NASA ESTO
Advanced Information Systems Technology (AIST)

NASA Earth Systems Digital Twins (ESDT)

Jacqueline Le Moigne

AMS'24, February 2024

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.

What is an Earth System Digital Twin (ESDT)?

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 allows for *computing forecasts of future states* under nominal assumptions and based on the current replica, and third offers the *capability to investigate many hypothetical scenarios* under varying impact assumptions.

It is an interactive and integrated multidomain, multiscale, digital replica of the state and temporal evolution of Earth systems that dynamically integrates:

- Relevant Earth system models and simulations
- Other relevant models (e.g., related to the world's infrastructure); continuous and timely (including near real time and direct readout) observations (e.g., space, air, ground, over/underwater, Internet of Things (IoT), socioeconomic)
- Long-time records
- Analytics and artificial intelligence tools.

Earth System Digital Twins (ESDTs) Components

ESTO
Earth Science Technology Office
AIST

Earth Systems Digital Twins (ESDTs) are an emerging capability for understanding, forecasting, and conjecturing the complex interconnections among Earth systems, including anthropomorphic forcings and impacts to humanity.

=> What Now? What Next? What If?

What now?

Digital Replica . . .

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

What next?

Forecasting . . .

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

What if?

Impact Assessment . . .

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

An ESDT includes:

- 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

What is different about Digital Twins?

1. **Continuous integration** of timely data (real- or near-real-time for some applications, “timely for others)

2. **Interactivity** with users => “playing with the models and the data” for policy/decision making and conjecturing/planning

3. Heavy use of **Machine Learning**
   - Data Fusion
   - Super-Resolution/Downscaling
   - Speeding up models => higher spatial and temporal resolution possible
   - Causal Reasoning

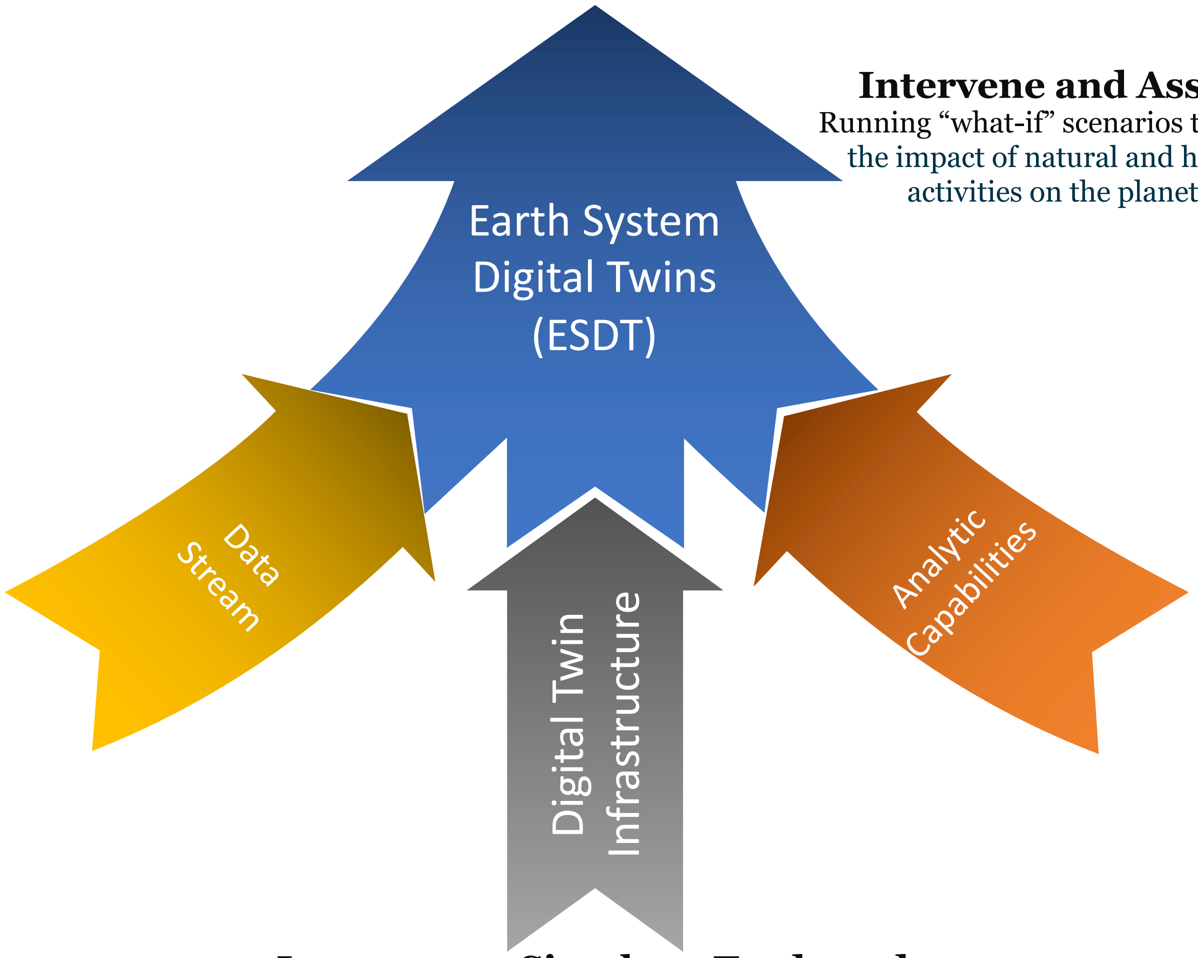
4. Integration of anthropomorphic forcing and **impact models**

