

STRATEGIC EVOLUTION OF NASA ESE DATA SYSTEMS (SEEDS)

Kenneth R. McDonald ^a, Liping Di ^b, Stephen W. Wharton ^c

^aNASA Goddard Space Flight Center (GSFC), Code 423.0, Greenbelt, MD 20771, USA.
mcdonald@rattler.gsfc.nasa.gov

^bLaboratory for Advanced Information Technologies and Standards (LAITS), George Mason University, 9801
Greenbelt Road, Suite 316-317, Lanham, MD 20706, USA.
lpd@rattler.gsfc.nasa.gov

^cNASA Goddard Space Flight Center (GSFC), Code 902, Greenbelt, MD 20771, USA.
wharton@daac.gsfc.nasa.gov

Commission II, WG II/3

KEY WORDS: Data and Information Systems, Remote Sensing, NASA, SEEDS, EOSDIS, ESE

ABSTRACT:

NASA's Earth Science Enterprise (ESE) is responsible for developing a scientific understanding of the Earth System and its response to natural or human-induced changes to enable improved predictions of climate, weather, and natural hazards. NASA ESE is planning for the evolution of the ESE science data and information systems to meet the research needs in the next ten years. To that end, ESE has chartered a team at the Goddard Space Flight Center to work with the communities that provide and use Earth science data to develop plans for the Strategic Evolution of ESE Data Systems (SEEDS) (previously referred to as NewDISS). SEEDS is currently a formulation activity with the goal of defining the scope and processes that will be employed by the ESE to foster an environment that will enable such evolution to occur. This paper discusses the approach and current status of the SEEDS formulation activities and some of the major challenges they are addressing.

1. BACKGROUND

The ESE has the requirement to assure that all the information, knowledge, and capabilities derived from its research program achieve maximum usefulness to the scientific communities, decision-makers, and the general public. Increasing our knowledge of the Earth system is the goal of NASA ESE's research program, which is complemented by an applications and a technology program (NASA ESE, 2002). In order to achieve the goal, NASA ESE invested billions of dollars in its Earth Observing System (EOS) in the past decade. Several EOS-era satellites have been launched and are collecting a huge amount of remote sensing data (Asrar and Dozier, 1994).

The original concept of the EOS Program was that all of the EOS data would be processed, archived, managed, and distributed by the EOS Data and Information System (EOSDIS) and its Distributed Active Archive Centers (DAACs) (Asrar and Greenstone, 1997). However, during its ten-year development cycle, the supported science community and NASA advisory panels have advocated for a more open, distributed set of data system and service providers (NRC, 1994, 1995, 1996). The rationale for such an opinion was that such an approach would capitalize on the expertise and resources of such an extended community of providers and that rapid innovation would be facilitated in such an environment.

NASA ESE has responded to this advice in multiple ways. First, they initiated a multiyear Earth Science Information Partner (ESIP) Experiment in which formed a federation of competitively selected data centers to explore the issues associated with distributed, heterogeneous data and information system and service providers (ESIP, 2002; NRC, 1998). At the same time, the EOSDIS architecture has evolved to accommodate the generation of data products by external processing systems developed under the direction of the EOS

instrument teams. Concurrently, the ESE chartered a study team called NewDISS to capture and consolidate the input from the community in a series of recommendations (NASA, 2001). An outcome of this study was the initiation of a formulation activity entitled the Strategic Evolution of ESE Data Systems (SEEDS) that is chartered to work with the Earth science user and data provider communities to generate approaches and plans for the future of ESE data and information systems (NASA, 2002).

2. SEEDS FORMULATION

The overall objective of the SEEDS Formulation Team is to recommend a framework and management strategy to enable evolution towards a future network of ESE data systems and providers that:

- Leverages the capabilities and lessons learned from the EOSDIS and other ESE data systems efforts.
- Encourages development and evolution of heterogeneous systems and services.
- Gives systems and service providers appropriate local control over data system design, implementation, and operation.
- Leverages competition, technology infusion, and reuse to improve system effectiveness.
- Ensures that products and services meet norms for utility and accessibility.
- Ensures that systems and products meet NASA security and survivability requirements.
- Monitors collective performance in meeting the Enterprise objectives and goals.

- Maintains sufficient organizational structure to allow effective resource management and implementation for NASA to carry out its science mission.

