CLASS Status and Evolution

John J Bates
NOAA's National Climatic Data Center

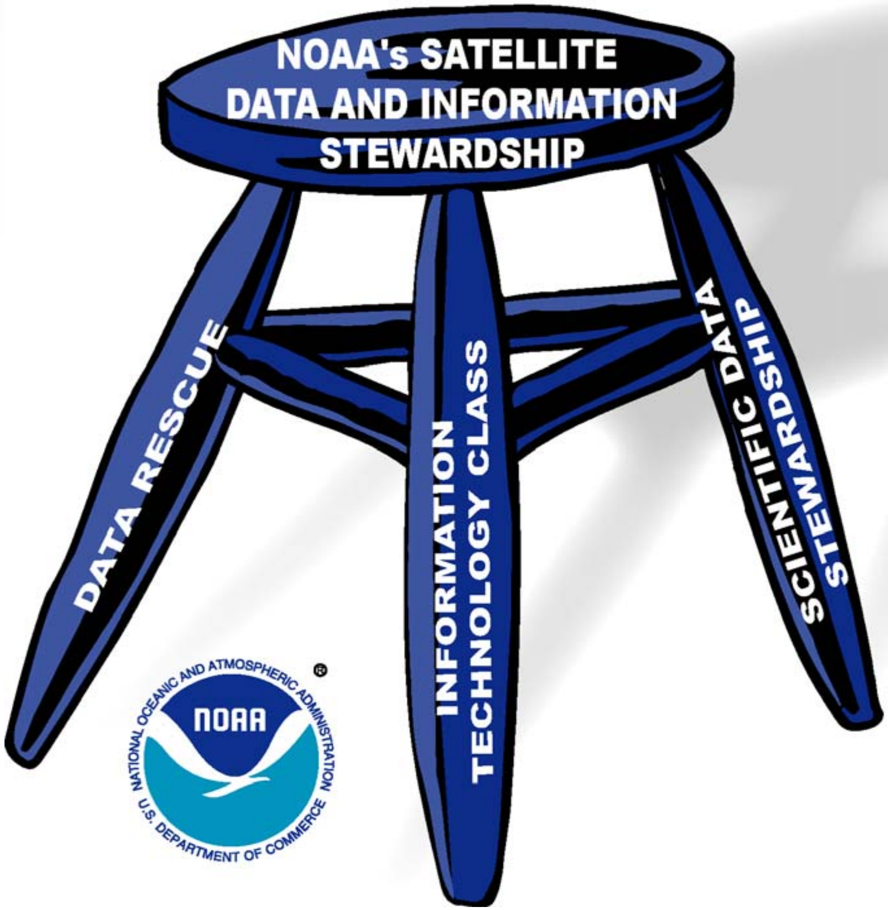
Outline

- History
- Current Status of CLASS
- Challenges - CLASS and NOAA's Integrated Data Environment
- Evolution of CLASS
- Opportunities for Inter-Agency and International Partnerships

CLASS history

- CLASS started as the satellite active archive in the early 1990s, largely to serve POES data and products
- GOES active archive began in the late 1990s
- Need to integrate POES and GOES and to accommodate planned large-array satellite and radar led to CLASS plans
- CLASS is now further evolving as a NOAA Enterprise solution for archiving the suite of NOAA observations in an integrated data environment

NOAA's National Data Centers -- Environmental Data Stewards



Data Management Scientific Data Stewardship is ownership, knowledge, utilization, and application of the data

CLASS is the Information Technology infrastructure (hardware and software environment, and tools) underpinning SDS

Data Rescue preserves and makes available historical data sets from obsolete media

