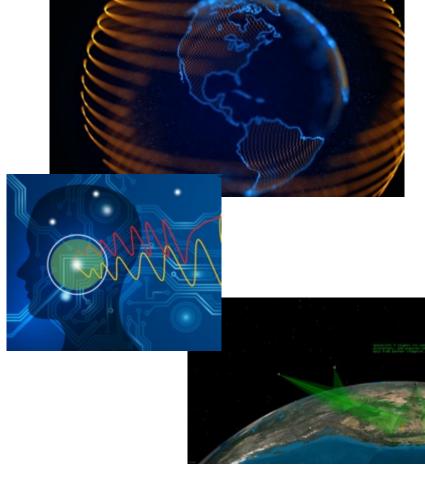


NASA ESTO Advanced Information Systems Technology (AIST)

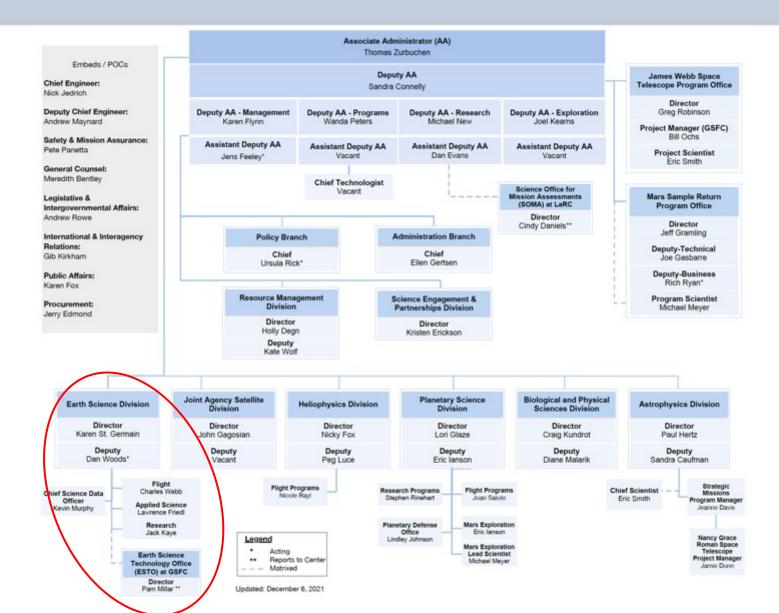
Jacqueline Le Moigne

November 2022



NASA Science Mission Directorate







ESTO leads technology development activities for the Earth Science Division. Through a science-driven competitive process it enables the next generation of instruments and information systems that advance our ability to study the Earth.

ESTO comprises five program lines:

| ATIP | Advanced Technology Initiatives Program |
|-------|--|
| IIP | Instrument Incubator Program |
| AIST | Advanced Information Systems Technology |
| DSI | |
| FIRET | Fire Technologies |



ESTO leads technology development activities for the Earth Science Division. Through a science-driven competitive process it enables the next generation of instruments and information systems that advance our ability to study the Earth.

ESTO comprises five program lines:

| ATIP | Advanced Technology Initiatives Program |
|-------|---|
| IIP | Instrument Incubator Program |
| AIST | Advanced Information Systems Technology |
| DSI | • Decadal Survey Incubation |
| FIRET | Fire Technologies |

ESTO and AIST Goals



• ESTO Goals

End-to-End Technology Development Approach

- o Identify technology needs based on Decadal Survey and annual requirements reviews
- Develop technologies through competitive peer-reviewed solicitations
- Assess the development strategy and maturity of funded technologies and leverage investments through internal NASA program synergy and partnerships with federal agencies, academia, and industry
- o Infuse maturing technologies into future missions and measurements

• AIST Goals

Innovate in technology development to enable:

- New and unique measurement collection capabilities through distributed sensing
- Optimizing Science missions return on investment through flexible and rapid information integration
- Agile Science investigations through data analytics and artificial intelligence tools and algorithms

AIST Objectives



Innovate in technologies that enable:

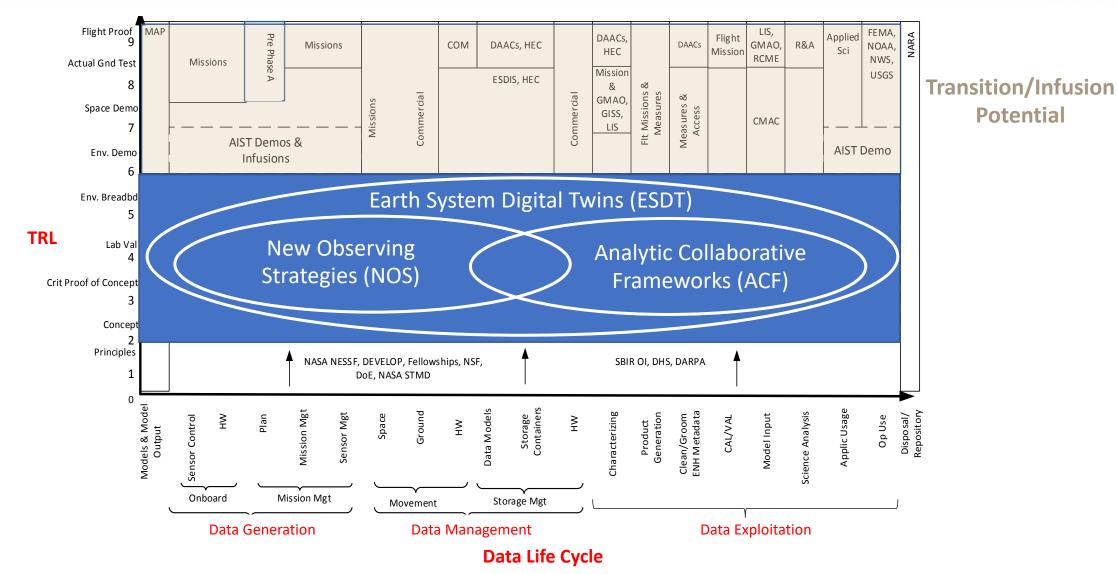
- O1. New observation measurements and new observing systems design and operations through intelligent, timely, dynamic, and coordinated distributed sensing
 > New Observing Strategies (NOS)
- O2. Agile science investigations that fully utilize the large amount of diverse observations using advanced analytic tools, visualizations, and computing environments, and that interact seamlessly with relevant observing systems => Analytic Collaborative Frameworks (ACF)
- O3. Developing integrated Earth Science frameworks that mirror the Earth with state-of-the-art models (Earth system models and others), timely and relevant observations, and analytic tools. This thrust will provide technology for enabling near- and long-term science^{*} and policy decisions => Earth System Digital Twins (ESDT)

More generally, provide "Science Data Intelligence"

^{* &}quot;Science decisions" including planning for the acquisition of new measurements; the development of new models or science analysis; the integration of Earth observations in novel ways; applications to inform choices, support decisions, and guide actions for societal benefit; etc.

AIST Program Scope

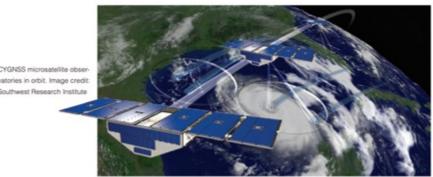






New Observing Strategies (NOS)

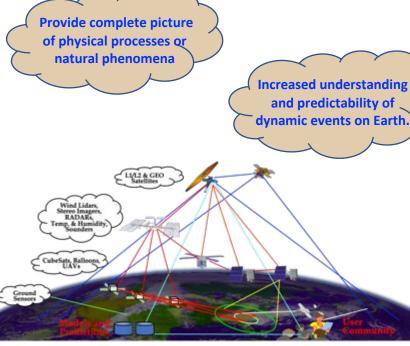
NOS for Optimizing Measurements Design & Dynamically Capturing full Science Events



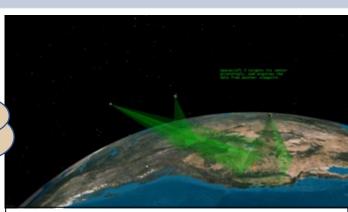
Distributed Spacecraft Mission (DSM): mission involving multiple spacecraft to achieve one or more common goals.

OBJECTIVES:

Multiple collaborative nodes from multiple organizations (NASA, OGAs, Industry, Academia, International) from multiple vantage points and in multiple dimensions (spatial, spectral, temporal, radiometric)



A SensorWeb is a distributed system of *sensing nodes* (space, air or ground) that are interconnected by a *communications fabric* and that functions as a single, highly coordinated, virtual instrument.