In order to achieve the objective, the Formulation Team has set up study teams to investigate specific subjects of concern to SEEDS and make recommendations. Currently there are seven study teams with members from government agencies, universities, and industry. The seven study teams and their tasks are the following:

- Standards for Near-Term Missions. The tasks of the team include considering ESE's near-term systematic measurement missions; recommending science data, metadata, and interoperability standards for applications; and incorporating advice and experience of mission science community in making recommendations.
- Levels of Service, Benchmarks and Cost Estimation. The study team will work with the science and applications communities to develop the minimum and recommended levels of service for core data sets and services required from ESE data management service providers. It will determine, from benchmarking, what data management services should cost, and develop a capability to perform end-to-end cost estimates for ESE data management services.
- Standards and Interfaces Processes. This study team will define a process for SEEDS to develop, adopt, evolve, and maintain standards and interfaces for data and information systems and services across the Earth Science Enterprise. The process should capitalize on the methods and experience of existing relevant data systems standards bodies (e.g. ISO, OGC) and NASA programs (e.g. EOSDIS, ESIP Federation).
- Data Life Cycle and Long-Term Archive: The tasks of this study team are to ensure safe handling of SEEDS-era data products as they migrate from data providers to active archive and long-term archive (LTA), even as numerous individuals and institutions take responsibility for the product during its life-cycle.
- Reference Architecture and Reuse Assessment: This study team tries to answer the following questions in its study. Should NASA/SEEDS invest in a software & component reuse effort? Should NASA/SEEDS invest in developing a reference architecture? If NASA/SEEDS should invest in either of these efforts (i.e. reuse and/or a reference architecture), what is the best method to assure effective and accountable community involvement, what is the best technical approach? Criteria for judging in priority order are cost savings over time; increased flexibility/responsiveness to new missions, science, and applications use of NASA data; and increase in effective and accountable community participation.
- Metrics Planning and Reporting: This study team will define appropriate metrics and reporting requirements for the participants in ESE Data Management Activities and demonstrate that proposed SEEDS organization structure can provide adequate accountability.
- Technology Needs and Infusion Plans: This study team will determine processes by which technology needs are identified and technology investments are infused into the evolving NewDISS. The study team will recommend ways for SEEDS to leverage the processes of NASA Earth Science Technology Office (ESTO)'s Advanced Information System Technology (AIST) Program ESTO,

involve ESE user community, and designate roles of ESTO AIST and SEEDS with regard to prototyping needs.

3. FORMULATION GUIDELINES AND ASSUMPTIONS

While each of the respective study tasks has their own focus topic, there are some general guidelines and assumptions that apply to the entire effort. Over the past decade, NASA ESE has made a substantial investment in the development of data and information systems. This is most evident in the EOSDIS Core System (ECS) but also includes unique components developed by the DAACs, the data processing systems developed by the instrument teams and a variety of other capabilities that are still actively used and maintained as a result of heritage missions and initiatives. The ESE is not intending nor does it have resources to undertake a wholesale replacement of these capabilities. Rather, it envisions the evolution of existing systems and services to maximize the return on those previous investments. Each of the studies is being performed within the context of the evolutionary nature of the overall effort.

A second guideline is for the formulation effort to be outward looking and is derived from the actual or perceived insular nature of previous NASA approaches to the development of data and information systems. Wherever appropriate, the SEEDS studies are addressing as wide a range of related activities as possible, within government, industry and academia in the U.S. and abroad. By taking this broad view, it is hoped that the recommendations that emerge from the SEEDS studies will be able to capitalize on this extended experience base and the best practices and latest technical approaches available.

One additional guideline that is directing the SEEDS formulation activity is that the community of users and providers of Earth science data and services should be actively and substantively involved in the definition and execution of the SEEDS processes, practices and policies. A number of the study teams have representatives of these communities either as members or consultants and all of the teams individually and collectively are making every opportunity to interact with the community, through interviews, meetings and workshops.

4. SAMPLE STUDY TASK – STANDARDS AND INTERFACES PROCESSES

A brief overview of the activities and current results of the Standards and Interfaces Processes study task illustrates the general approach of the SEEDS Formulation Team. The team performing this study has individuals with experience from previous NASA data systems development projects but also has representation from other federal agencies and academia. The approach has been to begin by assembling an extensive collection of information on the standards and standard interfaces activities of heritage data system development projects within NASA and other agencies and on the standards and standard processes employed by relevant formal standards bodies and organizations. These organizations include the ISO Technical Committee on Geographic Information/Geomatics (TC 211), the Open GIS Consortium (OGC), the World Wide Web Consortium (W3C), the Consultative Committee for Space Data Systems (CCSDS), the U.S. Federal Geographic Data Committee (FGDC) and the Internet Engineering Task Force (IETF). All of this information has been collected into a SEEDS Standards and Interfaces Survey Document.

