

# Data Stewardship

## NOAA's Programs for Archive, Access, and Producing Climate Data Records

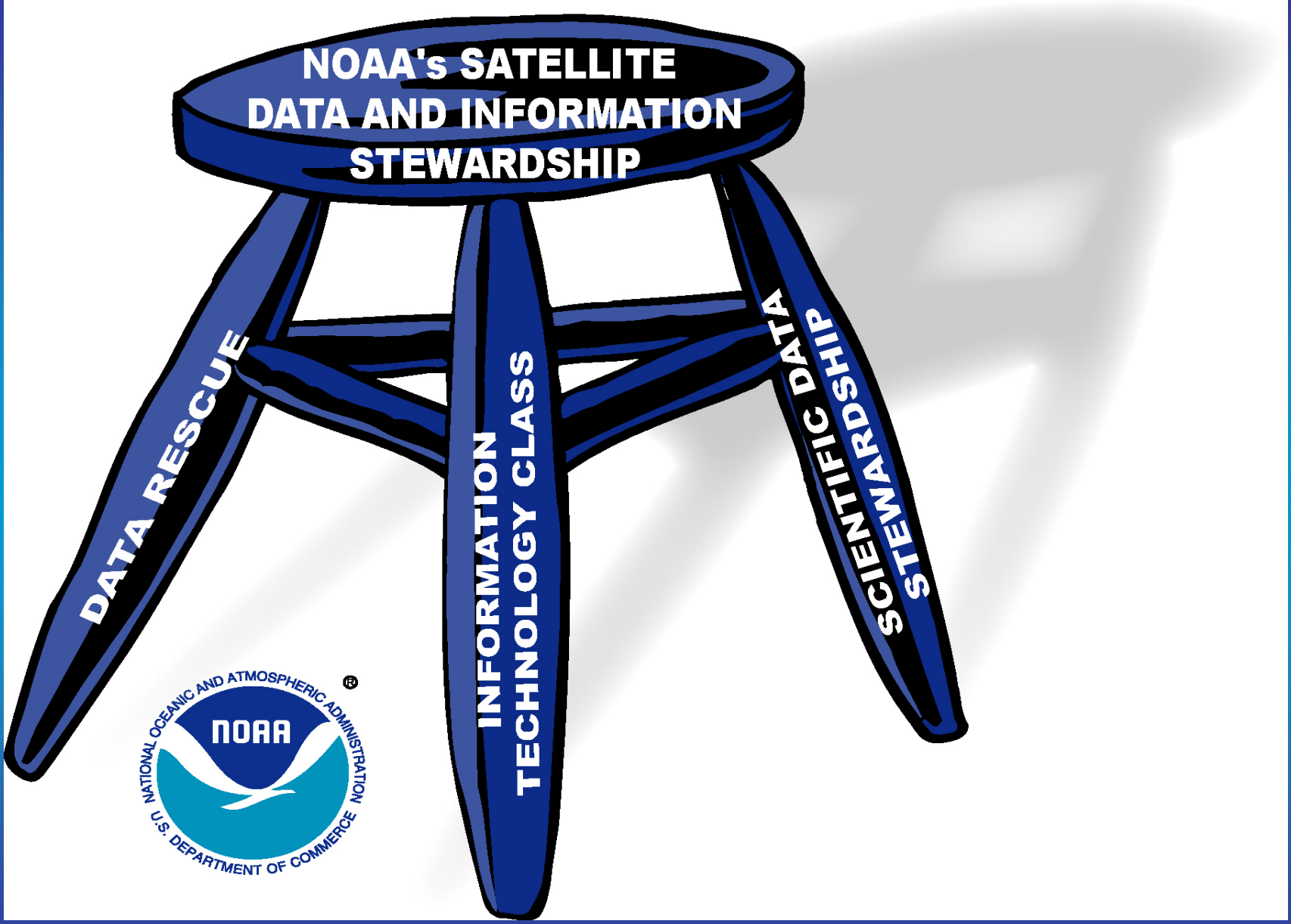
John Bates, *NOAA's National Climatic Data  
Center*