AIST Focus Areas towards ESDT

ESTO
Earth Science Technology Office
AIST

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

Interrogate, Simulate, Trade and Visualize

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

ESTO AIST Earth System Digital Twins (ESDTs)

Summary of Activities 2021-2023

ESTO Earth Science Technology Office AIST

- **AIST-21 Solicitation**, first US government Solicitation requesting Digital Twins Technology for Earth Science
- **16 current ESDT technology development projects** funded under the Advanced Information Systems Technology (AIST) program focusing on developing:
  - Underlying analytic capabilities to build Digital Replicas
  - Novel ESDT infrastructure technologies
  - Surrogate modeling and ML emulators
  - Preliminary prototypes including interconnected modeling.
- Workshops and other community meetings to explore science use cases, enabling technologies, frameworks, prototyping, interoperability, and federation:
  - **AIST ESDT Workshop and Report**: Oct 26-28, 2022
  - **Standards for Interoperable Digital Twins Workshop**: Sep 18, 2023
  - **ESDT Architecture Framework document**
- ESDTs will play a critical role in NASA's Earth Science to Action initiative.
- Collaboration with ESA, Destination Earth, CNES, and others

AIST ESDT webpage: https://esto.nasa.gov/earth-system-digital-twin/

AIST-21 Solicitation ESDT Awards

ESTO
Earth Science Technology Office
AIST

- **Analytic Collaborative Frameworks (ACF) Towards ESDT**

PI's Name	Organization	Title
Thomas Allen	Old Dominion University	Pixels for Public Health: Analytic Collaborative Framework to Enhance Coastal Resiliency of Vulnerable Populations in Hampton Roads, Virginia (VA)
Arlindo Da Silva	NASA Goddard Space Flight Center (GSFC)	An Analytic Collaborative Framework for the Earth System Observatory (ESO) Designated Observables
Thomas Huang	NASA Jet Propulsion Laboratory (JPL)	Fire Alarm: Science Data Platform for Wildfire and Air Quality

- **AI and ML-based Surrogate Modeling for ESDT**

PI's Name	Organization	Title
Allison Gray	Univ. of Washington, Seattle	A prototype Digital Twin of Air-Sea Interactions
Christopher Keller	Morgan State University (MSU)	Development of a next-generation ensemble prediction system for atmospheric composition
Gavin Schmidt	NASA Goddard Inst. for Space Studies (GISS)	Development of digital twin technologies for climate projections
Jouni Susiluoto	NASA Jet Propulsion Laboratory (JPL)	Kernel Flows: emulating complex models for massive data sets

AIST-21 Solicitation ESDT Awards (cont.)