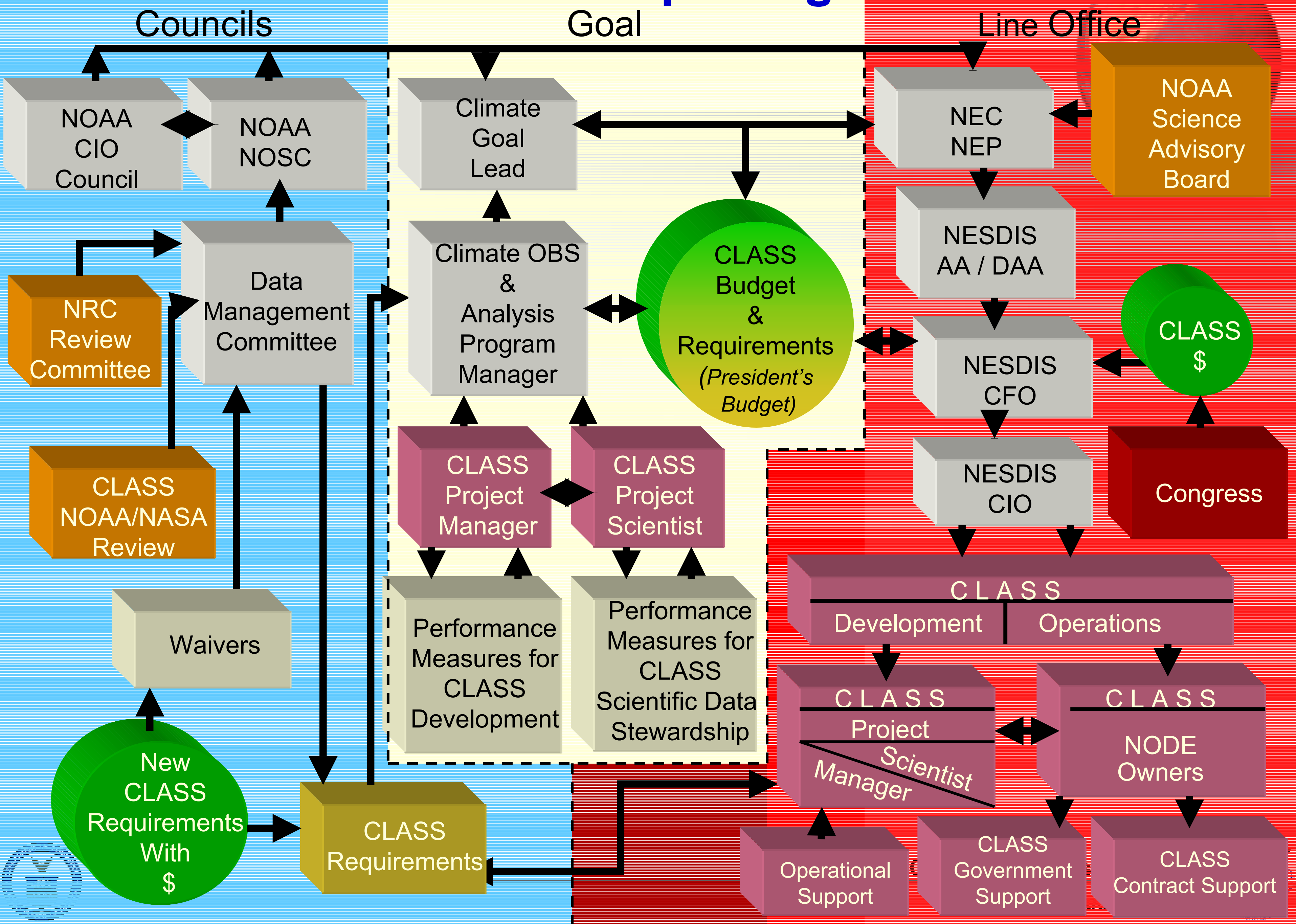
Current Status of CLASS

- All NOAA POES data (since 1978) and selected products are available (VTPR 1973-78 currently being rescued)
- All GOES since 1994 (i.e., GOES-Next) data are available (GOES data from 1979-1994 currently being rescued)
- CLASS will host Initial Joint Polar System (IJPS – Eumetsat Metop data)
- Initial capability for NASA EOS MODIS and Joint NPP developed
- Rescue of initial DMSP SSMI data under way (1978-1995 prior to NOAA SSMI operations)

CLASS Governance

- CLASS governance is undergoing 'growing pains'
- CLASS has grown to be a NOAA major program and is thus reviewed by the NOAA Observing System Council (NOSC)
- CLASS is also matrix managed with planning through the NOAA Climate Goal and execution through the NESDIS Line Office
- This has advantages and disadvantages…but is always a challenge…