Mitch Goldberg, *NOAA's Office of Research  
and Applications*



# Outline

- Background
- Defining Data Stewardship
- NOAA's Comprehensive Large Array data Stewardship System (CLASS)
- NOAA's Scientific Data Stewardship (SDS) Program
- Conclusions
- Need for International Partnerships - GEO



# Principles of Data Stewardship

1. Archive and access to fundamental measurements, products and metadata - CLASS
2. Data archaeology and improved use – CLASS/SDS
3. Careful monitoring of observing system performance for long-term applications - SDS
4. Generation of authoritative long-term records through validation of the calibration process, reprocessing, product generation and the blending of *in situ* and satellite measurements - SDS
5. Provide state of the environment information for decision makers and place the current state in its historical context - SDS



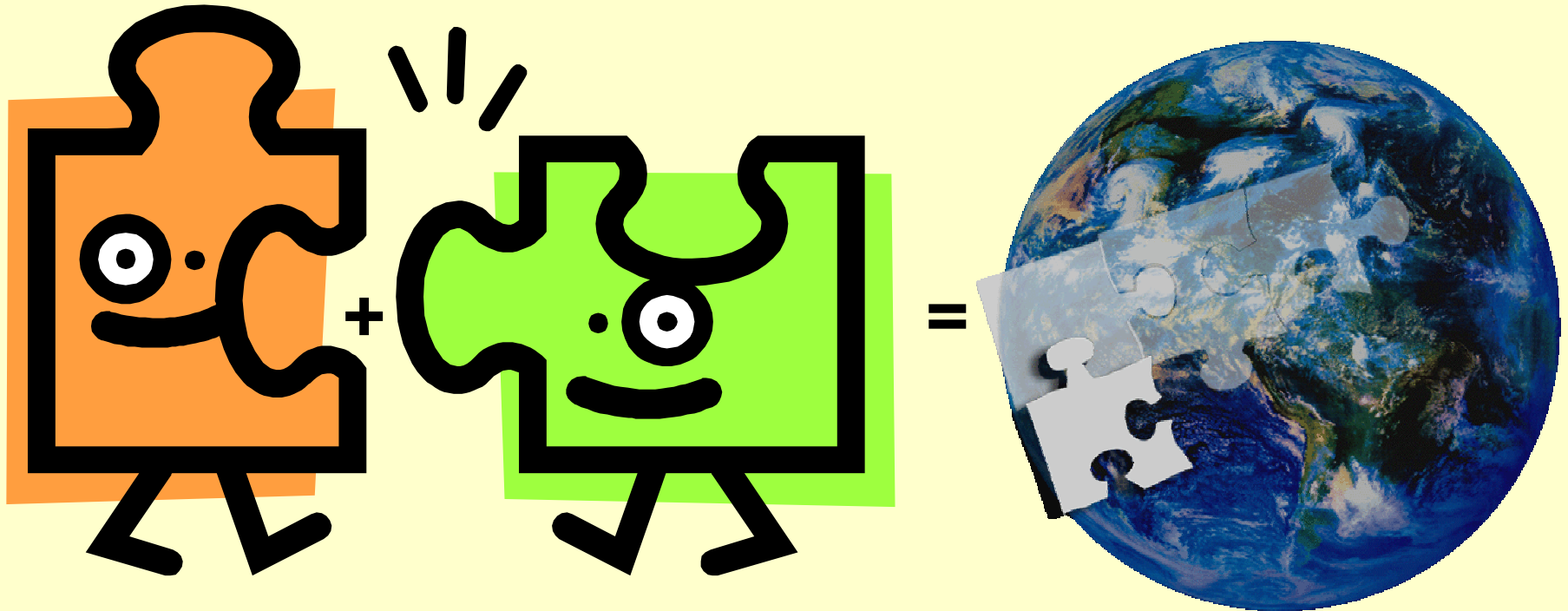
# CLASS Goals and Vision

NOAA's National Data Centers and their world-wide clientele of customers look to CLASS as the sole NOAA IT infrastructure project in which all current and future large array environmental data sets will reside. CLASS provides permanent, secure storage and safe, efficient access between the Data Centers and the customers.