- **ESDT Infrastructure**

PI's Name	Organization	Title
Thomas Clune	NASA Goddard Space Flight Center (GSFC)	A Framework for Global Cloud Resolving OSSEs
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
Matthias Katzfuss	Texas A&M University (TAMU)	A scalable probabilistic emulation and uncertainty quantification tool for Earth-system models
Tanu Malik	De Paul University	Reproducible Containers for Advancing Process-oriented Collaborative Analytics

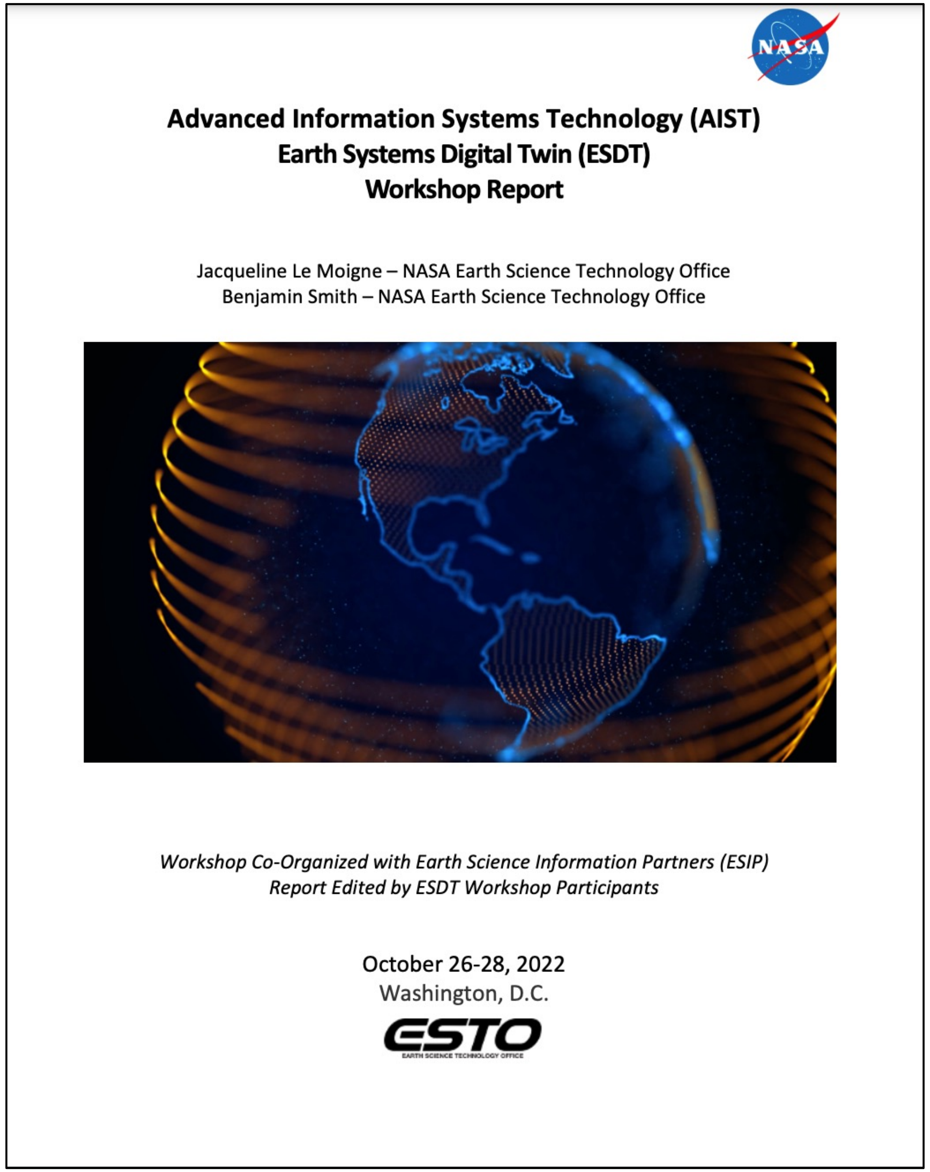
- **ESDT Prototypes**

PI's Name	Organization	Title
Rajat Bindlish	NASA Goddard Space Flight Center (GSFC)	Digital Twin Infrastructure Model for Agricultural Applications
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*
Thomas Huang	NASA JPL, GSFC and LaRC	Integrated Digital Earth Analysis System (IDEAS)
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
Alex Ruane	NASA Goddard Inst. for Space Studies (GISS)	An Urban Information System to Assess Neighborhood Climate Risk and Daily Exposures in Cities

