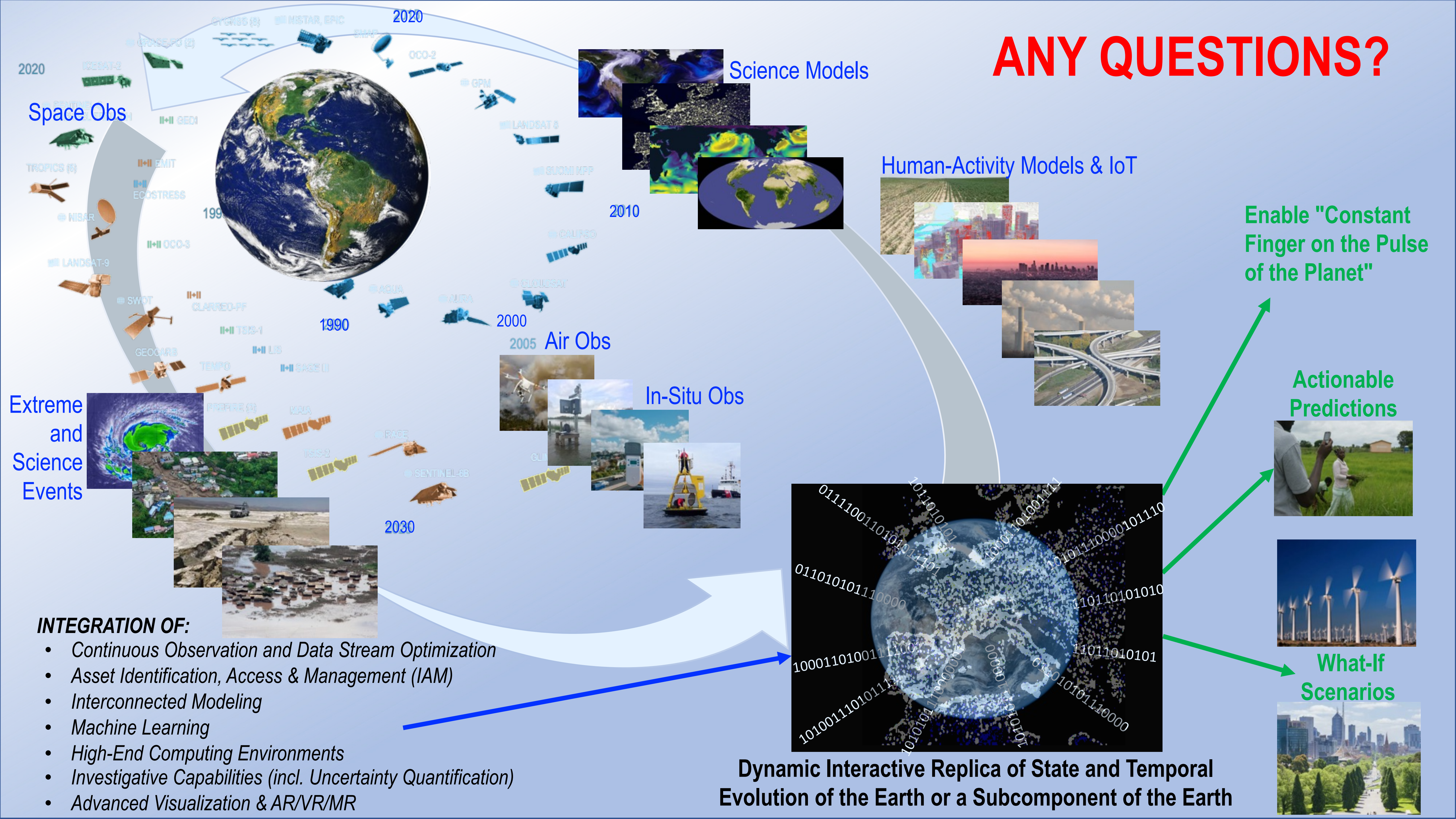
October 26-28, 2022
Washington DC

ESTO
EARTH SCIENCE TECHNOLOGY OFFICE

Projects Final Reports and Quads
Available on the ESTO Portfolio Webpage

http://www.estotechnology.us/techportfolio/

Choose ESTO Projects (left) and “All” (right)
Then Choose “Information Systems”



ANY QUESTIONS?
Space Obs
Science Models
Human-Activity Models & IoT
Air Obs
In-Situ Obs
Extreme and Science Events
2020
1990
2000
2005
2010
2030
Enable "Constant Finger on the Pulse of the Planet"
Actionable Predictions
What-If Scenarios
INTEGRATION OF:
• Continuous Observation and Data Stream Optimization
• Asset Identification, Access & Management (IAM)
• Interconnected Modeling
• Machine Learning
• High-End Computing Environments
• Investigative Capabilities (incl. Uncertainty Quantification)
• Advanced Visualization & AR/VR/MR
Dynamic Interactive Replica of State and Temporal Evolution of the Earth or a Subcomponent of the Earth

Earth Science Technology Office