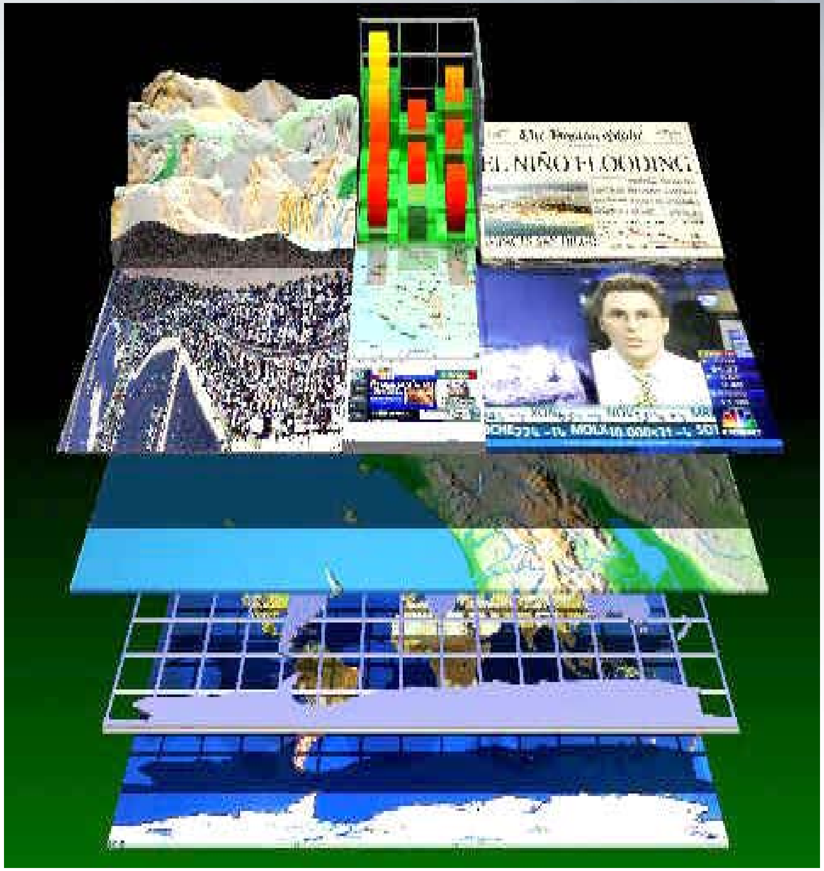
NOAA CLASS Reporting Structure



Data Management Challenges

Urgent need for improved cyber-infrastructure in NOAA, with CLASS leading the way

- New systems will lead to 100-fold increase in data volume
- Increasing need for interdisciplinary use of data
- Current systems already face challenges



Important societal issues require data from many observing systems

	Discipline Specific View	Whole System View
Atmospheric Observations		
Land Surface Observation		
Ocean Observations		
Space Observations		
Data Systems	**Current systems are program specific, focused, individually efficient.** **But incompatible, not integrated, isolated from one another and from wider environmental community**	**Coordinated, efficient, integrated, interoperable**

How is NOAA Responding?

- NOAA's GEO Integrated Data Environment
  - The umbrella for all of NOAA
- Active participation in USGEO & GEOSS
- Leadership in IOOS DMAC
- Scientific Data Stewardship Program
  - Climate data records from all data sources
- CLASS
  - Vehicle for data center's archive consolidation under GEO-IDE
- Enhanced access to data and products
  - Model data - NOMADS
  - Doppler Radar data
  - GIS applications/interfaces
  - New and legacy in situ data sets (e.g., transportation)

NOAA’s GEO Integrated Data Environment & CLASS

- **Scope** – NOAA-wide architecture development to integrate legacy systems and guide development of future NOAA environmental data management systems
- **Vision** – NOAA’s GEO-IDE is envisioned as a “system of systems” – a framework that provides effective and efficient integration of NOAA’s many quasi-independent systems
- **Foundation** – built upon agreed standards, principles and guidelines
- **Approach** – evolution of existing systems into a services-oriented architecture
- **Result** – a single system of systems (user perspective) that is used to access the data sets needed to address significant societal questions