The survey document has provided the foundation for the team to identify and characterize a potential approach and process options that can be used to support the ESE in making decisions on data and information standards. The overall process recommends that the enterprise consider the use of existing standards before initiating a new standard development activity. This use could be to adopt the existing standard as is or to constrain or extend it to meet the specific needs of the enterprise. If a development activity is deemed to be necessary, the process should consider alternate approaches to develop a standard (request for proposal (RFP), testbed/specification with external standards organization, collaboration, etc.). For either a new or an existing standard or standard interface, an approval process is required with appropriate representation from NASA officials, the community of data providers and users and other stakeholders. An important feature of the process of several standards organizations is that software and tools that implement and or support the standard must be available prior to its approval. This has been incorporated into the study team's proposed approval process.

The survey results and the process options that were derived from their analysis are being discussed with the ESE management and the community of providers, users and stakeholders through a series of meetings and workshops. The needs for associated activities such as user support, tool development, outreach and training are also topics that are being considered. The discussions also are addressing participation and the respective roles and responsibilities of NASA and the community of data providers and users in both the processes and the associated activities. A set of recommendations that incorporates the results of these interactions will be the final product of this study.

5. CONCLUSIONS AND FUTURE PLANS

While the final plans for SEEDS will depend on the results of both the studies, community interactions and decisions by NASA management, the current view is that SEEDS will not be responsible for implementing or operating any data systems. The NASA programs and projects selected to develop and operate new ESE flight missions and the research projects of the NASA-funded scientific communities will hold that responsibility. The results of these development efforts will be the more open, distributed and heterogeneous data systems that were envisioned in the earlier NewDISS studies.

The collection of study results produced by the SEEDS Formulation Team will be used to develop management guidelines, a standards framework and decision support capabilities for overall ESE data systems planning and coordination. These functions will ensure that the aggregation of the mission and science data systems can work in concert to meet the overall objectives of the enterprise and maximize the return on its investment in these efforts.

ACKNOWLEDGEMENTS

This paper describes work that is being performed by the entire SEEDS Formulation Team whose members are Catherine Corlan, Kathy Fontaine, Vanessa Griffin, Gail McConaughy, Ken McDonald, Karen Moe, H. Ramapriyan, Matt Schwaller, Richard Ullman and Stephen Wharton, all from NASA Goddard Space Flight Center.

REFERENCES

- Asrar, G., and Dozier, J., 1994. *EOS: Science Strategy for the Earth Observing System*. American Institute of Physics Press, Woodbury, New York 11797.
- Asrar, G., and Greenstone, R., 1997. *MTPE EOS Reference Handbook*. NASA Goddard Space Flight Center, Greenbelt, MD 20771.
- ESIP, 2002. The ESIP Homepage, <http://www.esipfed.org>.
- NASA, 2001. NewDISS, a 6-to 10-year Approach to Data Systems and Services for NASA's Earth Science Enterprise. http://lennier.gsfc.nasa.gov/seeds/ND_Reprt.pdf.
- NASA 2002. SEEDS: Strategic Evolution of ESE Data Systems. <http://lennier.gsfc.nasa.gov/seeds/index.html>.
- NASA ESE, 2002. Exploring Our Home Planet—Earth Science Enterprise Strategic Plan. http://www.earth.nasa.gov/visions/stratplan/ese_strategic_plan.pdf.
- NRC 1994. *Panel to Review EOSDIS Plans: Final Report*. National Academy Press, Washington, D.C., 88pp.
- NRC 1995. *A Review of the U.S. Global Change Research Program and NASA's Mission to Planet Earth/Earth Observing System*. National Academy Press, Washington, D. C., 96 pp.
- NRC, 1996. *Letter to Robert W. Corell with final conclusions on the U.S. Global Change Research Progra, and NASA's Mission to Planet Earth/Earth Observing System*. Washington, D. C., 18 pp.
- NRC, 1998. *Toward an Earth Science Enterprise Federation: Results from a Workshop*. Washington, D. C., 51 pp.