- A web-based data archive and distribution system for NOAA's environmental data
- A combined process to both reengineer legacy data storage and access systems and blend new and efficient technologies to ensure the stewardship of existing (e.g., POES, GOES, NEXRAD, in-situ) and rapidly approaching large-array data sets (e.g., NPP/NPOESS, EOS, METOP, NEXRAD).
- An aggressive plan to safeguard, enhance, expand and automate NOAA's capability to ingest, store, quality control, preserve, and access its vast environmental data holdings.
- A focused effort to ensure the **Information Technology (IT) infrastructure** is in place and working, before the arrival of significantly larger and more complex environmental data (e.g., NPP/NPOESS, and GOES-R).



# Synergy of CLASS & SDS



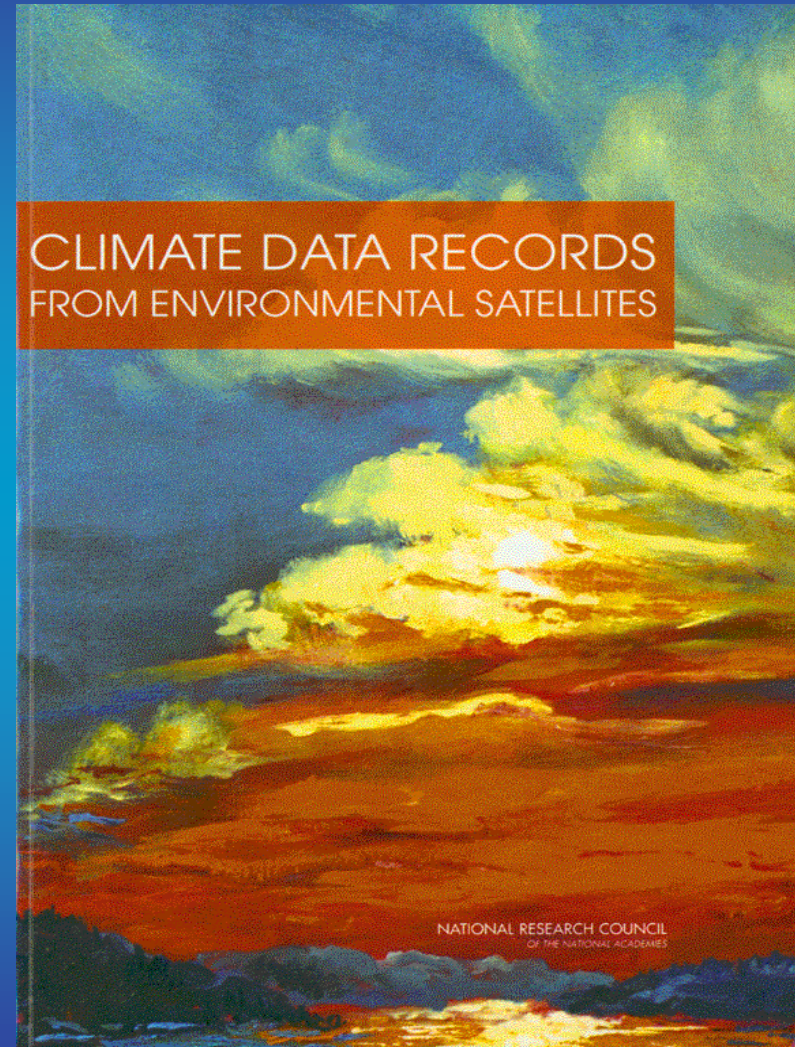
**CLASS + Scientific Data Stewardship = Climate Products and Information**

CLASS – NOAA's new archive; Comprehensive Large-Array data Stewardship System



# Background

- NOAA's Scientific Data Stewardship rooted in NRC dialogue and reports
- NOAA/NRC SDS leads
  - Bates
  - Goldberg



