



NASA SBIR 2011 Phase I Solicitation

S6.03 Algorithms and Tools for Science Data Processing, Discovery and Analysis, in State-of-the-Art Data Environments

Lead Center: GSFC

Participating Center(s): ARC, JPL, LaRC, MSFC, SSC

This subtopic seeks technical innovation and unique approaches for the processing, discovery and analysis of data from NASA science missions. Advances in such algorithms will support science data analysis and decision support systems related to current and future missions, and will support mission concepts for:

- All current operational missions (<http://www.nasa.gov/missions/current/index.html>).
- Future Earth Science Decadal Survey missions (<http://science.nasa.gov/earth-science/decadal-surveys>).
- The Landsat Data Continuity Mission (LDCM) (<http://ldcm.nasa.gov/>).
- The Joint Polar Satellite System (JPSS) (<http://www.nesdis.noaa.gov/pdf/jpss.pdf>).
- The Lunar Reconnaissance Orbiter mission (LRO) (<http://lunar.gsfc.nasa.gov/>).
- The Moon Mineralogy Mapper (M3) on Chandrayaan (<http://moonmineralogymapper.jpl.nasa.gov/>).
- The Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) (<http://crism.jhuapl.edu>).
- The Visual Infrared Mapping Spectrometer (VIMS) on Cassini (<http://saturn.jpl.nasa.gov/spacecraft/cassiniorbiterinstruments/instrumentscassinivims/>).
- The James Webb Space Telescope (JWST) (<http://www.jwst.nasa.gov/>).

Research proposed to this subtopic should demonstrate technical feasibility during Phase I, and in partnership with scientists show a path toward a Phase II prototype demonstration, with significant communication with missions and programs to ensure a successful Phase III infusion. It is highly desirable that the proposed projects lead to software that is infused into NASA programs and projects.

In the area of algorithms, innovations are sought in the following areas:

- Optimization of algorithms and computational methods to increase the utility of scientific research data for models, data assimilation, simulations, and visualizations. Success will be measured by both speed improvements and output validation.
- Improvement of data discovery, by identifying data gaps in real-time, and/or derive information through synthesis of data from multiple sources. The ultimate goal is to increase the value of data collected in terms of scientific discovery and application.
- Techniques for data analysis, that focus on data mining, data search, data fusion and data subsetting that scale to extremely large data sets in cloud, large cluster, or distributed computing environments.

In the area of tools, innovations are sought in the following areas:

- Frameworks and related tools such as open source frameworks or framework components that would enable sharing and validation of tools and algorithms.
- Integrated ecosystem of tools for developing and monitoring applications for high performance processing environments, including cloud computing, high performance cluster, and GPU processing environments, that support software development for science data discovery applications, including support for compilation, debugging, and parallelization.
- Integrated tools to collect, analyze, store, and present performance data for cloud computing and large scale cluster environments, including tools to collect data throughput of system hardware and software components such as node and network interconnects (GbE, 10 GbE, and Infiniband), storage area networks, and disk subsystems, and to allow extensibility for new metrics, and verification of the configuration and health of a system.

Tools and products developed under this subtopic may be used for broad public dissemination or within a narrow scientific community. These tools can be plug-ins or enhancements to existing software, on-line data/computing services, or new stand-alone applications or web services, provided that they promote interoperability and use standard protocols, file formats and Application Programming Interfaces (APIs) or prevalent applications. When appropriate, compliance with the FDGC (Federal Geographic Data Committee) and OGC (Open Geospatial Consortium) is recommended.