

Overview of NOAA Wide Data Management Issues



Tom Karl & Larry Tyminski
Co-Chairs
NOAA Data Management Committee

Opening Meeting
NOAA Science Advisory Board's
Data Archiving and Access Requirements Working Group (DAARWG)
December 7-8, 2006
Washington, DC

Outline

- Need for the SAB Data Archive & Access Requirements Working Group
- Data Management within NOAA's Line Offices / Goal Structure
- NOAA's Data Management Committee (DMC) role
 - Current DMC activities related to DAARWG
 - NOAA Data Management Report to Congress



NOAA Science Advisory Board

Data Archive and Access Requirements Working Group

• DAARWG Background

- Briefed and approved at the Winter 2006 NOAA SAB meeting
- Purpose: Provide recommendations to NOAA that:
 - help prioritize specific data sets & products that NOAA should archive
 - make specific recommendations with respect to data access consistent with established principles & guidelines

• DAARWG Rationale

- Exponential growth in NOAA data volume projected from:
 - Current and planned observing systems
 - NOAA and partners ability to generate new data products
 - NOAA's ability to integrate data from a variety of observing systems & disciplines
 - ...many others
- Value of ongoing advice from SAB external users and experts



NOAA Vision and Goals

NOAA VISION: *An informed society that uses a comprehensive understanding of the role of the oceans, coasts, and atmosphere in the global ecosystem to make the best social and economic decisions*

Five NOAA Mission Goals developed from stakeholders/internal review

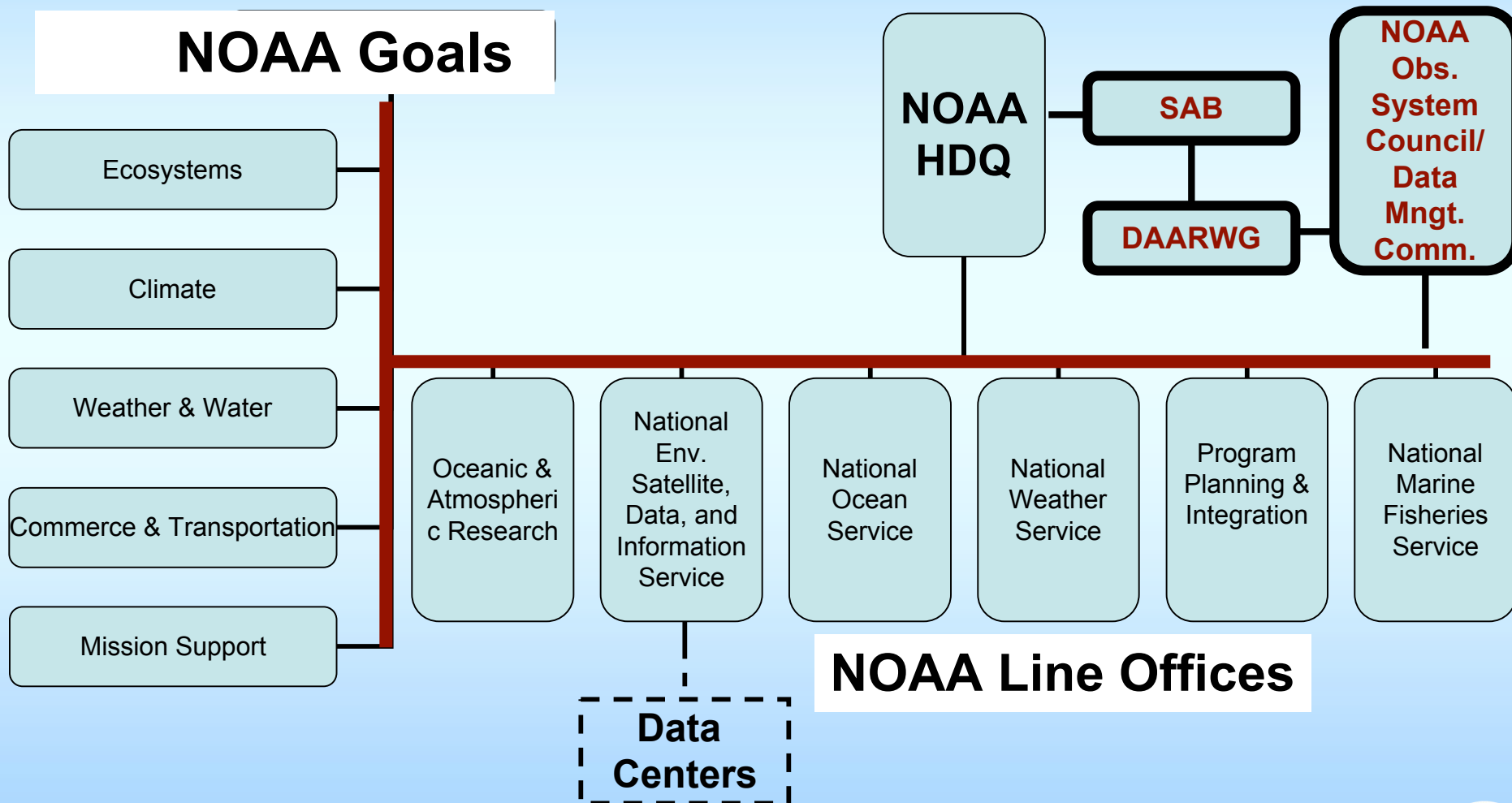
- **Ecosystems**
 - Protect, Restore, and Manage the Use of Coastal and Ocean Resources Through an Ecosystem Approach to Management
- **Climate**
 - Understand Climate Variability and Change to Enhance Society's Ability to Plan and Respond
- **Weather and Water**
 - Serve Society's Needs for Weather and Water Information
- **Commerce and Transportation**
 - Support the Nation's Commerce with Information for Safe, Efficient, Environmentally Sound Transportation
- **Mission Support**
 - Provide Critical Support for NOAA's Mission