A special case of DSM is an Intelligent and Collaborative Constellation (ICC) which involves the combination of:

- Real-time data understanding
- Situational awareness
- Problem solving;
- Planning and learning from experience
- Communications & cooperation between several S/C

Actively acquire data in coordination with other sensors, models in response to measurement needs and/or science events

1. Design and develop New Observing Concepts:

- From Decadal Survey or Model; Various size spacecraft; Systems of systems (Internet-of-Space); Various organizations
- Perform trades on sensor number/type, spacecraft, orbits; resolutions; onboard vs. on-the-ground computing; intersensor communications, etc.
- System being designed in advance as a mission or observing system or incrementally and dynamically over time
- 2. Respond to various science and applied science events of interest: Various overall observation timeframes; Various area coverages; Dynamic/Timely; Scheduling, re-targeting/re-pointing assets, as possible

System-of-Systems NOS-Testbed for technologies & concepts validation, demonstration, comparison and socialization

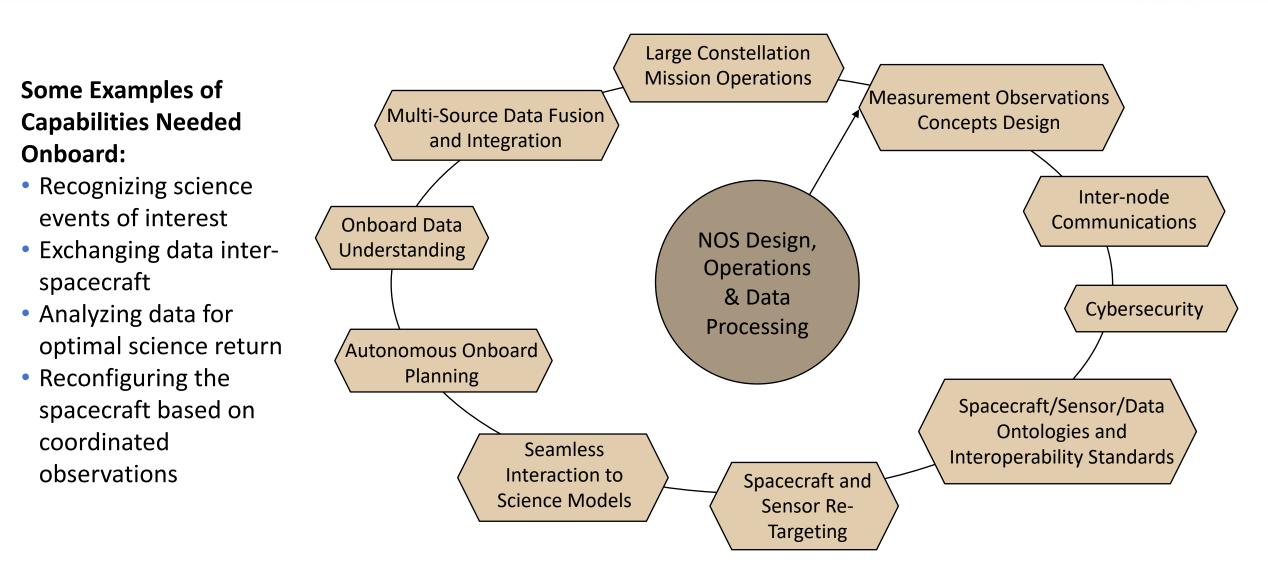
NOS Application Cases



| Mission Type <i>Timeframe</i> <i>Application</i> | Tactical Observing System Seconds-minutes Point event/phenomenon | Seconds-minutes Hours-days | |
|--|---|---|--|
| Example | Detect and observe volcanic activity | Increase spatial observation of primary forest burning as input into long-term Air Quality and Climate models | Select observing strategy to optimize all measurements that will improve hydrologic estimates |
| Functions | Detect emergent event Deploy observation assets | Deploy observation assets Digest information sources | Design observation system Digest information sources |
| Capabilities | Responsiveness Interaction Dynamics Adaptation | Resource allocation Coordination Data assimilation Prediction/ forecasting | Platform selection Coordination Data assimilation State estimation (belief) |

Technologies Needed for NOS







NOS-T Relevant

| PI's Name | Organization | Title | Synopsis |
|--------------------|---------------------------------|--|---|
| Mahta Moghaddam | U. of Southern California | SPCTOR: Sensing Policy Controller and OptimizeR | Multi-sensor coordinated operations and integration for soil moisture, using ground- based and UAVs "Sensing Agents". |
| Jim Carr | Carr Astro | StereoBit: Advanced Onboard Science Data Processing to Enable Future Earth Science | SmallSat/CubeSat high-level onboard science data processing demonstrated for multi- angle imagers, using SpaceCube processor and CMIS Instrument, and Structure from Motion (SfM). |
| Sreeja Nag | NASA ARC | D-SHIELD: Distributed Spacecraft with Heuristic Intelligence to Enable Logistical Decisions | Suite of scalable software tools - Scheduler, Science Simulator, Analyzer to schedule the payload ops of a large constellation based on DSM constraints (mech, orb), resources, and subsystems. Can run on ground or onboard. |
| Paul Grogan | Stevens Institute of Technology | Integrating TAT-C, STARS, and VCE for New Observing Strategy Mission Design | Inform selection and maturation of Pre-Phase A distributed space mission concept, by integrating: TAT-C: architecture enumeration and high-level evaluation (cost, coverage, quality); STARS: autonomous/adaptive sensor interaction (COLLABORATE); VCE: onboard computing and networking |

• OSSEs (Observing System Simulation Experiments)

| PI's Name | Organization | Title | Synopsis |
|------------------|----------------|--|--|
| Derek Posselt | NASA JPL | Parallel OSSE Toolkit | Fast-turnaround, scalable OSSE Toolkit to support both rapid and thorough exploration of the trade space of possible instrument configurations, with full assessment of the science fidelity, using cluster computing. |
| Bart Forman | U. of Maryland | Next Generation of Land Surface Remote Sensing | Create a terrestrial hydrology OSSE/mission planning tool with relevance to terrestrial snow, soil moisture, and vegetation using passive/active microwave RS, LiDAR, passive optical RS, hydrologic modeling, and data assimilation, using LIS and TAT-C. |
| Ethan Gutmann | UCAR | Future Snow Missions: Integrating SnowModel in LIS | Improve NASA modeling capabilities for snow OSSE, to plan and operate a future cost- effective snow mission by coupling the SnowModel modeling system into NASA LIS. |

AIST-21 NOS Awards



NOS for Smart Sensors and Onboard Intelligence

| PI's Name | Organization | Title | Synopsis |
|----------------------|--|--|---|
| William Blackwell | MIT Lincoln Labs | Sensor-in-the-Loop Testbed to Enable Versatile/Intelligent/Dynamic Earth Observation (VIDEO) | Develops a methodology and test approach for a scene measured by a sensor to be able to configure the sensor in real-time during the scene measurement. Will significantly improve the resolution of the retrieved atmospheric fields in regions in which that improvement is most beneficial, while conserving resources in other regions. Includes two components: (1) Radiometric Scene Generator (RSG) using advanced metamaterial and its associated control software; (2) Intelligent processing and configuration software using feature detection and ML, running onboard the sensor, to detect and react to changes by dynamically optimizing the sensor response functions. |
| James Carr | Carr Astronautics | Edge Intelligence for Hyperspectral Applications in Earth Science for New Observing Systems | Will use the SpaceCube processor and its Low-power Edge Artificial Intelligence Resilient Node (SC- LEARN) coprocessor powered by Google Coral Edge Tensor Processing Units (TPUs) to implement two Al science use cases in hyperspectral remote sensing: (1) Use learned spectral signatures of clear-sky scenes to retrieve surface reflectance and therefore increase the efficiency of collecting land observations on ~68% cloudy planet (e.g., for SBG); (2) Classify artificial light sources after training against a catalog of lighting types. SC-LEARN will fly on STP-H9/SCENIC to the ISS with a Headwall Photonics HyperspecMV hyperspectral imager. |
| James MacKinnon | NASA Goddard Space Flight Center | Multi-Path Fusion Machine Learning for New Observing System Design and Operations | Proposes to develop a system based on data fusion and multi-path neural network ML to aid in the design and operation of multi-sensor NOS concepts. Will build ML-enabled analytic tools and advanced computing environment capabilities for NOS workflows that utilize large amounts of diverse airborne and satellite observations. Using multiple neural networks working in parallel, it will first be demonstrated with a forest productivity use case, with fusion of lidar, spectrometry, satellite-derived climatology and ecosystem modeling providing insights into the driving environmental factors that influence productivity. Then will be used for sensitivity studies to guide sensor and mission requirements traceability. |
| Daniel Selva | Texas A&M Univ. | 3D-CHESS: Decentralized, distributed, dynamic and context-aware heterogeneous sensor systems | Proposes to demonstrate proof of concept for a context-aware Earth observing sensor web consisting of a set of nodes with a knowledge base, heterogeneous sensors, edge computing, and autonomous decision-making capabilities. Context awareness refers to the nodes' ability to gather, exchange, and leverage contextual information to improve decision making and planning. Will demonstrate and characterize the technology in a multi-sensor in-land hydrologic and ecologic monitoring system performing 4 inter-dependent missions: studying non-perennial rivers and extreme water storage fluctuations in reservoirs and detecting and tracking ice jams and algal blooms. |

AIST-21 NOS Awards (cont.)



NOS for UAS Integration and NOS Prototypes

| PI's Name | Organization | Title | Synopsis |
|----------------------|---|--|---|
| Meghan Chandarana | NASA Ames Research Center | Intelligent Long Endurance Observing System | Proposes the development of the Intelligent Long Endurance Observing System (ILEOS) to help scientists build plans to improve spatio-temporal resolution of climate-relevant gases by fusing coarse-grained sensor data from satellites and other sources and plan High-Altitude Long Endurance (HALE) UAS flights to obtain finer-grain data. ILEOS will also enable observations for longer periods and of environments not accessible through in-situ observations and field campaigns. 3 components: (1) the Target Generation Pipeline to identify candidate target scenes; (2) the Science Observation Planner using automated planning and scheduling technology to automatically generate a flight plan; and (3) a Scientists' User Interface. |
| Carl Legleiter | USGS | An Intelligent Systems Approach to Measuring Surface Flow Velocities in River Channels | Will develop a New Observing Strategy (NOS) for measuring streamflow from a UAS using an intelligent system. Using the USGS/NASA UAS-based payload for measuring surface flow velocities in rivers (USGS & NASA), consisting of thermal/visible cameras, a laser range finder, and an embedded compute (integrated within a common software middleware), it will address both quality control during routine streamgaging operations by quantifying uncertainty, as well as autonomous route-finding during hazardous flood conditions using inter-sensor communications. Will be implemented for real-time processing onboard the platform. |
| Carrie Vuyovich | NASA Goddard Space Flight Center | A New Snow Observing Strategy in Support of Hydrological Science and Applications | Will develop the Snow Observing System (SOS) considering the most critical snow data needs along with existing and expected observations, models, and a future snow satellite mission. It will estimate SWE and snow melt throughout the season, targeting obs with the greatest impact. It will: evaluate/combine observations from existing missions; create a hypothetical experiment to determine optimal observing strategy; assess value of new potential sensors, e.g., commercial SS for filling gaps and higher frequency obs. Higher density observations for early warning in regions where concerns for flood, drought or wildfires will also be studied. |

New Observing Strategies Testbed

- Technologies to be deployed should be first integrated into a working *breadboard* where the components can be debugged and performance and behavior characterized and tuned-up.
- A system of this complexity should not be expected to work without full integration and experimental characterization as a "system of systems"

Testbed Main Goals:

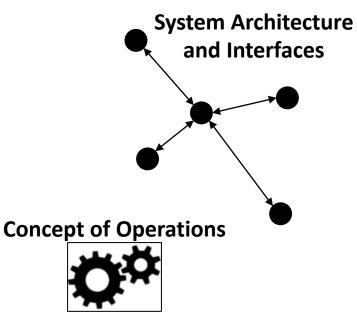
- 1. Validate new DSM/NOS technologies, independently and as a system
- 2. Demonstrate novel distributed operations concepts
- 3. Enable meaningful comparisons of competing technologies
- 4. Socialize new DSM technologies and concepts to the science community by significantly retiring the risk of integrating these new technologies.