# Key Elements of a Successful CDR Program

## CDR Organizational Elements

- High-level leadership council
- Advisory council to represent climate research community and other stakeholders
- Fundamental Climate Data Record (FCDR) Teams
- Thematic Climate Data Record (TCDR) Teams

## CDR Generation Elements

- High accuracy and stability of FCDRs
- Pre-launch characterization of sensors and lifetime monitoring
- Thorough calibration of sensors
- Well-defined criteria for TCDR selection
- Stakeholder involvement and feedback for TCDRs
- Well-defined criteria for TCDR validation
- Use of *in-situ* data for validation

## Sustaining CDR Elements

- Available resources for reprocessing CDRs as new information becomes available
- Provisions for feedback from scientific community
- Long-term commitment of resources for generation and archiving of CDRs and associated data

**Fundamental Climate Data Record (FCDR):** Time series of calibrated signals for a family of sensors together with the ancillary data used to calibrate them.

**Thematic Climate Data Record (TCDR):** Geophysical variables derived from FCDRs, often generated by blending satellite observations, in-situ data, and model output.



# Defining Scientific Data Stewardship

## Notional Functions of Scientific Data Stewardship for Climate

### Scientific Data Stewardship

*Real time and retrospective management of climate data*

#### Network Performance Monitoring

Observations & Metadata

Analyses and Quality Control

Feedbacks

Stewardship Teams

#### Climate Data Records

Archives

Reference Data Sets  
(Reprocessing)