NOAA Line Offices: *Execute activities required to achieve goals. Mission goals may involve the activities of more than one Line or Staff Office.*

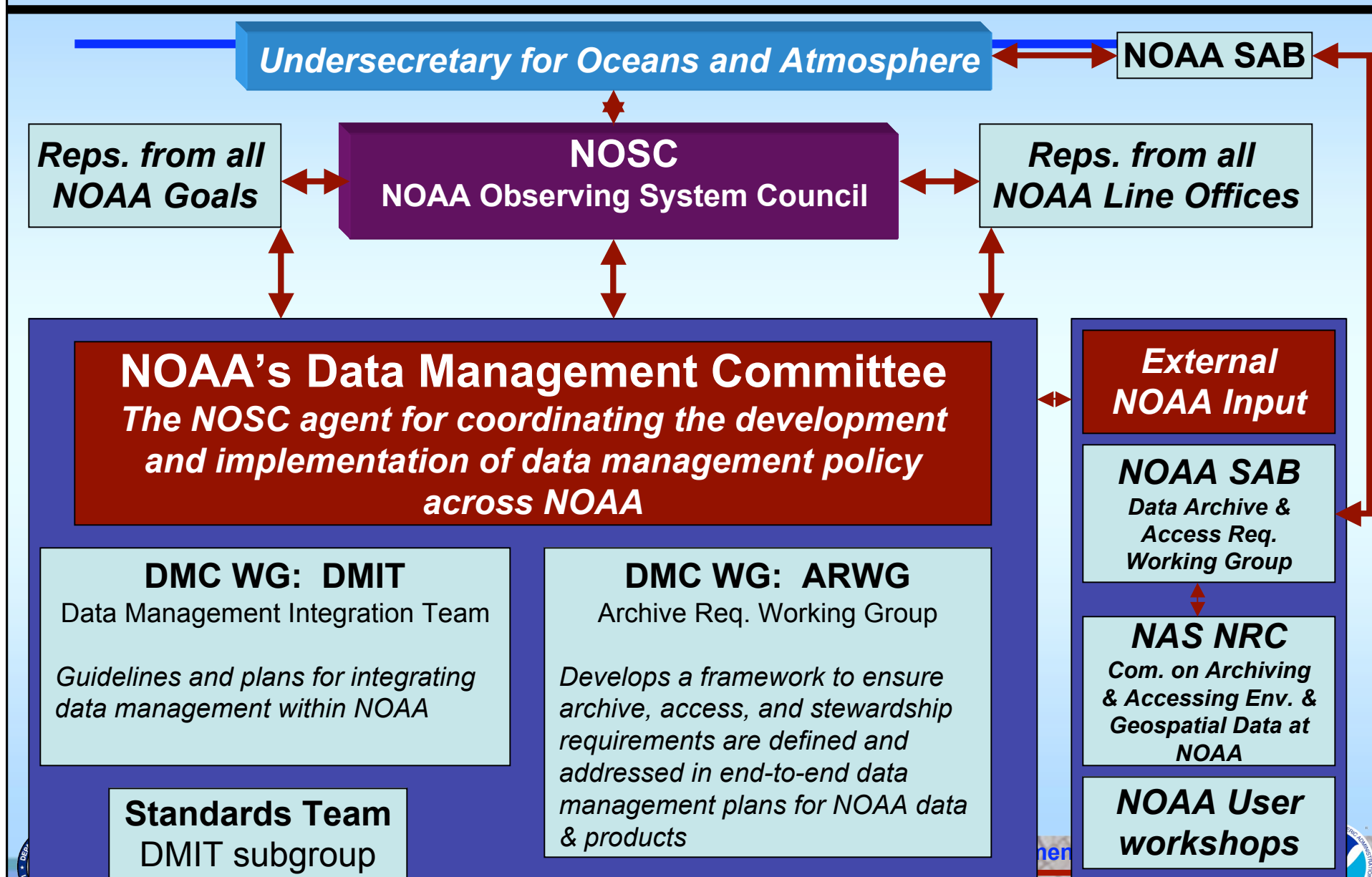


NOAA Organization

Matrixed Managed Goals and Line Offices



NOAA's Data Management Committee Reporting Structure