NOS-T framework objective:

Enable disparate organizations to propose and participate in developing NOS software and information technology

Governance Model

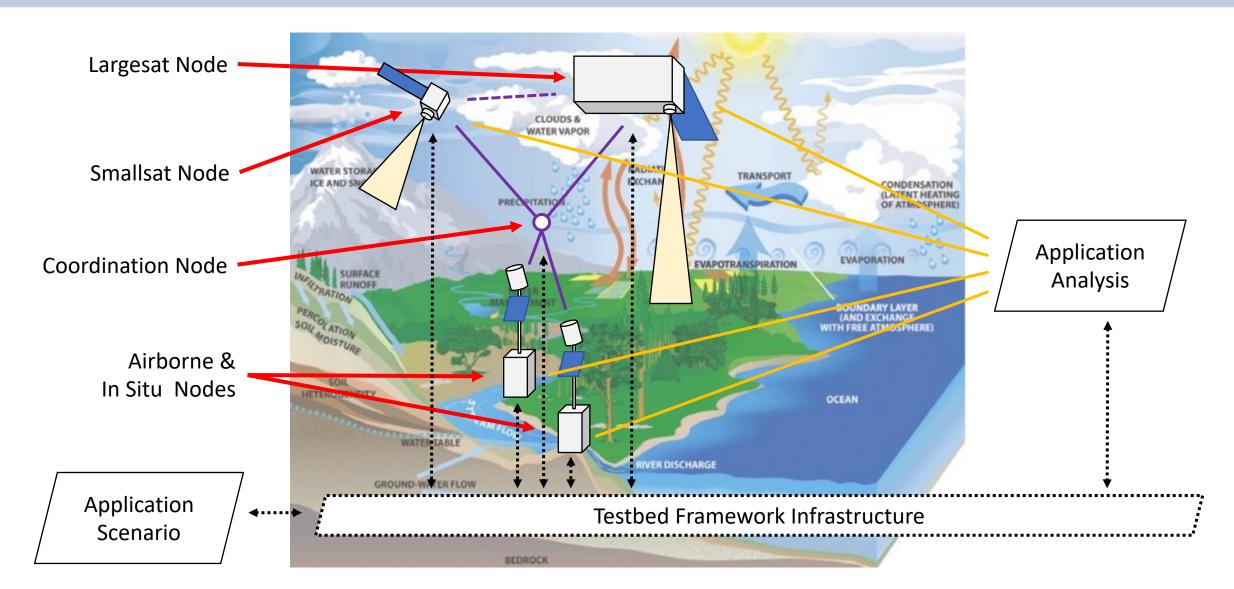






NOS Testbed Concept





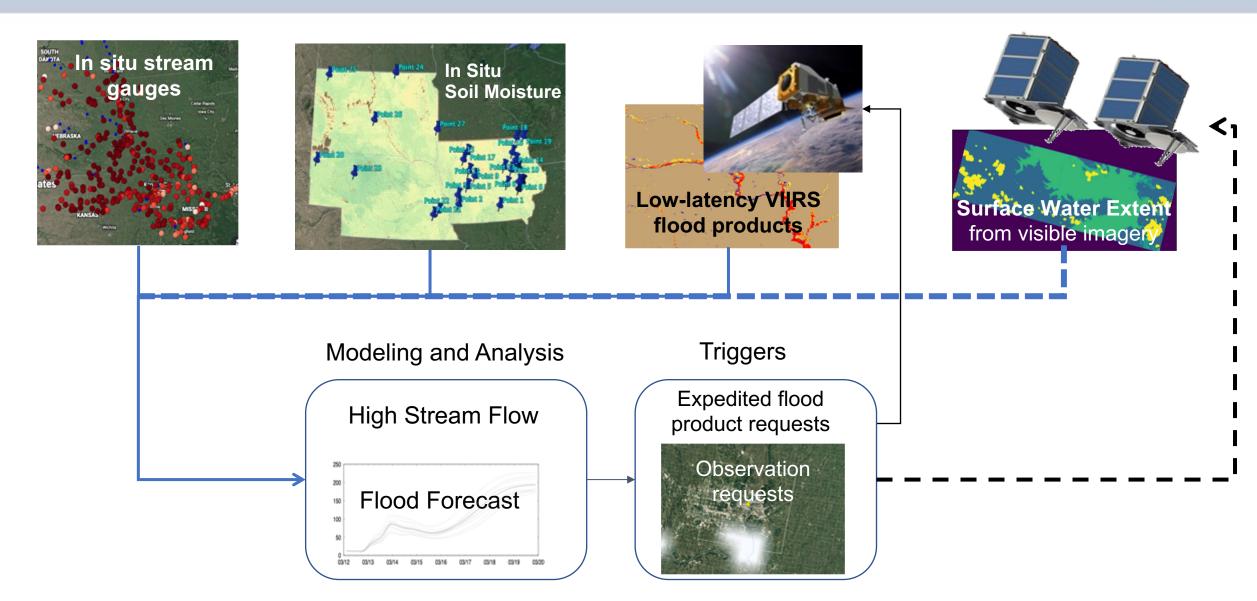
NOS-T Architecture and Pilot Projects



| PI's Name | Organization | Emails | Title | Synopsis |
|---|---|--|---|--|
| Tom McDermott & Paul Grogan & Jerry Sellers | Systems Engineering Research Center (SERC) | tmcdermo@stevens.edu; pgrogan@stevens.edu; jsellers@tsti.net | New Observing Strategies Testbed (NOS-T) Design and Development | Design the NOS-T framework to enable system-of systems experiments and testing; enable multi-party and geographically distributed participation and connected tests and operations; enables both open community and protected exchange of measurement data; provide a communications infrastructure; and simulate actual operational security challenges. |
| Chad Frost & Daniel Cellucci | NASA Ames | chad@nasa.gov; daniel.w.cellucci@nasa.gov | Earth Science "Tip and Cue" Technologies for a New Observing Strategy | Extend the capabilities of the Generalized Nanosatellite Avionics Testbed (G-NAT) and networked, state-of-the-art, miniaturized, tracking and sensing devices (termed 'tags'), developed in collaboration with USGS, to enable a tip-and-cue architecture for dynamically reconfigurable remote sensing. |
| Sujay Kumar & Rhae Sung Kim | NASA Goddard | sujay.v.kumar@nasa.gov; rhaesung.kim@nasa.gov | A Hydrology MIssion Design and Analysis System (H- MIDAS) | Extend LIS capabilities to: support the incorporation of distributed sensor observations fir hydrology; support the development of observation operators; perform data assimilation simulations and provide feedback to the observing systems. |
| Steve Chien & James Mason | NASA JPL | steve.a.chien@jpl.nasa.gov; james.mason@jpl.nasa.gov | Planning and Scheduling for Coordinated Observations | Develop a planning and scheduling framework for the NOS Testbed that will coordinate multiple observing assets (e.g. space, air, land) to perform coordinated and continuous measurements at varying scales (e.g. spatial, temporal). |
| Dan Crichton & Cedric David | NASA JPL | daniel.j.crichton@jpl.nasa.gov; cedric.david@jpl.nasa.gov | NOS Testbed Study and Science Use Cases Identification | Contribute to the definition of the NOS Testbed by identifying science use cases, observing assets, requirements, interfaces, and other design recommendations in close collaboration with the NOS Testbed Definition activity. |
| Louis Nguyen | NASA LaRC | | Ground Stations as a Service (GSaS) for Near Real-time Direct Broadcast Earth Science Satellite Data | Utilize GSaS to receive direct broadcast (DB) data from EOS to significantly reduce latency in acquiring LEO satellite observations (e.g., from 3-6 hours to 20-25 mins). It will provide ability to receive low latency LEO data without the need to own/maintain DB ground station; improve NASA Earth Science's ability to deliver lower latency products and therefore increasing optimal use; provide NOS with capability to schedule, coordinate, receive, and process DB data from EOS. |
| Jay Ellis | KBR/GSFC | nathaniel.j.ellis@nasa.gov | NOS Testbed Administration and Management | Administer and manage the NOS Testbed for disparate organizations to propose and participate in developing NOS software and information systems technology capabilities and services. |

NOS-Testbed Hydrology Demonstrations March 2021 – Historical Nebraska Flood + Live Mid-West Flood