Reference: 2022 ESDT Workshop Report

ESTO
Earth Science Technology Office
AIST

Report available
on AIST Website:
https://esto.nasa.gov/files/ESDT_Workshop_Report.pdf



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

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

October 26-28, 2022
Washington, D.C.

ESTO

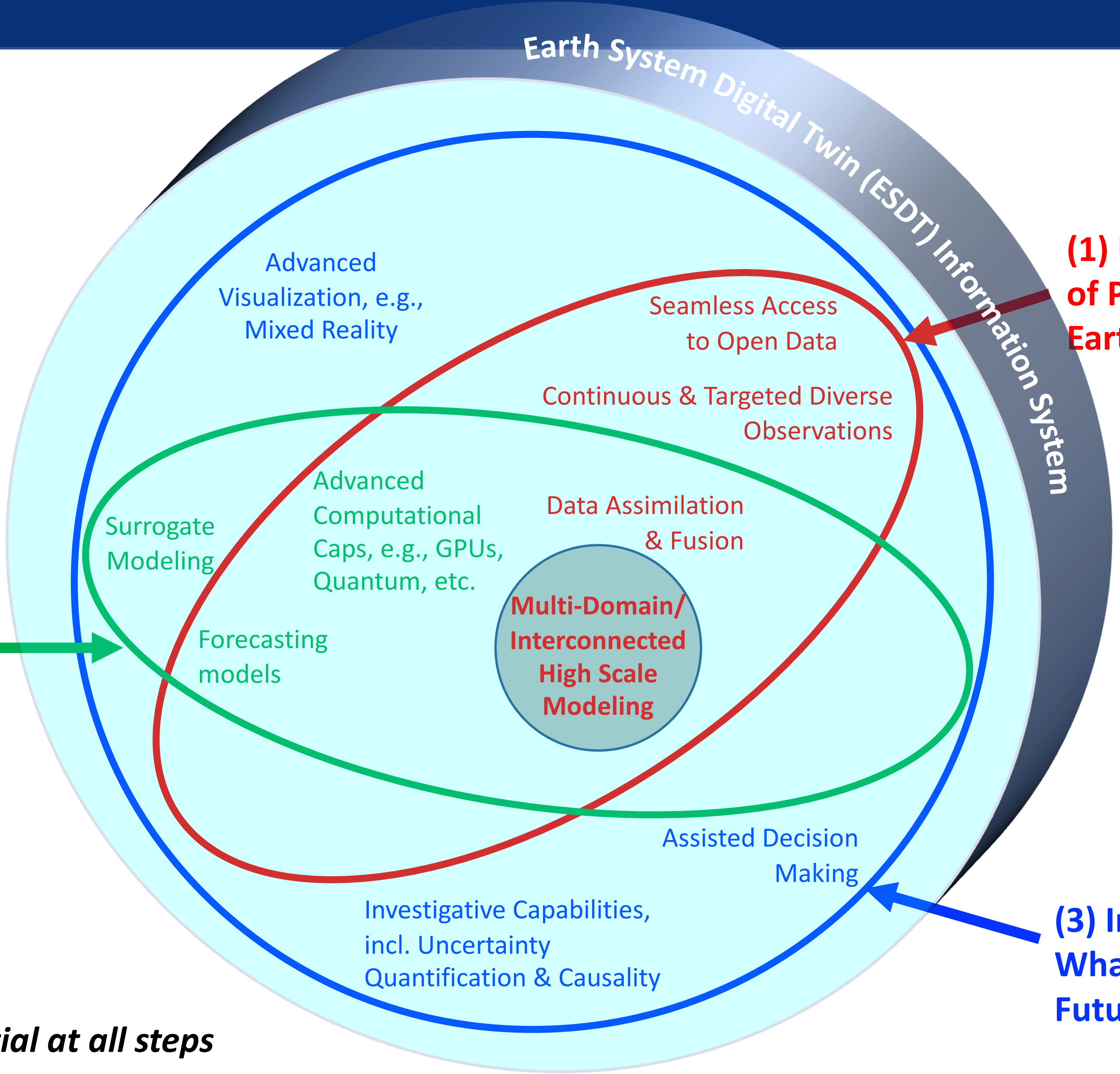
Workshop Summary

ESTO
Earth Science Technology Office
AIST

- Same digital replica can address needs of multiple users at various resolutions and for various applications
  - Scientific experts: ESDT capabilities built around ESMs
  - Science and applications users: more accurate forecasts and what-if simulations from varying initial and impact conditions.
  - Decision makers: what-if capability enables an exploration of alternatives and their impact on Earth systems and human activities, while the digital replica and forecasting capabilities provide a comprehensive interactive environment for understanding and monitoring current conditions and their evolution.
  - General public: ESDT inform daily activities and understand our changing planet.
- Global vs. local, multi-domain vs. thematic (e.g., some domains such as Climate or Weather will require a global Digital Twin while science areas such as Biodiversity might be more local)