GEO-IDE & CLASS Motivation

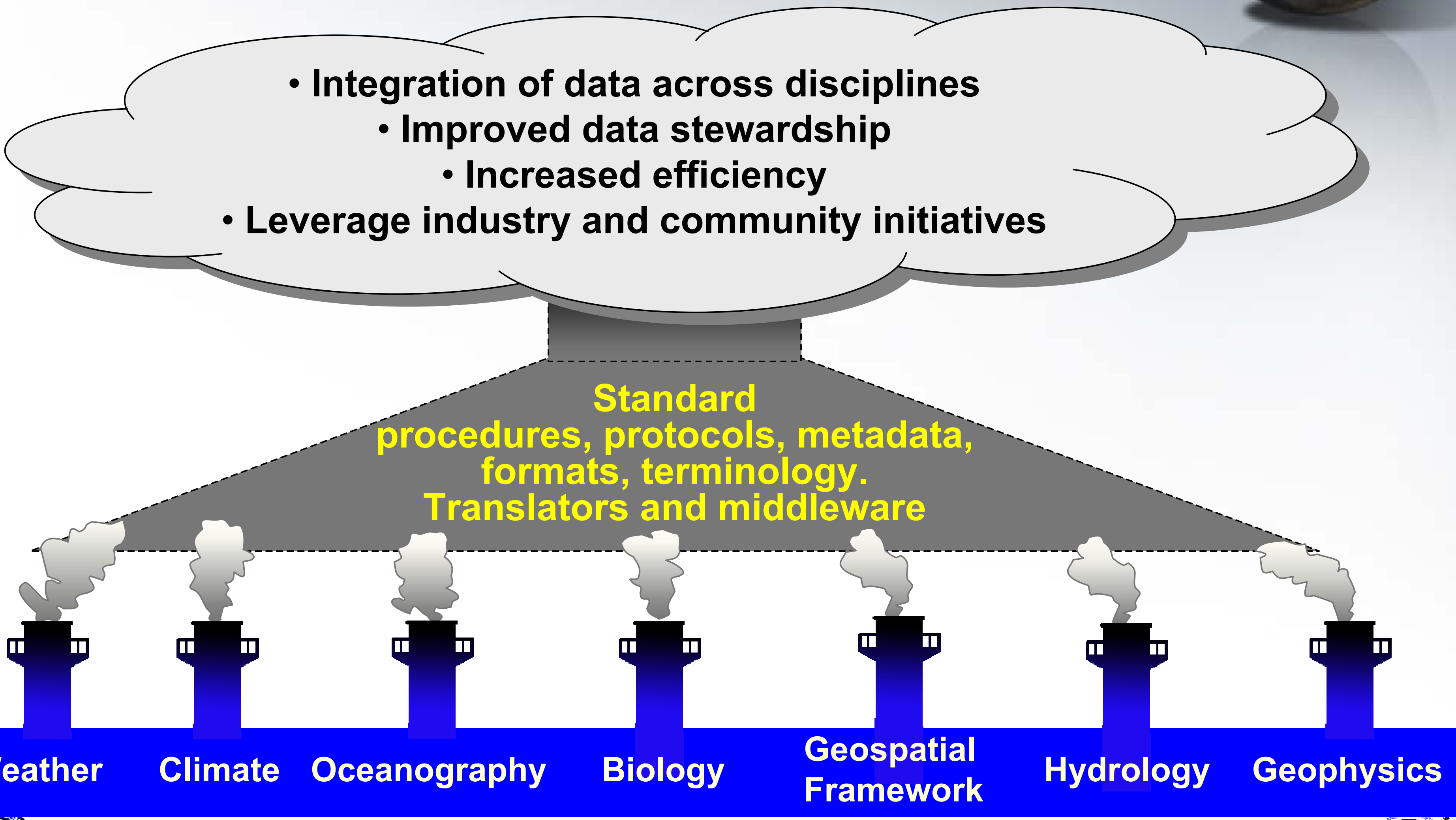
Correctly answering societal questions… What is the likelihood of ?... What is the impact of ?... What are the consequences of ?

- **Incompatible syntax (formats) and semantics (terminology) among science disciplines within NOAA. Thousands exist. Several examples:**
  - **Naming standards – Surface Air Temperature**
    - **Meteorology (WMO) named "Temperature/dry bulb temperature**
    - **Meteorology (air pollution) named "Boundary layer temperature"**
    - **Oceanography named "Air Temperature"**
  - **Location standards (latitude, longitude, elevation)**
    - **Lat/Lon can be degrees/minutes/seconds or degrees to tenths and hundredths**
    - **Latitude E/W, 0-180 positive and negative, or 0-360 running east or west**
    - **Z used to designate elevation in both atmosphere and ocean but positive is up in the atmosphere and down in the ocean**
  - **Formats (>50 formats used within NOAA; translators and standards needed)**
    - **GRIB, NetCDF, HDF and others used for gridded data**
    - **BUFR, NetCDF, and <u>many</u> others used for observations**

- **Potential for no answer or the "wrong" answer to important societal questions due to separate NOAA data management systems**