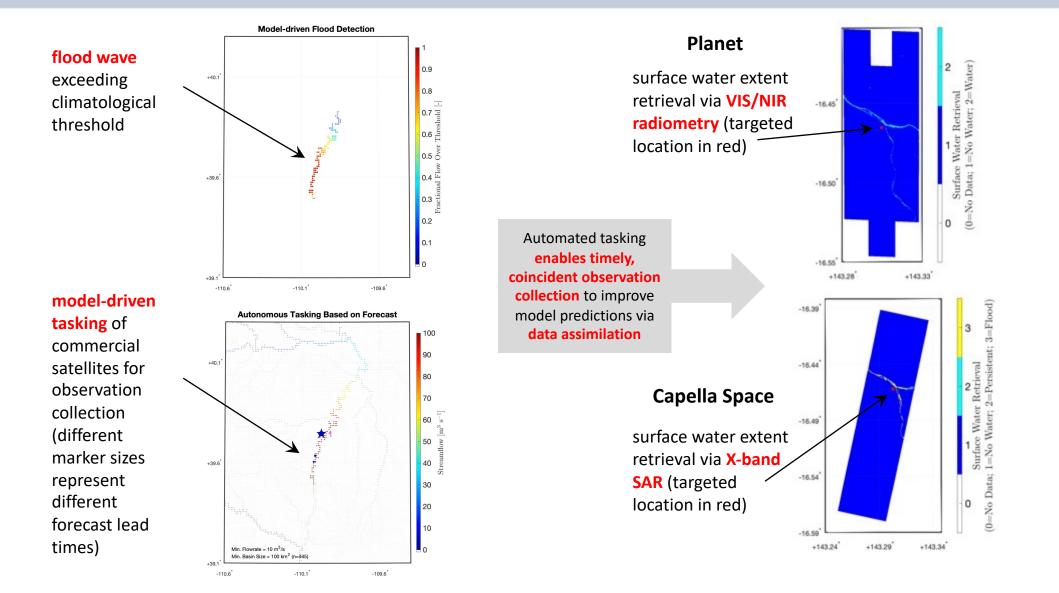




NOS-L: Autonomous, Model-driven Tasking of **Flood Events**



Successful NOS-L end-to-end live demonstration

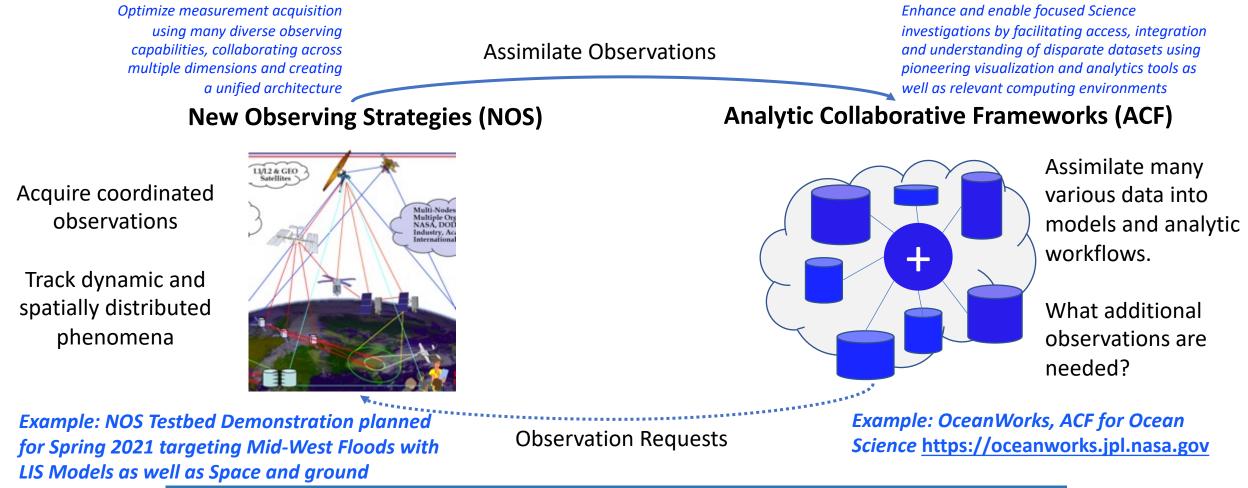




Analytic Collaborative Frameworks (ACF)

NOS and ACF for Science Data Intelligence





observations

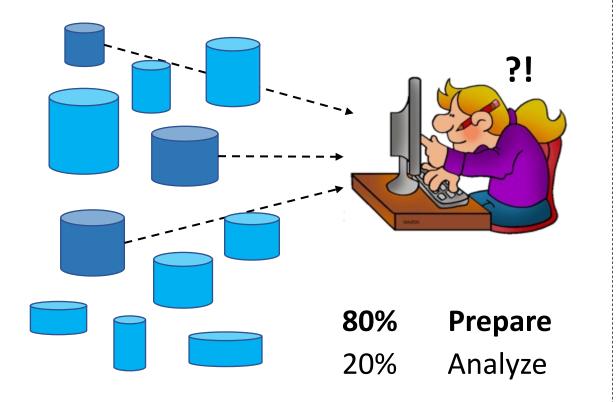
NOS+ACF acquires and integrates complementary and coincident data to build a more complete and in-depth picture of science phenomena

From Archives to Analytic Frameworks *Focus on the Science User*



Data Archives

Focus on data capture, storage, and management Each user has to find, download, integrate, and analyze



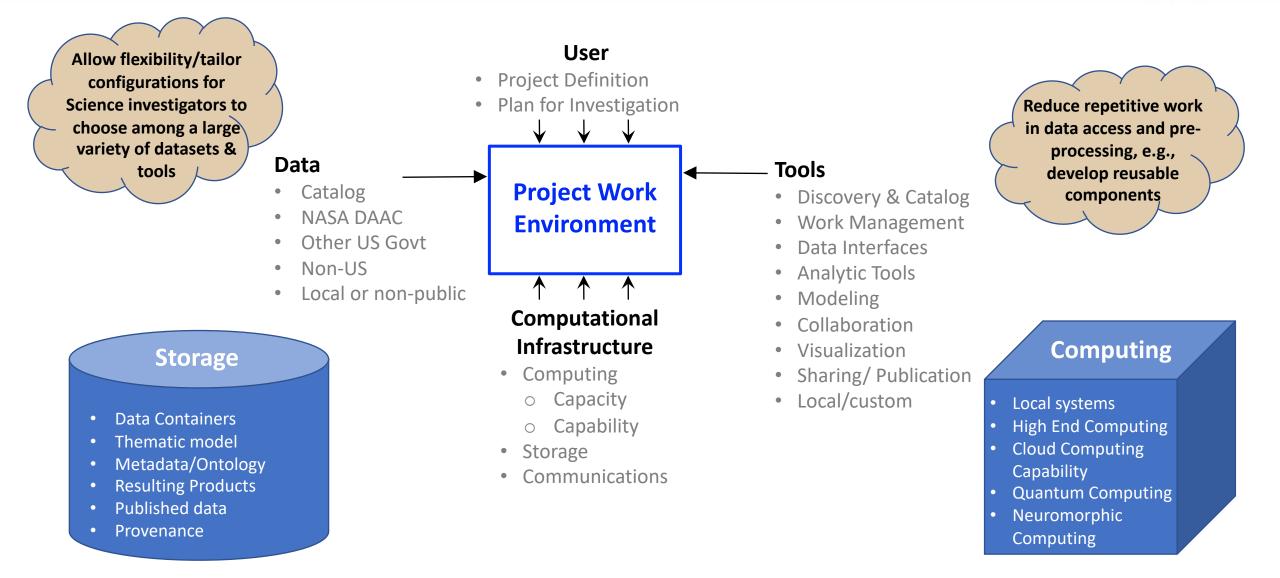
Analytic Centers Focus on the science user Integrated data analytics & tools tailored for a science discipline Search Analytics Science Applications Integration Visualization 20% Prepare Analyze 80%

Decision Support

Facilitates collaborative science across multiple missions and data sets

Analytic Collaborative Frameworks (ACF) Focus is on the Science User



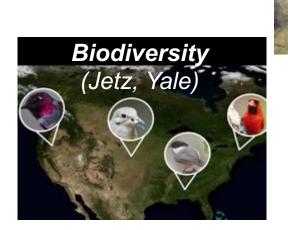


Analytic Collaborative Frameworks (ACF) support various Earth Science Disciplines



Air Quality

Precipitation (Beck, UAH)

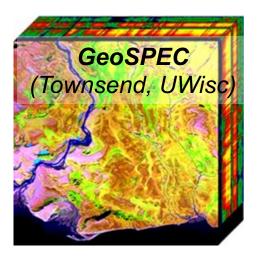


Methane

(Duren, UAz)