- Overall, could imagine a 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.
- Challenges of building optimal digital twins:
  - Interoperability, including standards and protocols to be built from the beginning: at syntactic, semantic, legal and organizational levels
  - How to organize each digital replica, including various types of raw data, Analysis Ready Data (ARD) and information, and using various solutions, including Data Cubes, Data Lakes, pointers, or on demand
  - Type and level of interactivity and refresh rate, visualizations and human interfaces
  - Advanced technologies in Machine Learning, explainability, uncertainty quantification, validation, etc.

AIST ESDT Capabilities

ESTO
Earth Science Technology Office
AIST



(1) Up-to-Date *Digital Replica* of Past and Current States of Earth System

(2) Real-Time (RT) or Near-RT *Forecasting* of Future States of Earth System

(3) Impact Assessment by running What-If Scenarios for Hypothetical Future States of Earth System

Note: *Machine Learning essential at all steps*

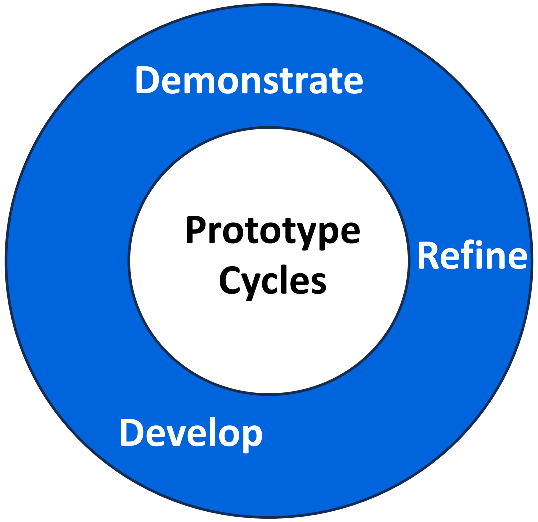
ESDT Science Use Cases/Scenarios

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.