GEO-IDE & CLASS

Bridging the gaps between stove-pipe systems



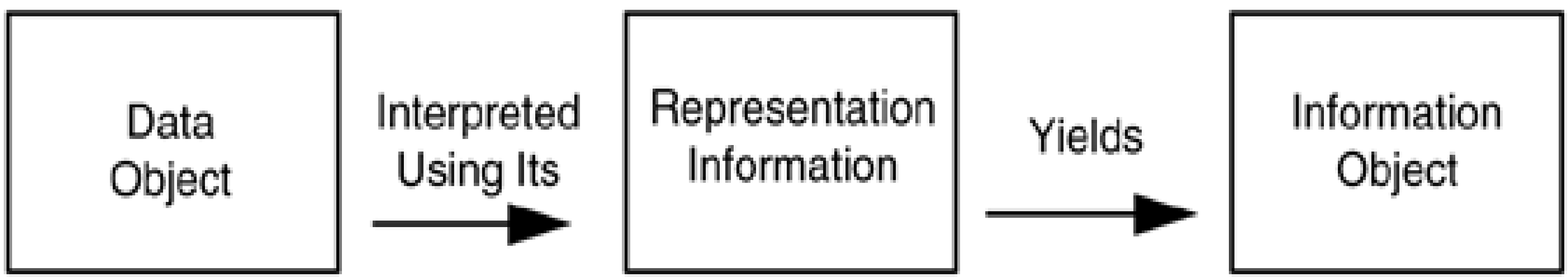
GEO-IDE & CLASS Rely Upon Standards – Existing Standards

Adopt, Adapt, Build only when necessary

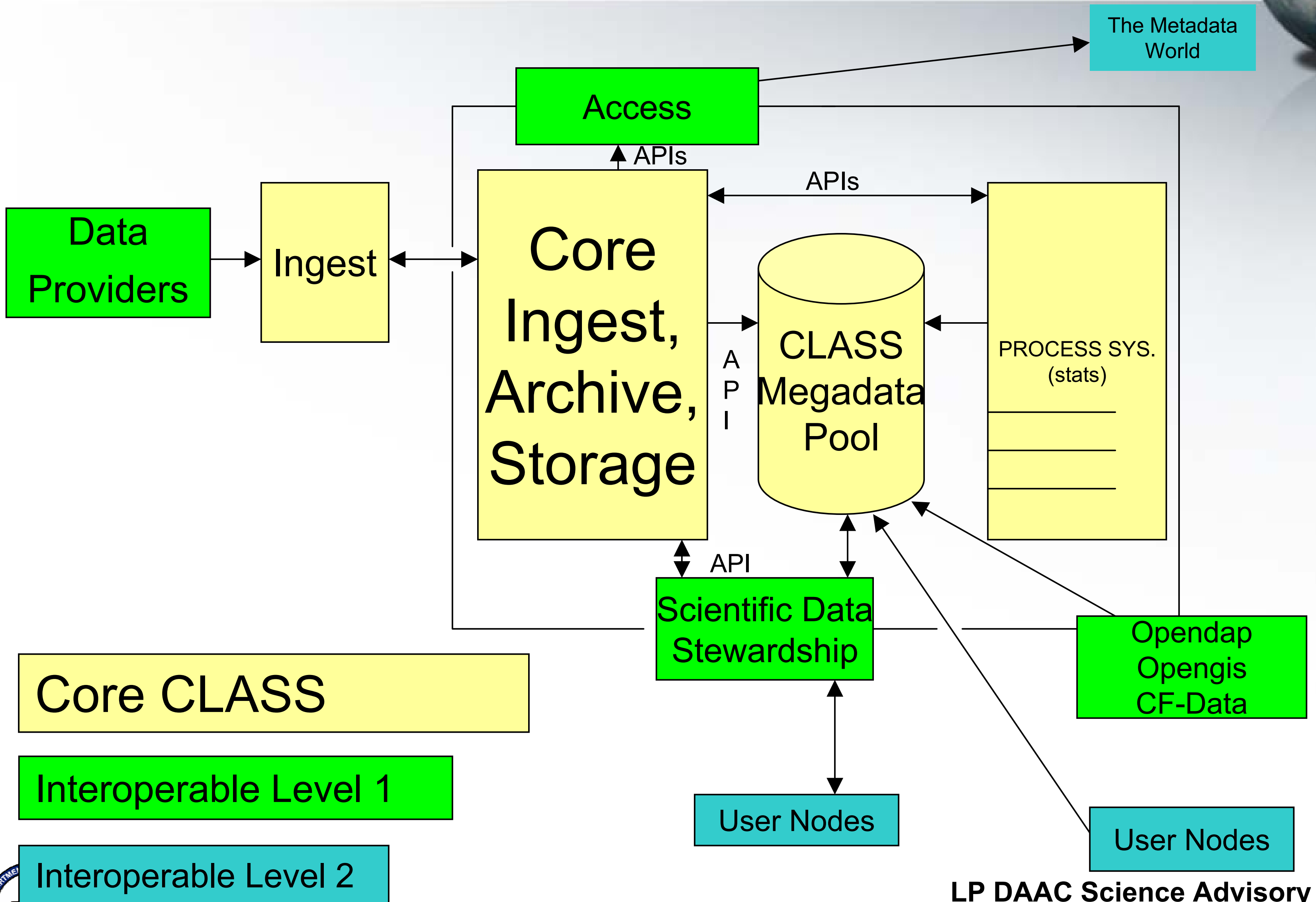
- Archive system standards
  - ISO Open Archival Information Systems
- Metadata standards
  - FGDC and ISO 19115 w/ remote sensing extensions,
- Format standards
  - XML Schemas, Spatial Databases (SQL), data formats (WMO, NetCDF, HDF, etc.)
- Standard names and terminology
- Open Geospatial Consortium (OGC) standards:
  - Features, Coverage (data), Geographic Markup Language
- Web Services Standards (World Wide Web Consortium)
- Process for evaluating/adopting standards within NOAA