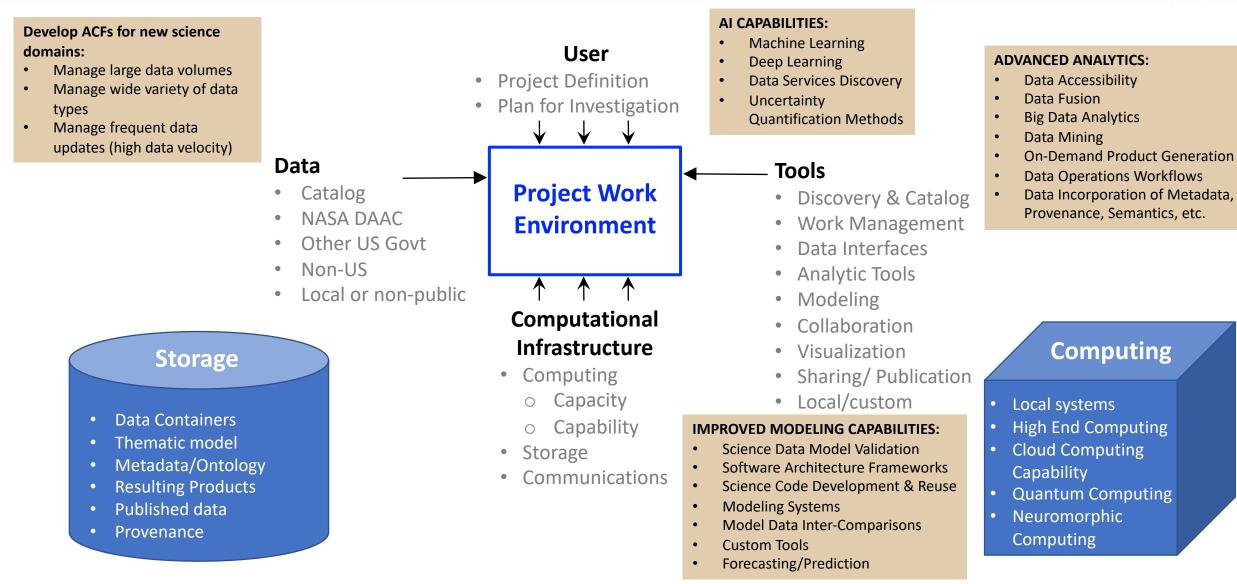


Aquaculture (Uz, GSFC)



Analytic Collaborative Frameworks (ACF) Focus is on the Science User





My Data dran lawers in the list to change the 2.00 Jupyter Notebook - Interactive Workbench AM-NOAA INFSDIS, Ire CONC-OSISAF, SMAP Rading 7,4527N, 77,1490W - 34,0695N, 70,2507W Mar 2015 - Nov 201

ACF Example – OceanWorks An Analytic Collaborative Framework (ACF) for Ocean Science



Dozens of Ocean Data Sets

- Stored in the cloud
- Integrated, ready to use
- Organized for fast search, subset and analysis

Search

Find relevant data sets

Cloud-based analytics

Analyze years of data over multiple data sets in seconds... without downloading data Run analytics across multiple data sets despite differences in scale, times, granules, etc.

Custom analytics

Scientists can also run their own custom tools and algorithms

Integrated data

Match up in-situ and remote sensing data, despite differences in scale and resolution

Visualization

Subsets, layers, animations. Integrates with ArcGIS and Jupyter Notebooks

https://oceanworks.jpl.nasa.gov



• Biodiversity ACF

| Award # | PI's Name | Organization | Title | Synopsis |
|--------------|------------------|--------------------------|--|--|
| AIST-18-0007 | Schollaert Uz | NASA GSFC | Supporting shellfish aquaculture in the Chesapeake bay using AI for water quality | Provide access to reliable information on a variety of environmental factors, not currently available at optimal scales in space and times, by using various data (sats and others) and AI for Pattern Recognition. |
| AIST-18-0031 | Moisan | NASA GSFC | NASA Evolutionary Programming Analytic Center (NEPAC) | Discover and apply novel algorithms for ocean chlorophyll using AI/ML (Genetic Programming) on satellite/in-situ obs and a user-friendly GUI to connect data and applications with HEC resources for improved science. |
| AIST-18-0034 | Jetz | Yale U. | Biodiversity - Environment Analytic Center | Near real-time monitoring of the biological pulse of our planet, using an online dashboard, taking into account various spatiotemporal resolutions, data uncertainty and biodiversity data biases, and supporting analysis, visualization and change detection across scales. |
| AIST-18-0043 | Townsend | U. Wisconsin, Madison | GeoSPEC: On-Demand Geospatial Spectroscopy Processing Environment on the Cloud | Develop a framework/processing workflow for on-demand cloud-based Hyperspectral/Spectroscopy Science Data Processing in preparation for SBG needs. Will provide options for new atmospheric & other types of corrections, possibilities for users' or commercial code. Will be tested with AVIRIS-Classic and –NG data. |
| AIST-18-0063 | Swenson | Duke University | Canopy condition to continental scale biodiversity forecasts | Characterize canopy condition from various spatio-temporal RS products (including drought indices and habitat structure) to predict supply of mast resources to herbivores (and threatened species) and visualize canopy condition and drought-stress maps |

• Land Cover ACF

| Award # | PI's Name | Organization | Title | Synopsis |
|--------------|-----------|----------------------|---|--|
| AIST-18-0020 | lves | U. Of WI, Madison | Valid time series analyses for satellite data | Develop new statistical tools to analyze large, time series of various remotely sensed datasets and provide statistical rigor and confidence to conclusions about patterns of change and to forecasts of future change, identifying patterns of annual trends, seasonal trends and phenological events, and analyzing the cause of these trends. |



• Air Quality ACF

| Award # | PI's Name | Organization | Title | Synopsis |
|--------------|-----------|------------------------|--|---|
| AIST-18-0011 | Martin | Washington U. | Development of GCHP to enable broad community access to high- resolution atmospheric composition modeling | Integrate atmospheric chemistry models online into Earth system models (ESMs) and offline using meteorological data, using the high-performance version of the GEOS-Chem global 3-D model of atmospheric chemistry (GCHP) and the Earth System Modeling Framework (ESMF) in its Modeling Analysis and Prediction Layer (MAPL) implementation. |
| AIST-18-0044 | Duren | NASA JPL | Multi-scale Methane Analytic Framework | ACF for methane data analysis spanning multiple observing systems and spatial scales with workflow optimization, analytic tools to characterize methane fluxes and physical processes, tools for data search and discovery, and a collaborative, web-based portal. |
| AIST-18-0072 | Henze | U. of CO, Boulder | Surrogate modeling for atmospheric chemistry and data assimilation | Advance computational tools available for AQ prediction, mitigation, and research by building a robust and computationally efficient chemical Data Assimilation system, merging research in compressive sampling and machine learning for large-scale dynamical systems and integrating multi-source data into an existing model. |
| AIST-18-0099 | Holm | City of Los Angeles | Predicting What We Breathe: Using Machine Learning to Understand Urban Air Quality | Link ground-based in situ and space-based remote sensing observations of major AQ components to classify patterns in urban air quality, enable the forecast of air pollution events, and identify similarities in AQ regimes between megacities around the globe, using science models and ML-based algorithms. |

• Precipitation ACF

| Award # | PI's Name | Organization | Title | Synopsis |
|--------------|-----------|-------------------------|--|--|
| AIST-18-0051 | Beck | U. Of AL, Hunstville | Cloud-based Analytic Framework for Precipitation Research | Leverage cloud-native technologies from the AIST-2016 VISAGE project to develop a Cloud-based ACF for Precipitation Research using a Deep Learning (CNNs) framework to provide an analysis-optimized cloud data store and access via on-demand cloud-based serverless tools. It will use coincident ground and space radar observations. |



• Disaster Management ACF

| Award # | PI's Name | Organization | Title | Synopsis |
|--------------|-----------|--------------|---|--|
| AIST-18-0055 | Coen | NCAR | Creation of a Wildfire Fire Analysis: Products to Enable Earth Science | Develop methods to create, test and assess wildland fire reanalysis products (standardized, gridded wildland fire information generated at regular intervals) using fire detection data, as well as coupled weather-wildland fire model and data assimilation. |
| AIST-18-0001 | Donnellan | NASA JPL | Quantifying Uncertainty and Kinematics of Earthquake Systems ACF (QUAKES-A) | Create a uniform crustal deformation reference model for the active plate margin of California by fusing data with widely varying spatial and temporal resolutions, quantifying uncertainty, developing data management and geospatial information services and providing collaboration and infusion into target communities. |
| AIST-18-0085 | Hua | NASA JPL | Smart On-Demand of SAR ARDs in Multi-Cloud & HPC | Enable full resolution time series analysis, high-accuracy flood and damage assessments with remote sensing SAR Analysis Ready Data (ARD), using Jupyter Notebooks and on- demand analysis across multi-cloud environments. |

• Cross-Cutting ACF Capabilities

| Award # | PI's Name | Organization | Title | Synopsis |
|--------------|-----------|-----------------------|---|--|
| AIST-18-0042 | Huffer | Lingua Logica | AMP: An Automated Metadata Pipeline | Automate and improve the use and reuse of NASA Earth Science data by developing a fully-automated metadata pipleline integrating ML and ontologies (SWEET) for a semantic, metadata mining from data. Developed in collaboration with GES DISC. |
| AIST-18-0059 | Zhang | Carnegie Mellon U. | Mining Chained Modules in Analytics Center Framework | Build a workflow tool as a building block for ACF, capable of recommending to Earth Scientists multiple software modules, already chained together as a workflow. The tool will leverage Jupyter Notebooks to mine software module usage history, develop algorithms to extract reusable chain of software modules, and develop an intelligent service that provides for personalized recommendations. |

AIST-21 ACF Awards



ACF Infrastructure

| PI's Name | Organization | Title | Synopsis |
|----------------|--------------------------------------|---|---|
| Ziad Haddad | NASA Jet Propulsion Laboratory | Thematic Observation Search, Segmentation, Collation and Analysis (TOS2CA) system | TOS2CA is a user-driven, data-centric system that can identify, collate, statistically characterize and serve Earth system data relevant to a given phenomenon relevant to ESO. It will facilitate the collation and analysis of data from disparate sources, help scientists to establish science- traceability requirements, quantify detection thresholds, define uncertainty requirements and establish data sufficiency to formulate truly innovative missions. Components will include: 1) a user-driven thematic data collector; 2) a statistical analyzer; and 3) a user-friendly visualization and data exploration toolkit. |
| Huikyo Lee | NASA Jet Propulsion Laboratory | Open Climate Workbench to support efficient and innovative analysis of NASA's high-resolution observations and modeling datasets | The Regional Climate Model Evaluation System (RCMES/JPL &UCLA) performs systematic evaluation of climate models and is powered by the Open Climate Workbench (OCW). The goal is to develop OCW v2.0 by extending the capabilities of OCW for characterizing, compressing, analyzing, and visualizing observational and model datasets with high spatial and temporal resolutions. It will run on AWS Cloud with special emphasis on developing two use cases: air quality impacts due to wildfires and elevation-dependent warming. |

AIST-21 ACF Awards (cont.)