ESDT Benefits to NASA Earth Science
Mapping ESDT Use Cases to Earth Action

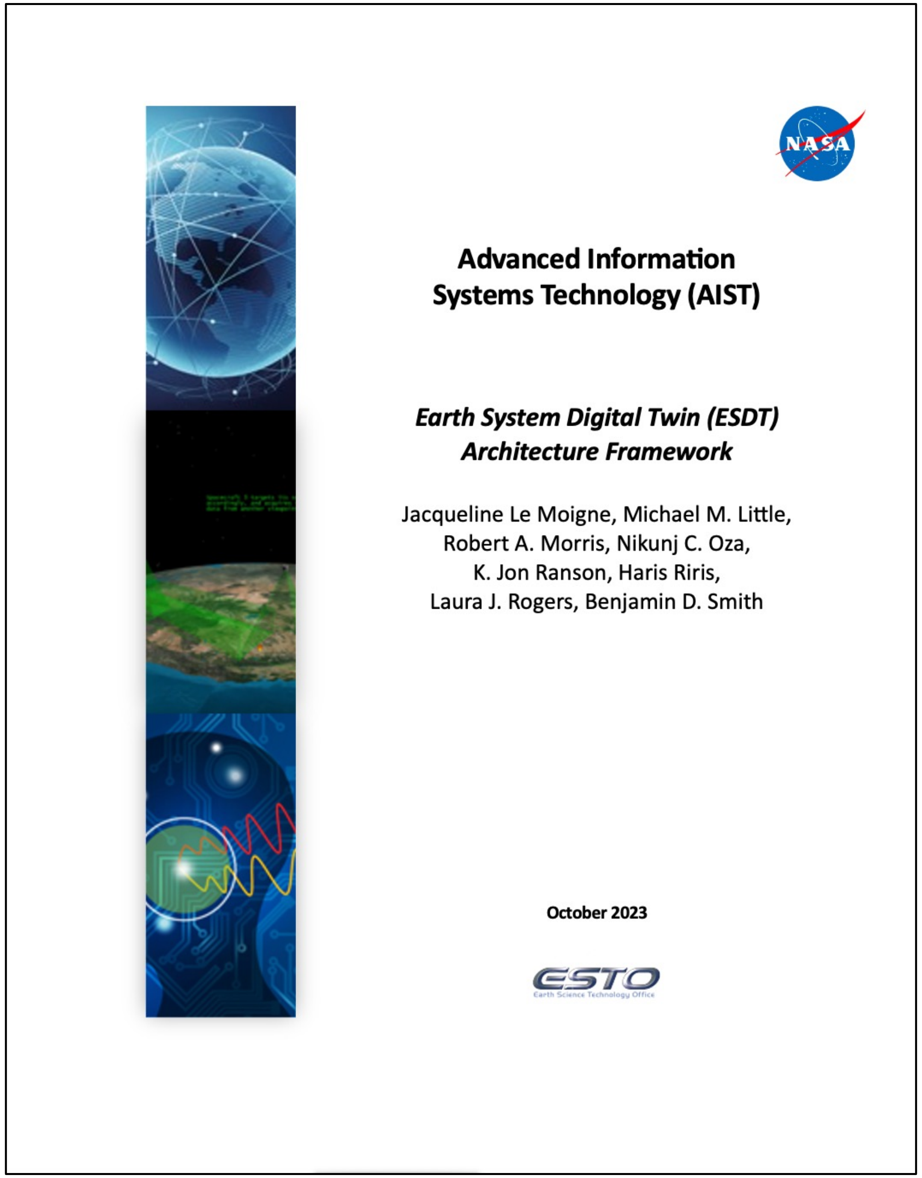
Current AIST ESDT Use Cases
Wildfires
Ocean Carbon
Water Cycle
Central Africa Carbon Corridors
Atmospheric Boundary Layer
Coastal Zone Digital Twin



Potential Earth Action Focus Areas
Greenhouse gas measurement and monitoring
Agriculture
Wildfire Risk and Recovery
Sea Level, Coastal Risk and Resilience
Air Quality
Environmental Health, Water Quality and Infectious Disease
Energy and Sustainable Infrastructure
Disasters and Extreme Events
Water Resources
Biodiversity and Ecological Conservation
National Security
Terrestrial Carbon Accounting
Human-Ocean interactions

Reference: ESDT Architecture Framework Document

Document available
on AIST Website:
https://esto.nasa.gov/files/AIST/ESDT_ArchitectureFramework.pdf


Advanced Information
Systems Technology (AIST)

