NASA SBIR 2004 Phase I Solicitation

E3  Advanced Information Systems Technology For Earth Science

The objectives of the Advanced Information System Technology (AIST) Topic are to develop innovative technologies that enable new, or enhance existing, mission and science measurement capabilities for problems closely aligned to the NASA Earth Science Enterprise and, upon completion, provide these capabilities to the broadest set of NASA missions across the agency. The Earth Science Enterprise acquires, processes and delivers very large (gigabyte to terabyte) volumes of remote sensing and related data to public and government entities that apply this information to understand and solve problems in Earth Science. Currently, NASA’s Earth Science Enterprise (ESE) operates 18 orbiting platforms with 80 sensors making scientific measurements of the complex Earth system. Information technology is currently employed throughout ESE’s space and ground systems and the AIST Topic is soliciting technologies that apply to the end-to-end system functions. Target capabilities fall into five major themes: Data Collection and Handling, Transmission and Dissemination, Search, Access, Analysis and Display, and Systems Management.

Results from the AIST Topic will:

- Reduce the risk, cost, size, and development time of NASA’s ESE space-based and ground-based information systems,
- Increase the accessibility and utility of Earth science data,
- Enable new Earth observation measurements and information products, and
- Develop information technologies that enable planetary scale observing systems in support of NASA’s exploration and discovery vision.

Subtopics

E3.01 Automation and Planning
Lead Center: ARC
Participating Center(s): GSFC
The Automation and Planning Subtopic solicits proposals that allow either spacecraft or ground systems to robustly perform complex tasks given high-level goals with minimal human direction. Technology innovations include, but are not limited to: 1) automation and autonomous systems that support high-level command abstraction; 2) efficient and effective techniques for processing large volumes of data (commonly available on the Internet) into useful information; 3) intelligent search of large, distributed data archives, and data discovery through searches of heterogeneous data sets and architecture; and 4) automation of routine, labor intensive tasks that either increase reliability or throughput of current process. Specific areas of interest include the following:

- Search agents that support applications involving the use of NASA data;
- Methods that support the robust production of data products given a set of high-level goals and constraints;
- Autonomous data collection including the coordination of space or airborne platforms while adhering to a set of data collection goals and resource constraints;
- Autonomous data logging devices (software, or hardware and software) supporting a variety of weather and climate sensors, capable of ground-based operation in a wide variety of environmental conditions; such systems would probably be solar powered with accurate time stamping;
- Planning and scheduling methods related to Earth Science Mission objectives;
- System and subsystem health and maintenance, both space- and ground-based;
- Distributed decision making, using multiple agents, and/or mixed autonomous systems;
- Automated software testing;
- Verification and validation of automated systems;
- Automatic software generation and processing algorithms;
- Control of Field Programmable Gate-Arrays (FPGA) to provide real-time products.

E3.02 Distributed Information Systems and Numerical Simulation

Lead Center: ARC
Participating Center(s): GSFC

This subtopic seeks advances in tools, techniques, and technologies for distributed information systems and large-scale numerical simulation. The goal of this work is to create an autonomous information and computing environment that enables NASA scientists to work naturally with distributed teams and resources to dramatically reduce total time-to-solution (i.e., time to discovery, understanding, or prediction), vastly increase the feasible scale and complexity of analysis and data assimilation, and greatly accelerate model advancement cycles. Areas of interest follow below.
Distributed Information Systems

- Core services (autonomous software systems) for automated, scalable, and reliable management of distributed, dynamic, and heterogeneous computing, data, and instrument resources. Services of interest (which may be based on Open Grid Service Infrastructure) include those for authentication and security, resource and service discovery, resource scheduling, event monitoring, uniform access to compute and data resources, and efficient and reliable data transfer.

- Higher level services, including those for job management, resource brokering, workflow management, portlet (i.e., application-specific graphical user interface) building, and collaboration.

- Services for management of distributed, heterogeneous information, including replica management, intuitive interfaces, and instantiation on demand or “virtualized data.” These services would be used, for example, to access and manipulate NASA’s wealth of geospatial and remote sensing data.

- Science portals for cross-disciplinary discovery, understanding, and prediction, encapsulating services for single sign-on access, semantic resource and service discovery, workflow composition and management, remote collaboration, and results analysis and visualization.

- Tools for rapidly porting and hosting science applications in a distributed environment. These applications were written for an integrated, or workstation, environment using standard programming languages or tools such as Matlab, Interactive Data Language (IDL), or Mathematica.

Large-Scale Numerical Simulation

- Tools for automating large-scale modeling, simulation, and analysis, including those for managing computational ensembles, performing model-optimization studies, interactive computational steering, and maintaining progress in long-running computations in spite of unreliable computing, data, and network resources.

