NASA’s Strategy for Enabling the Discovery, Access, and Use of Earth Science Data

Francis Lindsay, PhD
Martha Maiden
Science Mission Directorate
NASA Headquarters

IEEE International Geoscience and Remote Sensing Symposium 2006,
Denver CO
August 3, 2006
Data Systems to Support Science Communities

To achieve the widest possible support for these disparate communities NASA data systems must be capable of:

1) providing a stable and robust data system infrastructure (networks, computing facilities, data processing, algorithm development, data storage and data dissemination), and ...

2) be readily adaptable and flexible to needs and requirements that are science domain specific.
Improving NASA’s Earth Science Data Systems

* Improve the ease of data handling -- satellite observation/model data remains voluminous.
* _Build Once, Use Many Times._ Incorporate any new techniques as services that can be employed in multiple instances within NASA science systems.
* Leverage NASA existing data system capabilities while including pertinent community developments.
NASA’s Core & Community Data System Views

* ‘Core’ data system elements reflect NASA’s responsibility for managing Earth science satellite mission data characterized by the continuity of research, access, and usability.

* The core comprises all the hardware, software, physical infrastructure, and intellectual capital NASA recognizes as necessary for performing its tasks in Earth science data system management.

* The ‘community’ elements are those pieces or capabilities developed and deployed largely outside the NASA core elements and are characterized by their ‘evolvability’ and innovation.
<table>
<thead>
<tr>
<th><strong>CORE</strong></th>
<th><strong>COMMUNITY</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Projects Subject to ‘Senior Review’</td>
<td>Projects Competitively Selected</td>
</tr>
<tr>
<td>Substantive NASA Oversight</td>
<td>‘Light Touch’ Oversight w/ Significant Community Involvement</td>
</tr>
<tr>
<td>Tight Integration of Data System Tools, Services and Functions</td>
<td>Community-based Tools and Services Loosely-Coupled</td>
</tr>
<tr>
<td>Employ Well Established Information Technologies</td>
<td>Employ ‘Edgy’ or Emerging Technologies</td>
</tr>
</tbody>
</table>
Data Discovery & Access (core)

* Global Change Master Directory (GCMD)
* Data Pools (On-line)
* ECHO (middleware to EOS data)
* Evolution of EOSDIS Elements (ongoing study)
Evolution of EOSDIS

NASA’s research communities have access to all EOS data through services at least as rich as any contemporary science information system, for example:

- The physical location of data storage is irrelevant
- Data access latency is no longer an impediment
- Finding data is based on common search engines (e.g., Google 2015)
- Services are primarily invoked by machine-to-machine interfaces
- Multiple data and metadata streams can be seamlessly combined
- Custom processing (e.g., subsetting, averaging, reprojection) provides only the data needed, the way they are needed
- Best practice standard interfaces and protocols are universally employed
- Open interfaces and best practice standard protocols are universally employed

The research and value-added provider communities use EOS data interoperably with any other relevant data sources and systems.

The EOS archive holdings are regularly peer reviewed for scientific merit.
Evolution of EOSDIS

Mechanisms to collect and preserve the pedigree of derived data products are readily available.

Processing and data are mobile: processing can be moved to data and/or data can be moved to processing.

NASA data systems have evolved into components that allow fine-grained control over cost drivers.

Expert knowledge is readily accessible to enable researchers to understand and use the data.

Community feedback directly to those responsible for a given system element is readily available.
NASA’s DATA SYSTEM COMMUNITIES
NASA Funding Opportunities (ROSES)  
(incorporate community innovation)

✴ REASoN: Principle solicitation to add needed Earth science research and/or science data products.

✴ ESTO AIST: Develop and mature improved information system technology - risk reduction.

✴ Space Science AISR: Space Science data system testbeds and prototypes.

✴ DECISION: Applied sciences solicitation for improving and creating ‘solutions networks’ and support DSS tools.

✴ ACCESS: Leverage existing capabilities to improve NASA’s Earth science data systems.
NASA’s Data System ACCESS Program
(Advancing Collaborative Connections for Earth System Science)

✴ ... to enhance and improve existing pieces of the data and information systems infrastructure that support NASA’s Earth science research communities.

✴ ... to strengthen the interconnectedness of NASA’s heterogeneous data system components by leveraging proven information technologies, protocols, and practices.

✴ ... to augment and improve current data system capabilities for the benefit of NASA’s Earth science communities.
ACCESS Goals (ROSES 2005)

- Focus of the ACCESS 2005 Announcement:
  1. Enhance or create tools and services to support evolution to science measurement processing systems.
  2. Tools and services to support and/or enhance NASA’s Science Focus Area (SFA) communities.

- Cooperative Agreement funding for 2 year award periods with a possibility of a 3rd year extension.

- Currently 16 ACCESS cooperative agreement-based projects.
Core/Community Implementations
(more than a concept)

**ECHO** *(EOS Clearinghouse)*
- Middleware to improve the discovery and access of NASA data. Users build custom clients for their community.
- Developed as a part of NASA’s core data systems but fully adaptable to the needs/requirements of science communities.

**OPeNDAP** *(Open-source Project for a Network Data Access Protocol)*
- OPeNDAP provides software which makes local data accessible to remote locations regardless of local storage format. Provides tools for transforming existing applications into OPeNDAP clients.
- Broad use among community members and used as data servers from NASA’s core systems.
Earth Science Data System Working Groups (recommendation from SEEDS study)

- **Metrics Planning and Reporting** - develop flexible metrics collection that reflects progress of data system in supporting Science Mission Directorate programs.

- **Standards Processes** - catalyst for the adoption of community-based standards (e.g., oceans, precipitation groups)

- **Technology Infusion** - characterizing the key features of a future data system capabilities vision

- **Software Reuse** - provides recommendations for NASA HQ consideration to encourage reuse
Data System Software Policy

* NASA solicits and selects the best solutions to its information technology needs from industry and academia through open solicitation and peer review process.

* NASA emphasizes the use of COTS and the re-use of existing innovative solutions, and the cost efficiency associated with them, as important criteria in its selection and awarding its grants and contracts.

* For purposes of peer review evaluations, the use of COTS software, software developed through open source, ‘freeware’, or government-sponsored projects to improve, modernize, or expand NASA’s data and information systems is of equal consideration.

* The criteria for evaluation of the overall software or tool proposed is the - effective total life-cycle cost of development and maintenance - design is appropriate to usage for the targeted users - design utilizes the mix of COTS, open source, free-ware, and new development
Conclusions

- NASA’s Earth science data systems continue to improve making discovery and access to data and services easier.
- Many of the current challenges are being overcome by strategic use emerging information technologies.
- Other improvements are being made fostered by the valuable inputs from NASA’s large science and technology user communities.
- The practices in place that denote our ‘core’ and ‘community’ data system elements allows NASA to remain nimble in evolving its data systems while leveraging the best practices in use by our science communities.
- This structure will ensure that NASA is able to meet its challenging goal of providing critical Earth science data to the widest possible science and user communities.
thanks!