Current DMC activities related to DAARWG



7

December 7-8, 2006

Data Archive & Access Requirements Working Group
Washington, DC



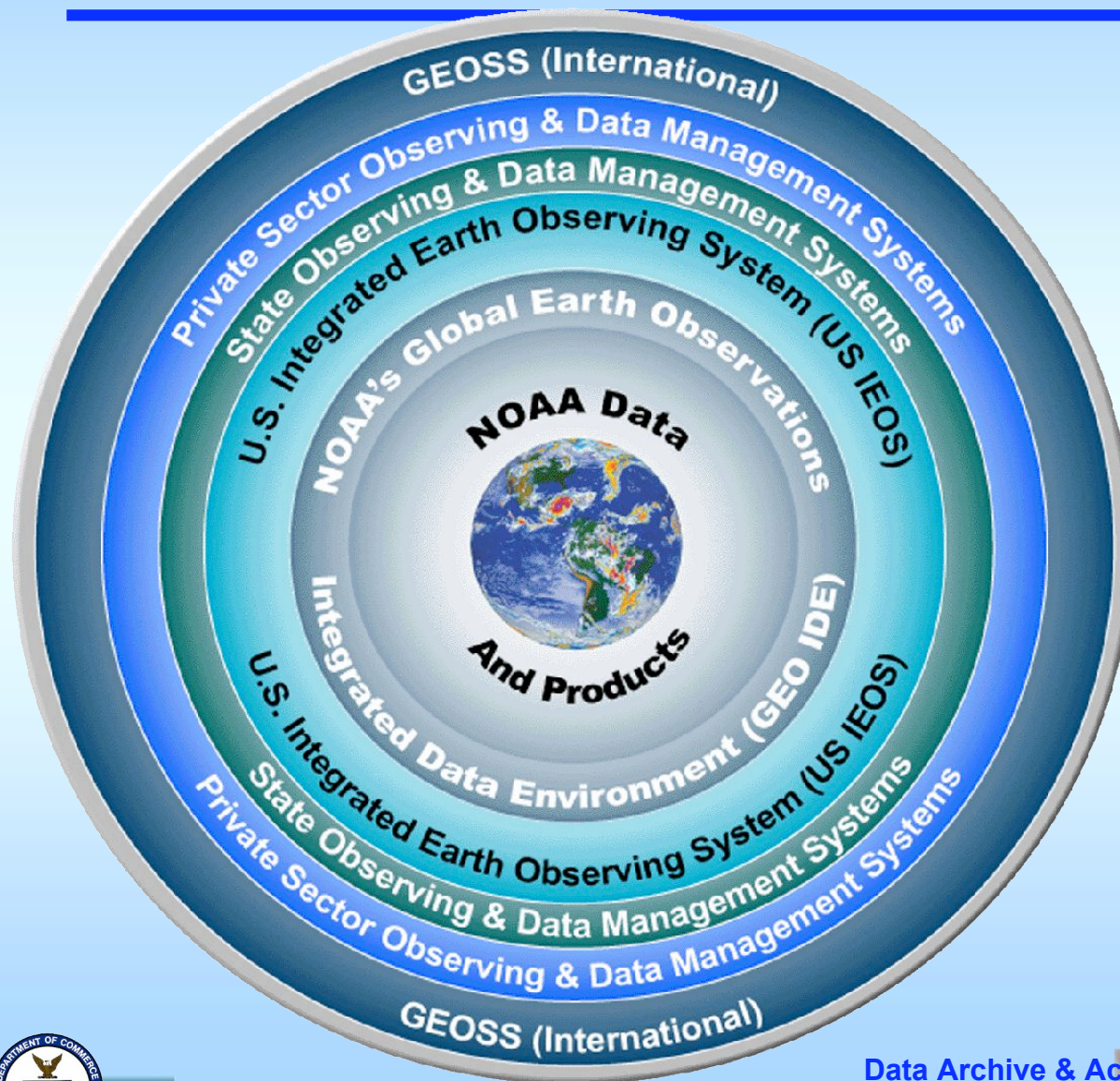
Data Access & Archive Req. Committee Reports by the National Research Council