Machine Learning for Modeling

| PI's Name | Organization | Title | Synopsis |
|-------------------------------|---|--|--|
| Yehuda Bock | Scripps Institute, Univ. of CA, San Diego | | Proposes to create open-source software to provide a rich, interactive environment where machine learning (ML) models are used as collaborator to direct the attention of the human analyst to non-physical artifacts and real transient events that require interpretation. The proposed system will be realized through two coupled sub-systems: a novel "back-end" ML software called the Transient and Artifact Continuous Learning System (TACLS), and a significant upgrade to existing "front end" interactive MGViz user environment, originally designed to view displacement time series and their underlying metadata, to now interact and display layers of spatiotemporal information. Will use Scripp's archive of thousands of artifacts and transients. |
| Stephanie Schollaert Uz | NASA Goddard Space Flight Center | Integration of Observations and Models into Machine Learning for Coastal Water Quality | Proposes to build upon ongoing collaborations with state agencies managing water water resources around the Chesapeake Bay to monitor water quality and ecosystem properties and how they change over time and space. Previous AISTpromising results used multispectral optical, medium spatial resolution satellite data trained using geophysical model variables within a machine learning (ML) architecture. Utilizing higher spatial resolution from commercial satellites, feature maps from many sensors of varying spatial, spectral, and temporal resolutions will be derived to determine the minimum set of requirements for RS of water quality, e.g. water clarity, phytoplankton blooms, and the detection of pollutants. |
| Brian Wilson | NASA Jet Propulsion Laboratory | SLICE: Semi-supervised Learning from Images of a Changing Earth | Proposes to investigate and characterize the efficacy of multiple Self- and Semi-Supervised Learning (SSL) techniques for representative image problems on Earth imagery, and then select the best for further infusion into mission and science workflows. Three top-level goals of the SLICE system are: (1) Establish the SLICE framework and platform on the AWS Cloud and supercomputing environments; (2) Investigate and characterize the accuracy of multiple SSL models (i.e. SimCLRv2, DINO, EsViT) on a variety of relevant remote sensing tasks with minimum labels (e.g., ocean phenomena); (3) Build and publish self- and semi-supervised learning models with a focus on the upper ocean small-scale processes in anticipation of several on-going and upcoming NASA missions (i.e. SWOT, WaCM, and PACE). |



ACF Prototypes

| PI's Name | Organization | Title | Synopsis |
|------------------|---|--|---|
| Colin Gleason | Univ. of Massachusetts at Amherst | A hosted analytic collaborative framework for global river water quantity and quality from SWOT, Landsat, and Sentinel-2 | Proposes to integrate data from the soon-to-be launched SWOT mission with traditional optical imagery in a single ACF to simultaneously co-predict river water quantity and quality at a scale not currently possible. The existing Confluence system will be extended and the outputs of the proposed ACF will be: (1) Unique seamless data environment for SWOT and optical data; (2) Extended library of algorithms for water quantity; (3) Novel library of computer vision algorithms for water quality; and (4) Automated computational environment to produce river water quantity and quality products, globally. Multiplying the mass flux of water (via SWOT) by its constituent concentrations (via optical data) provides a constituent mass loading (e.g., sediment, algae) in the world's river systems and therefore a direct benefit to society and ecosystems. |
| Seungwon Lee | NASA Jet Propulsion Laboratory | Ecological Projection Analytic Collaborative Framework (EcoPro) | Proposes to build EcoPro to support multidisciplinary teams conducting ecological projection studies, collaborations, applications, and new observation strategy developments. EcoPro will contain: (1) an analytic toolkit to perform multidisciplinary analyses; (2) a data gateway to organize, store, and access key input and output datasets; and (3) a web portal to publish and visualize the results of the studies and to provide a virtual collaborative workspace. The goals are to provide capabilities to advance ecological projection on multi-decadal timescales, as well as ecological forecasting on shorter timescales. |



Earth System Digital Twins (ESDT)

AIST Earth System Digital Twin(s)



AIST defines an Earth System Digital Twin (ESDT) as 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.

Effective ESDTs enable users to run hypothetical scenarios to improve the understanding, prediction of and mitigation/response to Earth system processes, natural phenomena and human activities as well as their many interactions. An ESDT is a type of integrated information system that, for example, enables continuous assessment of impact from naturally occurring and/or human activities on physical and natural environments.

AIST ESDT strategic goals are to:

- Develop information system frameworks to provide continuous and accurate representations of systems as they change over time;
- Mirror various Earth Science systems and utilize the combination of Data Analytics, Artificial Intelligence, Digital Thread*, and state-of-the-art models to help predict the Earth's response to various phenomena;
- Provide the tools to conduct "what if" investigations that can result in actionable predictions.

^{*} The digital thread designates the communication framework that links all digital twin data flow throughout its lifecycle.

Earth System Digital Twins Components



Digital Replica . . .

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

Forecasting . . .

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

Impact Assessment . . .

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

Earth System Digital Twins Components



Digital Replica

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

Forecasting What next

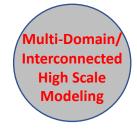
An integrated picture of how Earth systems will evolve in the future from the current state.

Impact Assessment

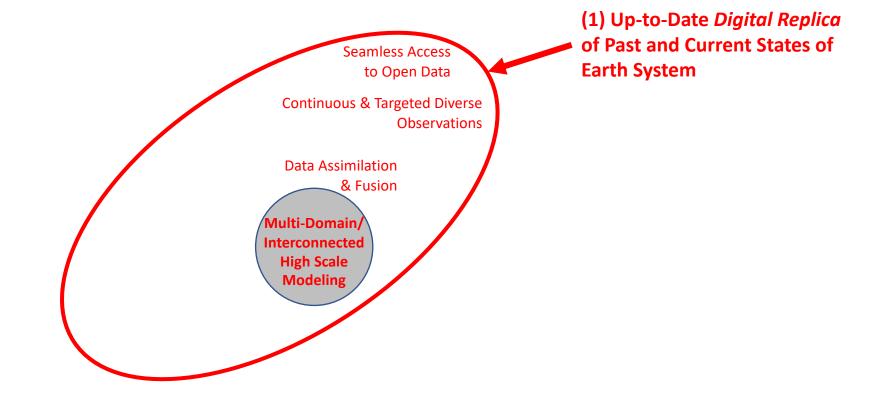
An integrated picture of how Earth systems could evolve under different hypothetical what-if scenarios.