Evolution of CLASS

- CLASS currently only takes responsibility for data objects
- To ensure information preservation, we must have a steward for the representation information
- Preserving information is a people business that connects data providers with users
- NOAA’s National Data Centers (NNDCs) are in an optimal position to do that



CLASS Long-Term Architecture Evolution



Recent Developments in the Earth Sciences Data Management Community

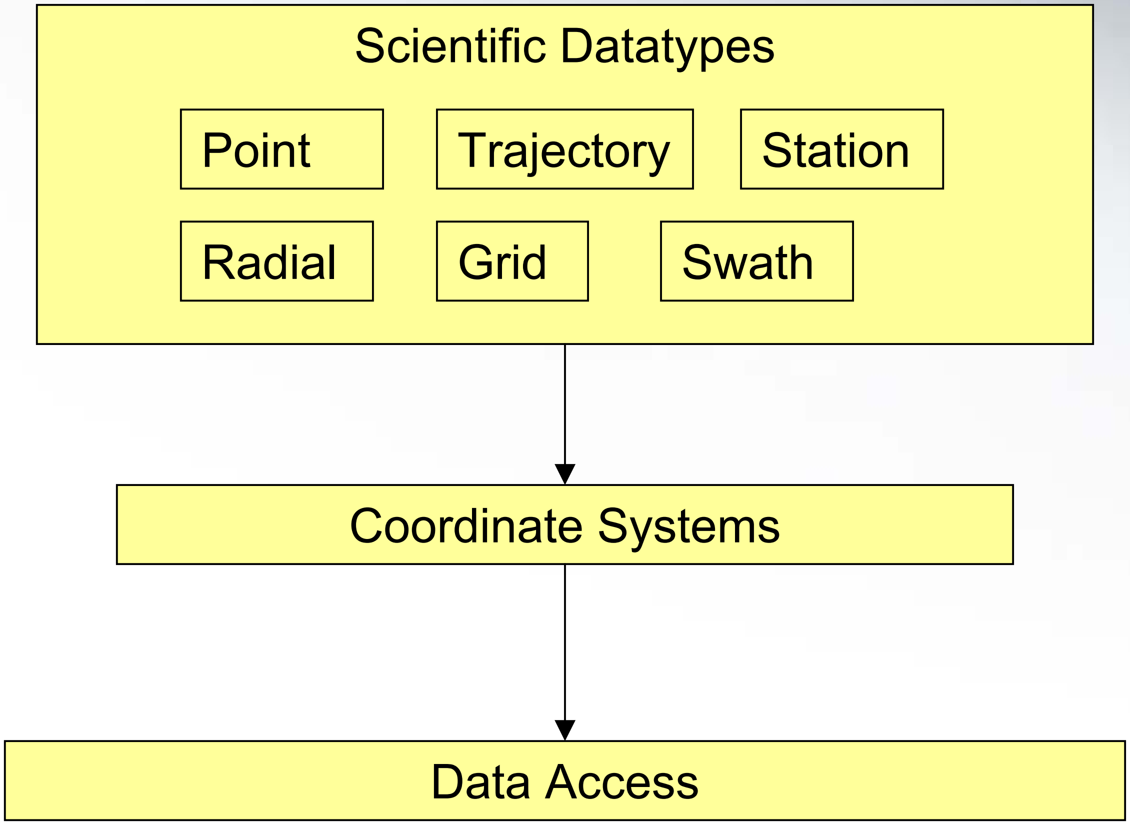
- EOSDIS core system (ECS) is evolving/ending
- ECS is being replaced with a new approach:
  - Define vision and principles
  - Prioritize to do list
  - Build a little, learn a little; repeat
  - These are extreme programming/continuous evolution project lifecycles vs. old waterfall/slow spiral lifecycles
- Common Data Model merges OPeNDAP, NetCDF, and HDF and allows a community-wide approach to interoperability principles
- Leverage this community through increased participation of Data Centers and CLASS in the Federation of Earth Science Information Partners (ESIP-Fed; NOAA, thanks to G. Withee, is now a sponsoring Agency along with NASA)