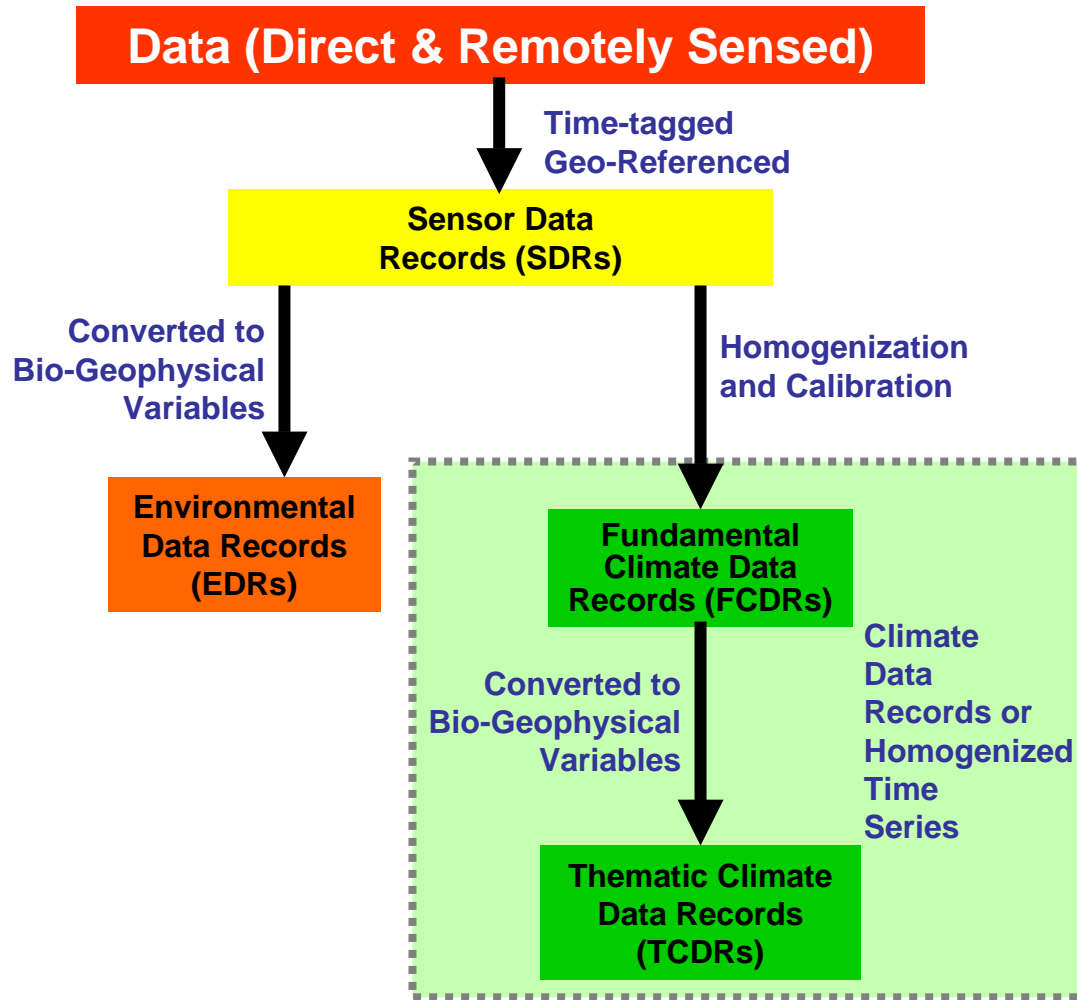
Model Re-analyses

Climate Quality Products

Climate Analyses

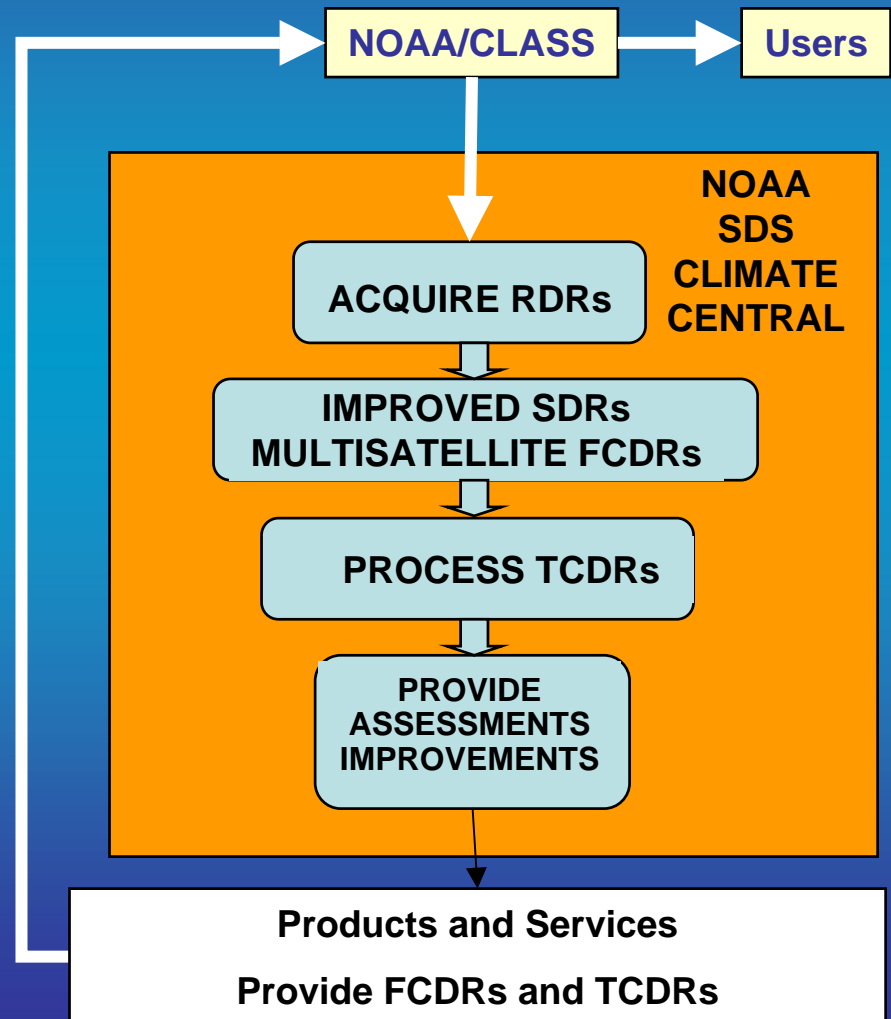
# Defining CDRs

## Climate Data Records



# NOAA CDR Reprocessing - Climate Central Requirements

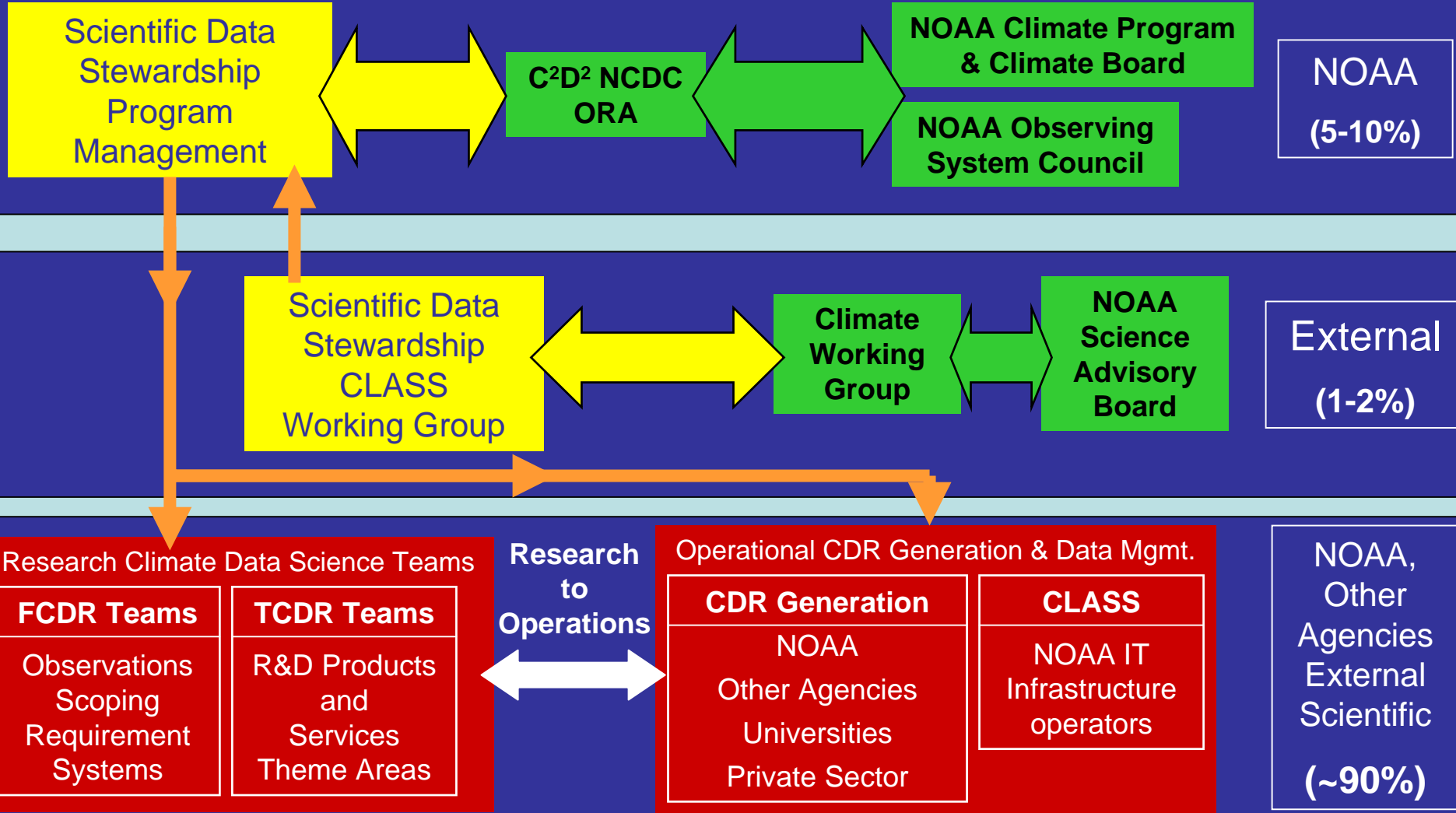
- For reprocessing, SDS requires an IDPS-like system, (Climate Central) to process:  
RDR → SDR →  
Multi-satellite FCORs & TCDRs
- SDS interdependent with CLASS, e.g., large data set I/O
- CLASS and SDS mutually dependent
- Requires NASA & W&W NDE linkage for enterprise solution