- **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

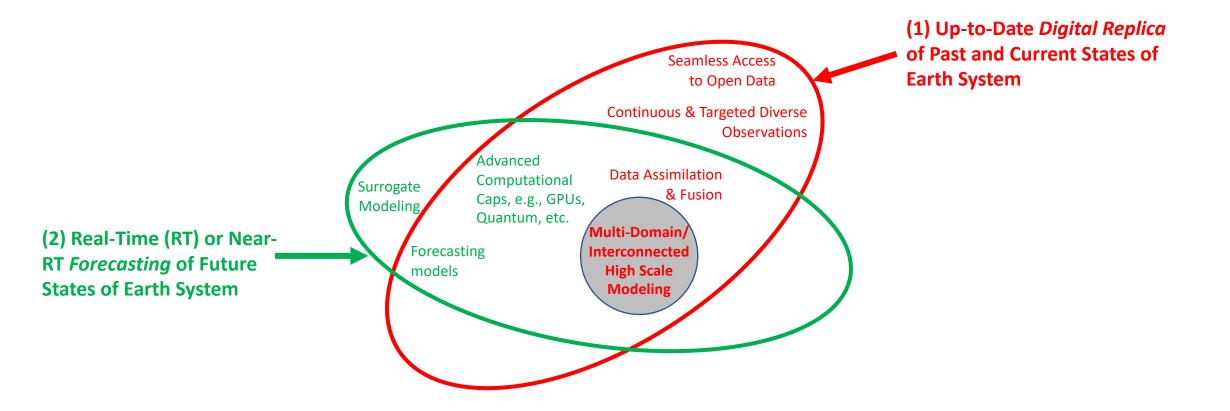




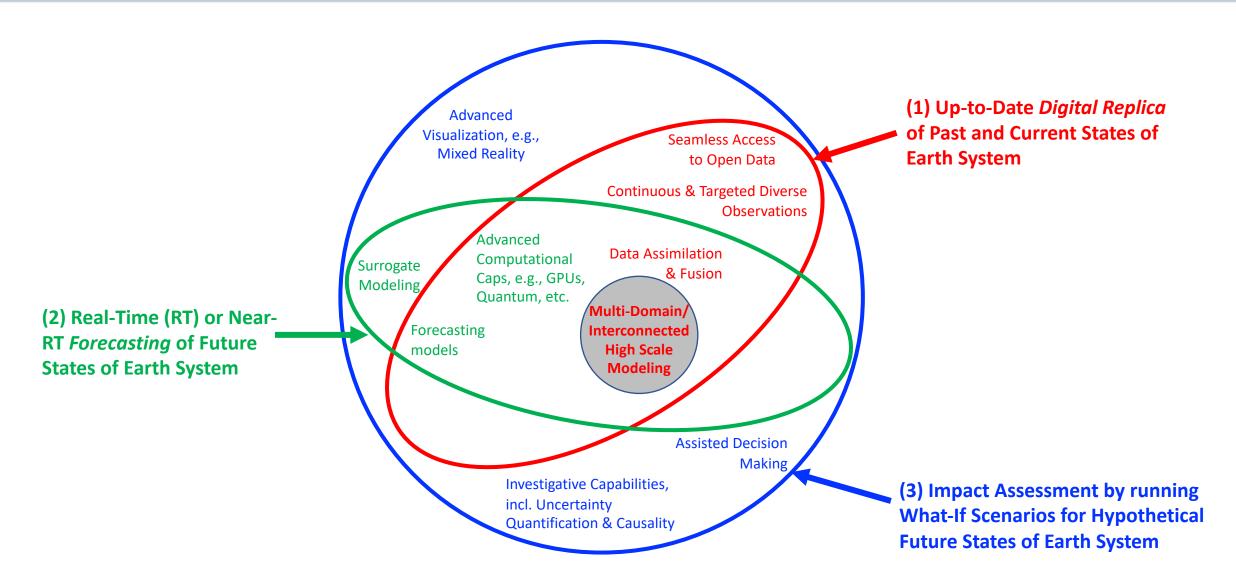






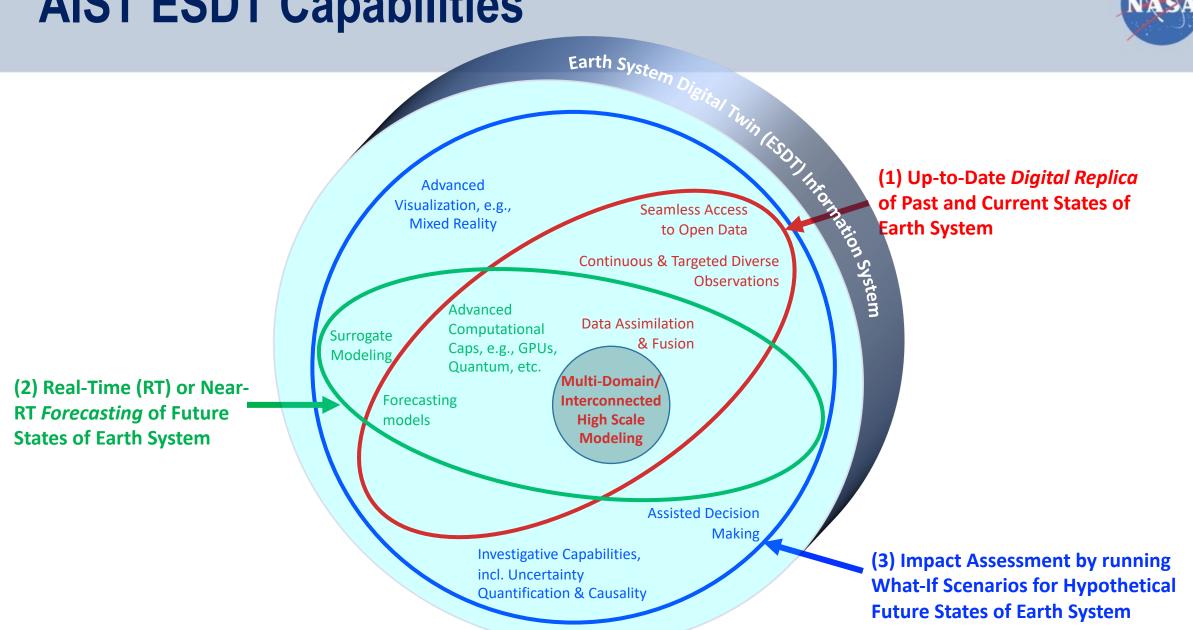




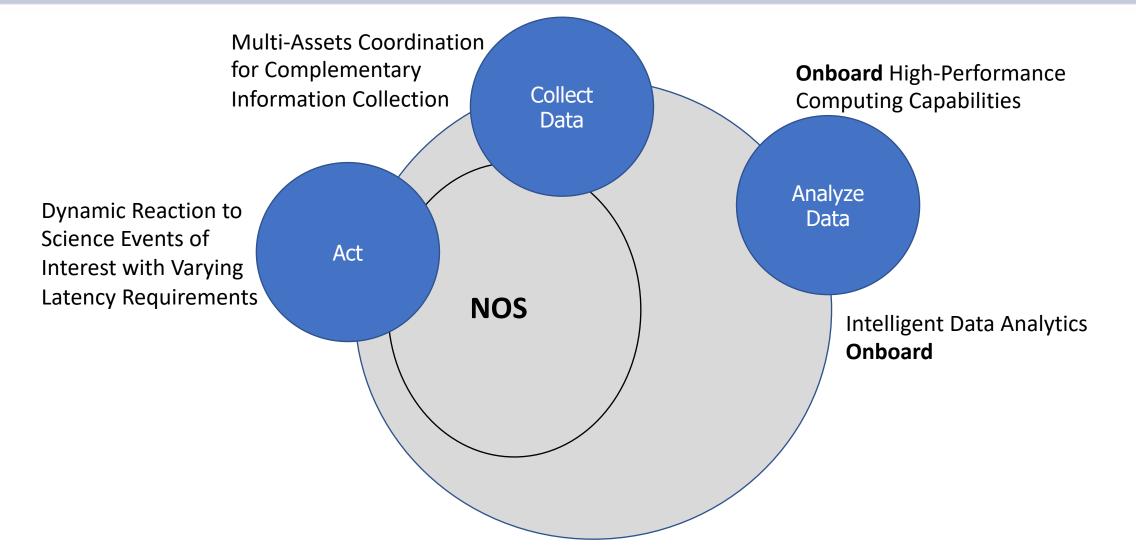


AIST ESDT Capabilities





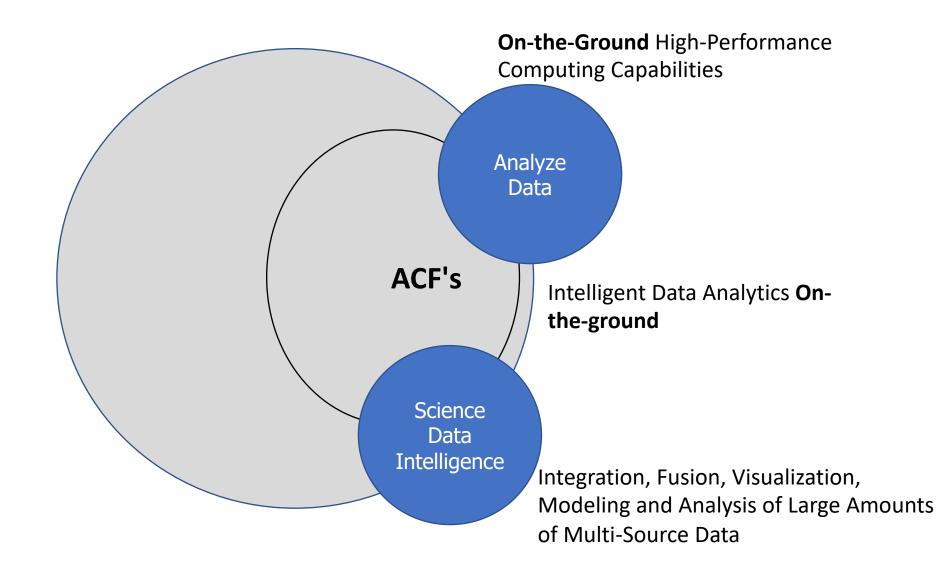
ESDT = New AIST-21 Thrust Continuous Integration of NOS and ACF Technologies



NASA

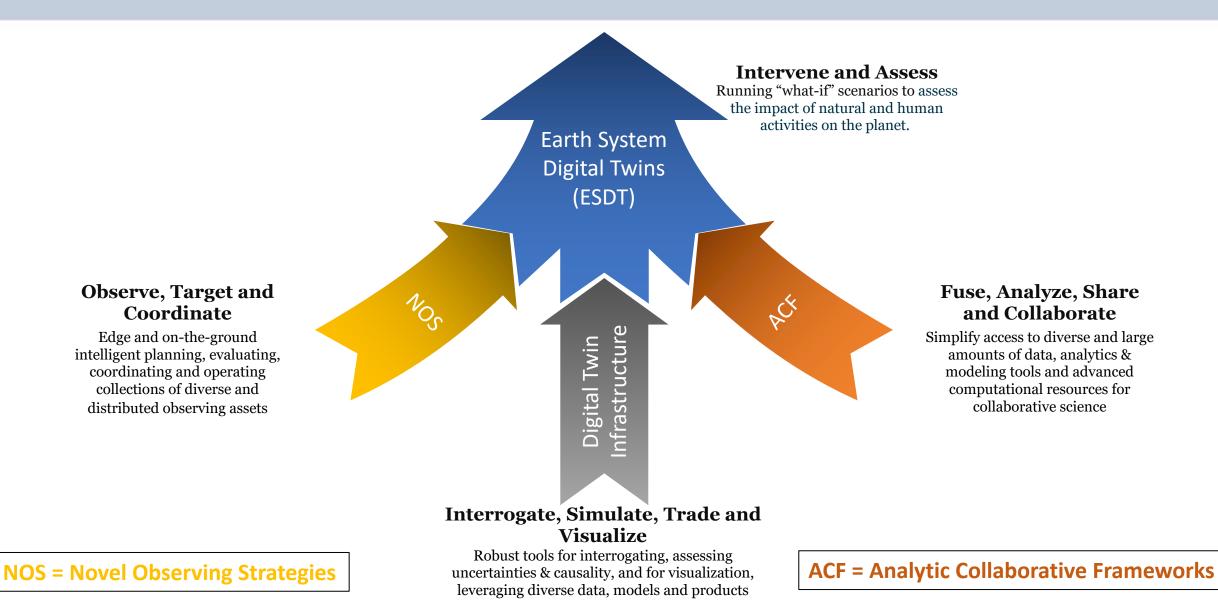
ESDT = New AIST-21 Thrust Continuous Integration of NOS and ACF Technologies