Evolution of CLASS – Making the Data Management Problem Tractable Using a Common Data Model

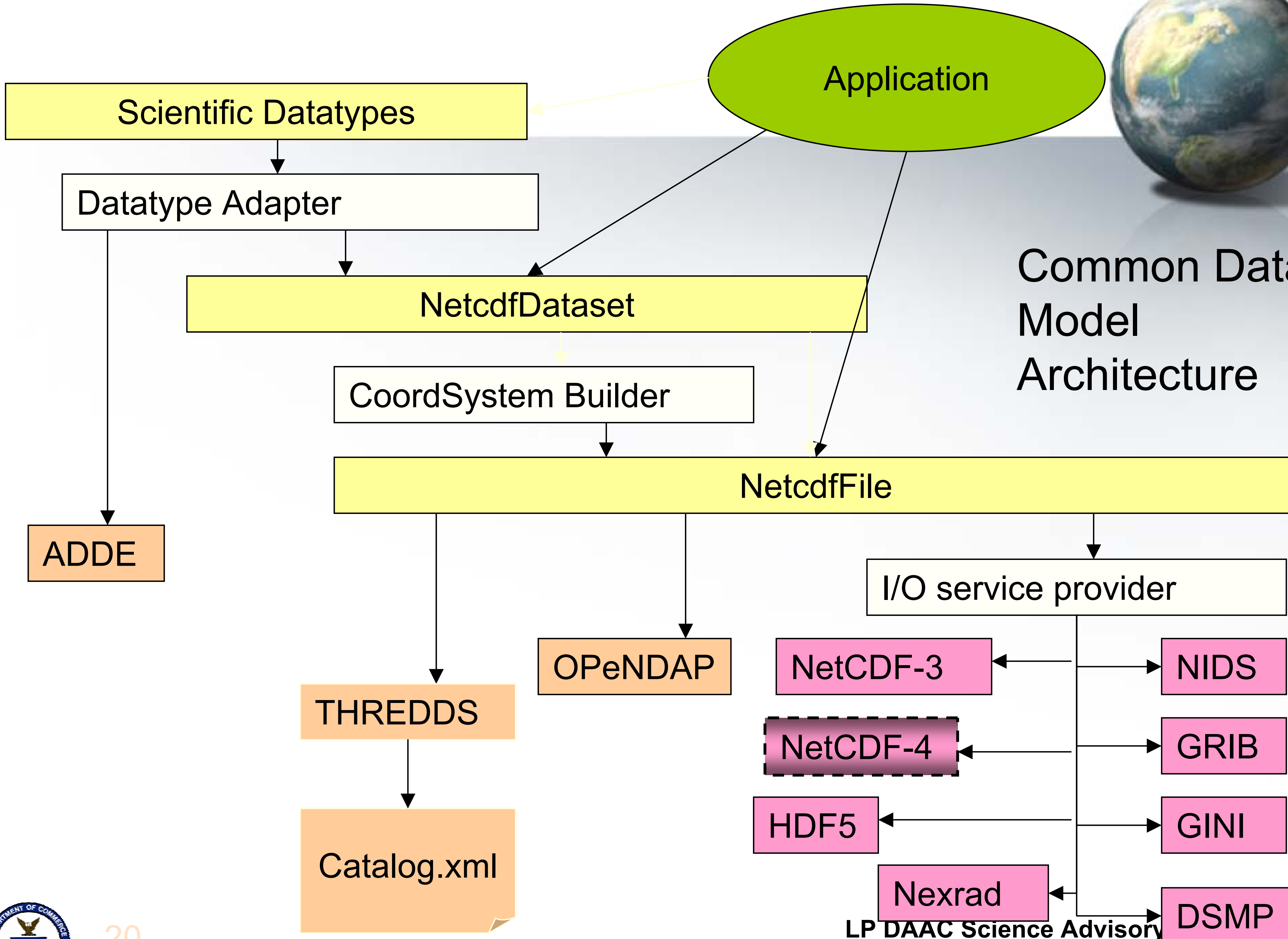
What's a Common Data Model?

- Its about scientific data: storing, accessing
- It's an abstraction
- Equivalent to an abstract object model in object oriented programming
- An **Abstract Data Model** describes data objects and what methods you can use on them
- This allows reuse of major portions of code and so reduces the volume of work (i.e., total cost) from an N x N dataset problem, where N gets very large, to a M x M data type problem, where M is constrained to a limited number.