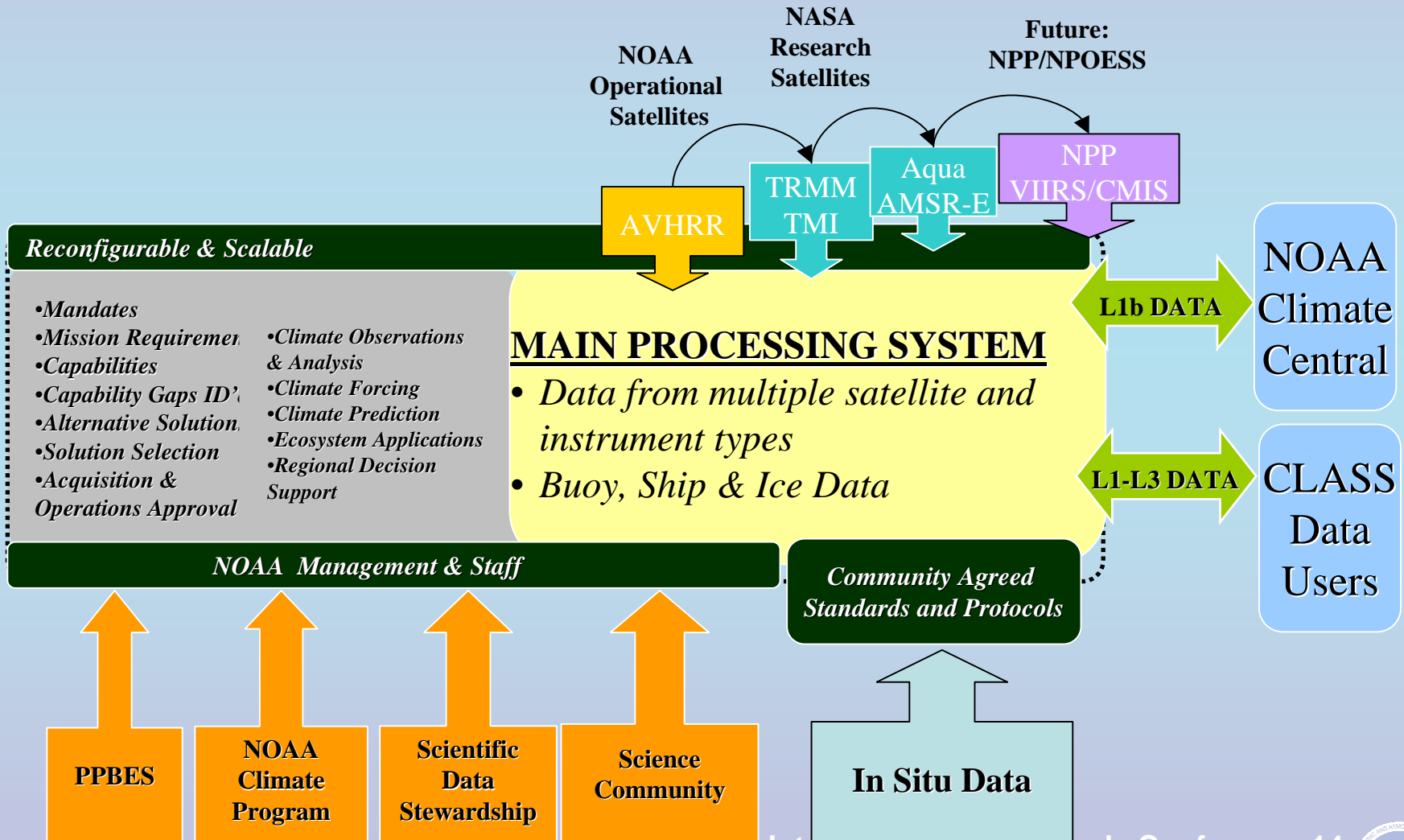
# NOAA's Scientific Data Stewardship Program