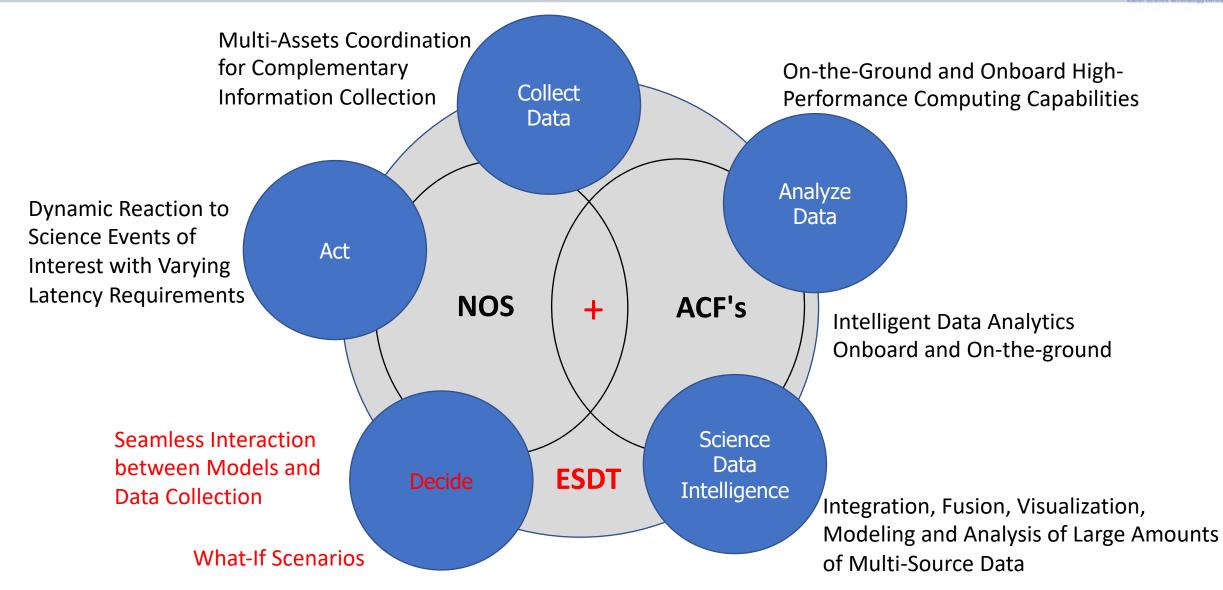


ESDT one of 3 AIST Thrusts





ESDT = New AIST-21 Thrust Continuous Integration of NOS and ACF Technologies





ESDT Technologies Requested in AIST-21



- Technologies for agile interaction and interoperability between measurement acquisition (NOS or NOS-like) and science investigations (ACF or ACF-like)
- Frameworks that enable data ingest from multiple, integrated models, and/or moving from mono-discipline to multi-discipline inter-related systems
- Leveraging of Model-Based System Engineering (MBSE) frameworks for the development and sustainment of Earth Systems Digital Twins, especially based on the integration and coordination of NOS and ACF systems
- Digital Thread developments to link all digital twin capabilities (design, performance data, product data, operational status data, event status data), to enable design requirements, records, provenance, and system reorientations to be easily reviewed and address issues within the digital twin system
- Concepts and technologies for developing "federated ESDTs" in which multiple individual ESDTs interact and can be integrated as the layers of broader ESDTs
- Novel AI (not limited to ML) techniques enabling systems to quickly request, integrate, and fuse diverse and timely Earth observations into ESDTs
- Investigative technologies to facilitate "what-if" investigations inherent to ESDT systems, including but not limited to:
 - Multi-scale simulations, statistics, uncertainty quantification, and causality methodologies
 - Computational algorithms and methodologies involving high-end computing, such as GPUs or other hardware systems that will enable running large permutations of what-if scenarios using large amounts of data and high-resolution and high-fidelity models
 - Statistical methodologies that optimize the computational efficiency of such "what-if" investigations
 - Innovative and simple user interfaces and visualization methods based on Augmented Reality (AR) and Mixed Reality (MR) techniques and capable of visualizing complex systems of systems

AIST-21 ACF for ESDT Awards



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



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. |



• Al-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. |



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. |



Earth System Digital Twins ESDT Workshop Washington, DC – October 26-28, 2022

Why a Workshop on Earth Systems Digital Twins (ESDT)?



- Bring together science and technology communities to explore the use and benefits of ESDT and their enabling technologies.
- Develop ESDT reference use cases and corresponding technology needs to guide the development of ESDT technologies that will be needed by NASA Earth Science within the next five to ten years.

WORKSHOP GOALS:

- 1. Identify driving Earth science use cases that will benefit from unique ESDT capabilities
- 2. Identify emerging technologies that will enable such ESDT systems within the next five to ten years
- 3. Understand opportunities and technologies for federating ESDTs and creating more capable systems

Workshop Goals: Some Overarching Questions

NASA

- What are Digital Twins?
- What are the benefits of ESDT to NASA Earth Science?
- What are the main differences between Earth System Models (ESMs) and Earth System Digital Twins (ESDT) (e.g., model resolutions, connection to impact models, overall interactive information system, others)? How can we integrate/coordinate ESMs with ESDT?
- What are the main architecture components of an ESDT? What could various ESDT architectures look like?
- Should we develop a common Digital Twin Engine?
- What are the AIST Technologies that need to be enhanced or developed?
- 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 various data, models, ESDT interoperate? Which basic interfaces and standards will be required?
- How to prioritize "quick wins" (short-term prototypes) and incrementally enhance digital twin investments over time?
- How do we validate ESDT (e.g., using historical data, etc.)? How to quantify uncertainty?
- How will we federate future ESDT?
- Which sustainable digital twin governance model should be adopted to address software configuration changes, security and full life cycle management?

Workshop Agenda https://esdt2022.sched.com/



THURSDAY Oct 27, 2022

| 8a | 8:30 | GATHERING and Welcome |
|-------|-------|---|
| 8:45 | 9:40 | Lightning talks I |
| 9:40 | 10:00 | Science reference scenarios overview |
| 10:00 | 10:30 | BREAK |
| 10:30 | Noon | Breakout Session A: Science Use Cases |
| NOON | 1pm | LUNCH |
| 1:00 | 1:30 | Session A Brief-outs |
| 1:30 | 3:00 | Lightning talks II |
| 3:00 | 3:30 | BREAK |
| 3:30 | 5:00 | Breakout Session B: Technologies and gaps |
| 5:00 | 5:30 | Session B Brief-outs |
| 5:30 | | Adjourn for the day |

FRIDAY Oct 28, 2022 (Half Day)

| 8a | 8:30 | Welcome/plan for the day |
|-------|-------|---|
| 8:30 | 10:00 | Breakout Session C: ESDT Systems Vision & Federation |
| 10:00 | 10:30 | BREAK |
| 10:30 | 11:15 | Session C Brief-outs |
| 11:15 | 12:00 | Wrap up discussion |
| 12:00 | 12:15 | Closing remarks |
| 12:15 | | Lunch |
| 1:15 | | Workshop Ends |

WEDNESDAY October 26, 2022

| 9:30 | 10:00 | Gathering | |
|-------|-------|--|--|
| 10:00 | 10:30 | Welcome Plenary | Susan Shingledecker, ESIP Jacqueline Le Moigne, ESTO AIST |
| 10:30 | 11:30 | Panel: Digital Twins for NASA Earth Science | |
| 11:30 | Noon | Interactive discussion | |
| NOON | 1pm | LUNCH | |
| 1:00 | 1:30 | Earth system digital twins and the European Destination Earth initiative | Peter Bauer, ECMWF |
| 1:30 | 2:30 | Panel: Federating Earth System Digital Twins | |
| 2:30 | 3:00 | BREAK | |
| 3:00 | 3:30 | Global Digital Twin of the Earth System Environment in NOAA - for Monitoring and Prediction | Sid Boukabara, NOAA |
| 3:30 | 4:30 | Panel: Technologies for Earth System Digital Twins | |
| 4:30 | 5:00 | Wrap up discussion | |

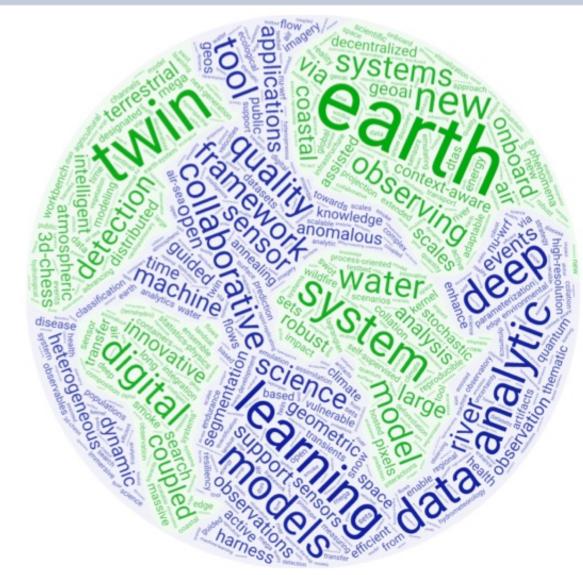
ESDT Science Use Cases/Scenarios



| ESDT Domain | Abstract |
|------------------------------------|--|
| 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 (e.g., urbanization, land use change), to understand coastal ecological changes and impacts to ecosystem services; ocean, land, and atmospheric system to understand feedback processes, such as storm intensification and sea level rise, and their impact on coastal communities and the blue economy; assessing feasibility and impacts of the 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 <i>such as</i> 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 the underlying processes and their relationship to climate and air quality, and 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. |

Questions?





AIST-21 Word Cloud of Selected Proposals Titles