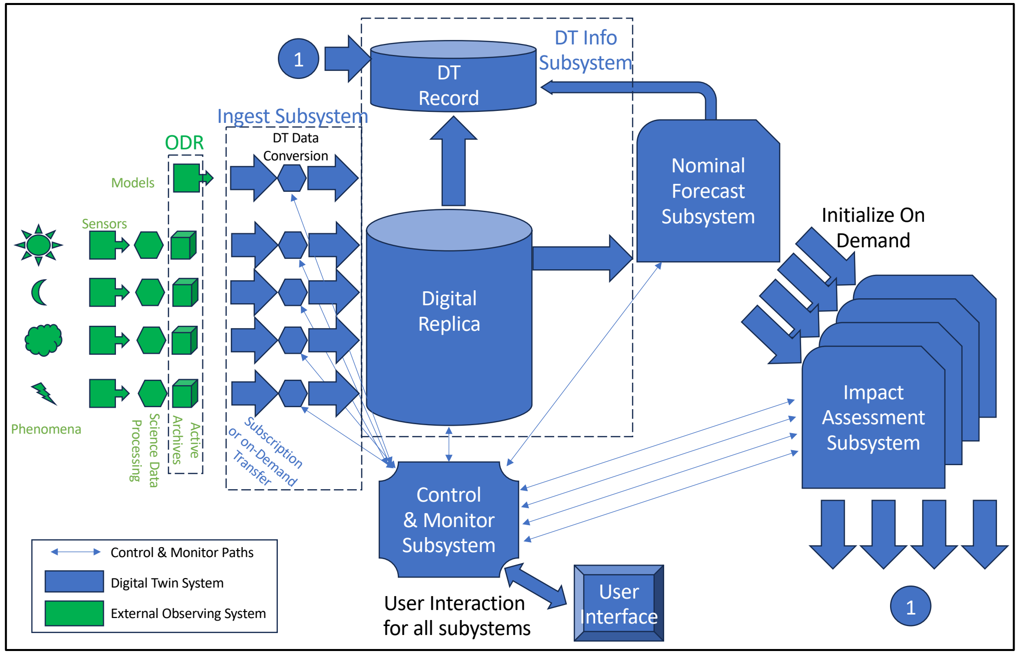
Earth System Digital Twin (ESDT)
Architecture Framework

Jacqueline Le Moigne, Michael M. Little,
Robert A. Morris, Nikunj C. Oza,
K. Jon Ranson, Haris Riris,
Laura J. Rogers, Benjamin D. Smith

October 2023

ESDT Architecture Framework *Considerations*

- **Consider architecture principles**
  - Modularity
  - Process automation and error checking
  - Comply with Open Source Science principles from SPD-41A
  - Permit evolution of concepts and uses and reasonable addition of new components
  - Provide the Glue to stitch together all ESD capabilities
  - Open-standard interfaces enabling opportunities for broader use
  - Interfaces for federation with other ESDTs
  - User interfaces for a range of skill levels and interests (i.e.,"from farmer to scientist")