- Tools for computer system performance modeling, prediction, and optimization for real applications.

- Techniques and tools for application parallelization and performance analysis.

- Tools for effective load balancing, and high reliability, availability, and serviceability (RAS) in commodity clusters and other large-scale computing systems.

- Novel supercomputing approaches using FPGAs, graphics processors, and other novel architectures and technologies.

E3.03 Geospatial Data Analysis Processing and Visualization Technologies

Lead Center: SSC
Participating Center(s): GSFC
Proposals are sought for the development of advanced technologies in support of scientific, commercial, and educational application of ESE and other remote sensing data. Focus areas are to provide tools for processing, analysis, interpretation, and visualization of remotely sensed data sets. ESE benchmarks practical uses of NASA-sponsored observations from remote sensing systems and predictions from scientific research and modeling. Specific interest exists in the development of technologies contributing to decision support systems, and model development and operation. For more information on decision support models under evaluation, please visit [http://earth.nasa.gov/eseapps/index.html](http://earth.nasa.gov/eseapps/index.html). Areas of specific interest include the following:

- Unique, innovative data reduction, rapid analysis and data exploitation methodologies and algorithms of information from remotely sensed data sets, e.g., automated feature extraction, data mining, etc.;
- Algorithms and approaches to enable the efficient production of data products from active imaging systems, e.g., multipoint data resampling, digital elevation model creation, etc.;
- Data merge and fusion software for efficient production and real-time delivery of digital products of ESE Mission and other remote sensing data sets, e.g., weather observation and land use and land cover data sets;
- Innovative approaches for incorporation of GPS data into *in situ* data collection operations with dynamic links to spatial databases including environmental models;
- Image enhancement algorithms for improving spatial, spectral, and geometric image attributes;
- Innovative approaches for the querying and assimilation of application-specific datasets from disparate and distributed databases from government, academic and commercial sources into a common framework for data analysis;
- Innovative approaches for querying of application-specific data sets from disparate, distributed databases in government, academic, and commercial data warehouses into a common framework for data analysis; and
- Innovative visualization technologies contributing to the analysis of data through the display and visualization of some or all of the above data types including providing the linkages and user interface between the cartographic model and attribute databases.

**E3.04 Data Management and Visualization**

**Lead Center: GSFC**

This subtopic focuses on innovative approaches to managing and visualizing large collections of Earth science data in a highly distributed and networked environment.

- Develop technologies that support long term data management, storage, search, and retrieval of very large, distributed, geospatial Earth science data sets, including the development of object based storage...
devices, file systems that promote long term data maintenance and recovery from user errors, and global compression techniques that optimize data backup operations.

- Develop techniques to manage and locate data in a distributed metadata catalog environment and provide tools to create, use, and then tear down wide area high speed Storage Area Network (SAN) access to remote data sets.

- Develop tools and techniques that enable high bandwidth scientific collaboration in a distributed environment, and allow data viewing, real-time data browse, and general purpose rendering of multivariate geospatial scientific data sets using georectification, data overlays, data reduction, and data encoding across widely differing data types and formats.

- Design and implement 3-D virtual reality environments for scientific data visualization that will enable users to ‘fly’ through the data space to locate specific areas of interest, and make use of novel 3-D presentation techniques which minimize or eliminate the need for special user devices such as goggles or helmets.

E3.05 On-Board Science for Decisions and Actions

Lead Center: ARC

Current sensors can collect more data than is possible to transmit to the ground for analysis. One solution is to incorporate intelligence in the sensor or platform to prioritize or summarize the data and send down high priority or synoptic data. In the future, a sensor-web capability will demand this remote onboard autonomy and intelligence about the kind and content of data being collected to support rapid decision-making and tasking. This subtopic is interested in developing new methods to autonomously understand ES data in support of making rapid decisions and taking actions under two themes:

**Onboard Satellite Data Processing and Intelligent Sensor Control**

Software technologies that support the configuration of sensors, satellites, and sensor webs of space-based resources. Examples include capabilities that allow the reconfiguration or retargeting of sensors in response to user demand or significant events. Also included in this category is onboard processing of sensor data through the use of processing architectures and reconfigurable computing environments, as well as technologies that support or enable the generation of data products for direct distribution to users.

**Onboard Satellite Data Organization, Analysis, and Storage**

Software technologies that support the storage, handling, analysis, and interpretation of data. Examples include innovations in the enhancement, classification, or feature extraction processes. Also included are data mining, intelligent agent applications for tracking data, distributed heterogeneous frameworks (including open system interfaces and protocols), and data and/or metadata structures to support autonomous data handling, as well as compaction (lossless) or compression of data for storage and transmission.