## Governance and Management Structure

FY06  
EDSM \$2.5M  
C<sup>2</sup>D<sup>2</sup> \$1 to \$1.5M



# NOAA's Scientific Data Stewardship Thematic Climate Data Records (TCDRs) Sea Surface Temperature



# Conclusions

- NOAA SDS is primary USA program for CDR development
- NOAA leadership for sustained program is required to fill lawful mandates
- User involvement in all aspects of SDS





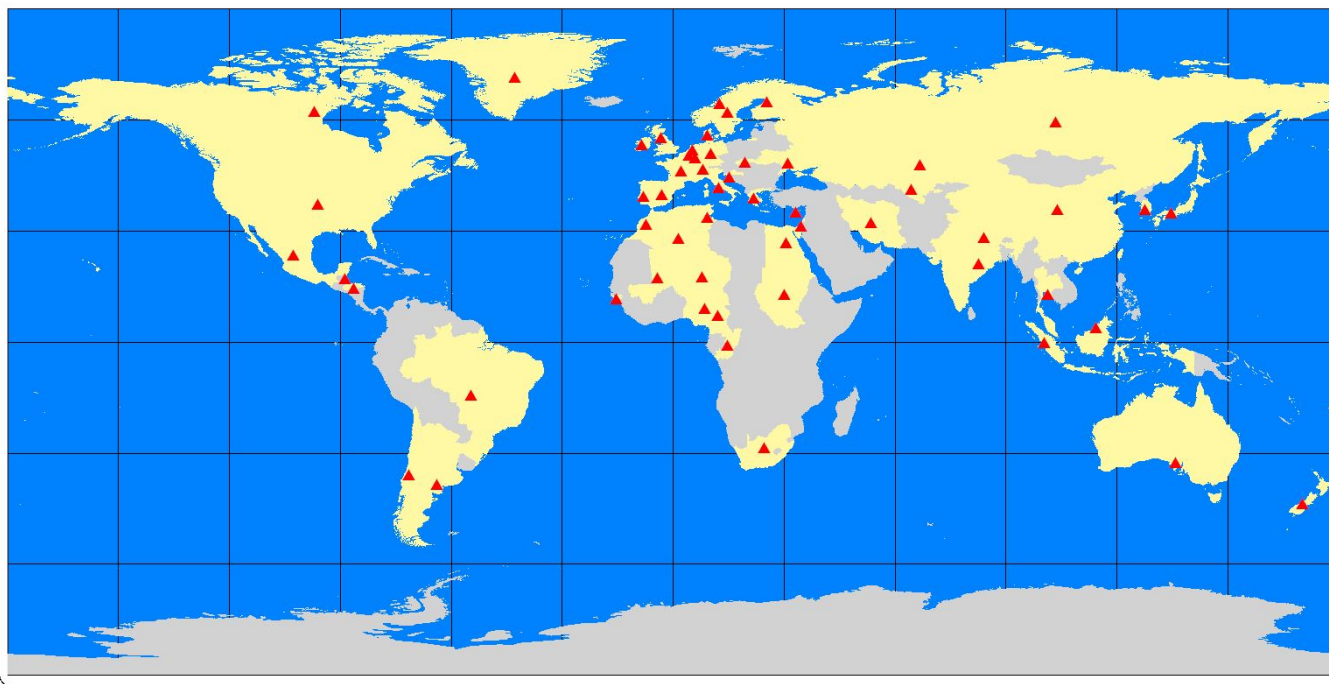
# Need for International Partnerships Group on Earth Observations GEO

## Group on Earth Observations - Member Countries

55 Members of GEO + EC  
05.03.05



Sources:  
Map - ESRI  
Members - <http://earthobservations.org>  
Design - Ministry of Environment, Israel



International TOVS Study Conference, 14<sup>th</sup>, ITSC-14, Beijing, China, 25-31 May 2005.  
Madison, WI, University of Wisconsin-Madison, Space Science and Engineering Center,  
Cooperative Institute for Meteorological Satellite Studies, 2005.