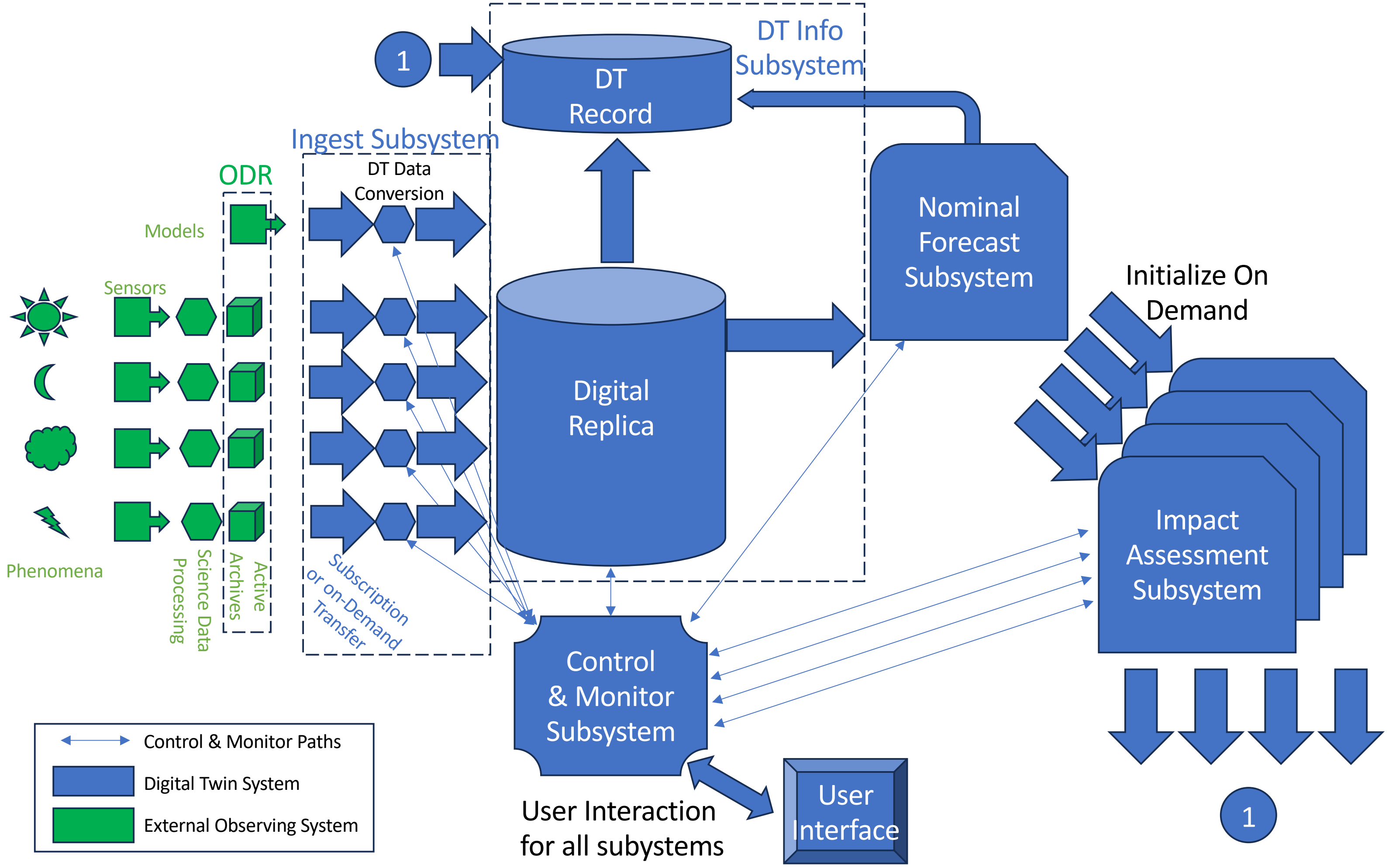


- **Enable component technology developers to consider their place in the overall architecture**
  - Re-use beyond a single architecture
  - Identify technology gaps and what is required to fill them

ESDT Conceptual System Diagram

ESTO Earth Science Technology Office AIST

An ESDT architecture must consider the full range of components and their relationships



Functional components:

- Observational Data Repository (ODR)
- Ingest Subsystem (ISS)
- DT Information Subsystem (DISS)
- Nominal Forecast Subsystem (NFSS)
- Impact Assessment Subsystem (IASS)
- Control and Monitor Subsystem (CMSS)
- User Interface Subsystem (UISS)

Architecture design may combine components or group them differently

Conclusion/Next Steps

AIST-23 Solicitation Future Selections => 2 or 3 end-to-end ESDT Prototypes (2024-2027)

Coastal Zone Digital Twin (NASA, NOAA and CNES) => 1st Prototype expected early 2025

Some Overarching Questions:

- How will various data, models, ESDT interoperate/be federated? Which basic interfaces/standards/protocols will be required? Syntactic, semantic, legal and organizational levels
- What are the main architecture components of an ESDT?
  - What is the role of Machine Learning for ESDT?
  - What is the role of Open Science for ESDT?
  - Which computational resources will be required? Cloud, GPU's, Quantum, Neuromorphic, etc.?
  - How will continuous data will be integrated? How often will digital replica be refreshed? Which user interfaces?
- How do we validate ESDT (e.g., using historical data, etc.)? How to quantify uncertainty?
- Which sustainable digital twin governance model should be adopted to address software configuration changes, security and full life cycle management?

ESTO
Earth Science Technology Office