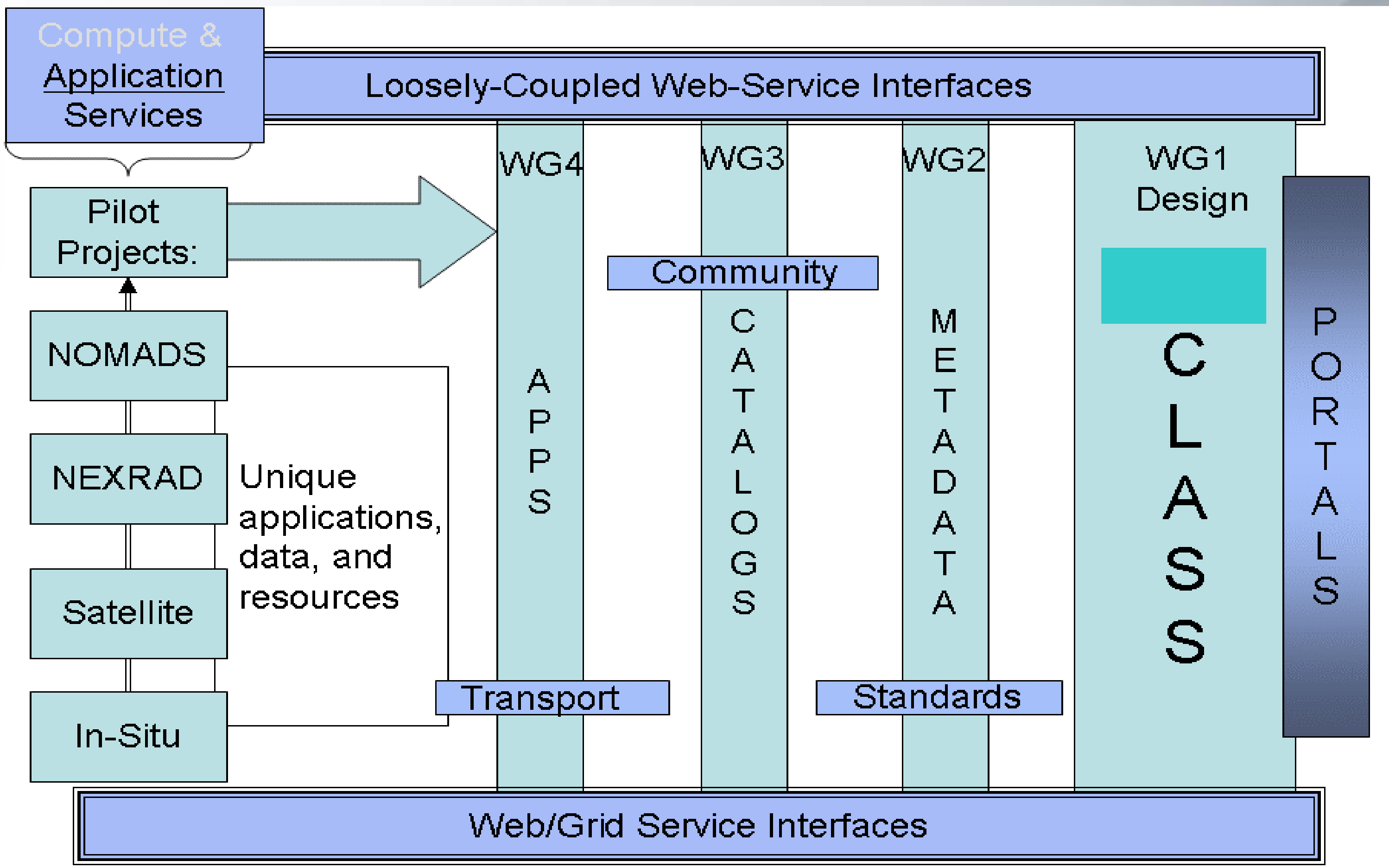
Common Data Model Layers



Common Data Model Architecture



Initial Common Data Model Work at the NNDCs – *Apply, Assess, Share Experience*



Opportunities for Inter-Agency and International Partnerships

- Build on what already exists…
- For Inter-Agency, the Federation of Earth System Information partners (ESIP) is making excellent progress
  - Technical working groups and application areas are the basis for getting work done
- For International, evolve existing partnerships
  - Work with Eumetsat on METOP-1 archive
  - Work with and evolve GEWEX radiation panel data management working group
- NOAA-NASA MOU on Research-Applications provides a new opportunity for partnership