- **NRC Committee established by DMC in Nov 2005**
 - **Committee Goals:**
 - Develop a list of principles & guidelines regarding what types of data should be permanently archived and what types of data could be stored for shorter durations under budgetary constraints
 - Develop a list of principles and guidelines on how best to provide access to different variables, data sets and derived products
 - **Activities**
 - DMC coordinated meeting held Mar 2006 on NOAA data archive
 - Phase 1 NRC report complete May 2006 on data archive principles
 - DMC coordinated meeting held Oct 3-4 on NOAA data access
 - Phase 2 NRC report to be completed Summer 2007 report

NRC Principles & Guidelines Reports will be used by NOAA's Science Advisory Board's Data Archive & Access Requirements Working Group (DAARWG) to recommend prioritization of specific NOAA data sets



GEO-IDE - an essential component of environmental information management for NOAA



Integrated observing, data processing and information management systems

Connected by NOAA's Integrated Data Environment

Contributes to U.S. Global Earth Observation System (USGEO) and International Global Earth Observing System of Systems (GEOSS).

GEO IDE Documents

Implementation Plan & Standards

NOAA

**Global Earth Observation
Integrated Data Environment
Implementation Plan**

---DRAFT---
Version 1.2
13 April 2006

Prepared by:
U.S. Department of Commerce
National Oceanic and Atmospheric Administration
Data Management Committee (DMC)
Data Management Integration Team

NOAA Data Management Integration Team (DMIT) Standards
(Version 4.2 - 15 February 2006)

- Recommendation 1.1 - Discovery-level metadata content standards
- Recommendation 1.2 - Discovery-level keyword lexicon
- Recommendation 1.3 - Discovery-level metadata representation/exchange standard
- Recommendation 1.4 - Catalogue search protocol specification
- Recommendation 1.5 - Comprehensive use-level metadata
- Recommendation 2.1 - File transfer protocols and APIs
- Recommendation 3.1 - Database access methods
- Recommendation 4.1 - Web Services
- Recommendation 5 - Data and product format standards for delivery
- Recommendation 5.5 - XML schemas for NOAA data
- Recommendation 6 - Standards for accuracy and content of geospatial data

2006 Standards Document

Table of Contents

1. Introduction	1
2. Management, Leadership and Support.....	3
3. Standards.....	6
4. Skills Development.....	9
5. Applications Management	10
6. Technology Management.....	11
7. Web Services	12

2006 Implementation Plan and Table of Contents



User Input for NOAA Data Archive & Access

NOAA Data Users' Workshops

- **Recent NOAA Workshops**
 - primarily NESDIS led
 - Data users from Academia, Business, Government
 - 2003: Boulder, CO (375 users)
 - 2004: Kansas City, MO (NWS cosponsor) (70users)
 - 2005: Asheville, NC (NWS assisted) (204 users)
- **Next Workshop**
 - To be held Spring 2008 in DC/ Baltimore area
 - Better cross-section of NOAA Line Offices and Goals
 - **Possible theme: Delivering Integrated Data & Information in the 21st Century**



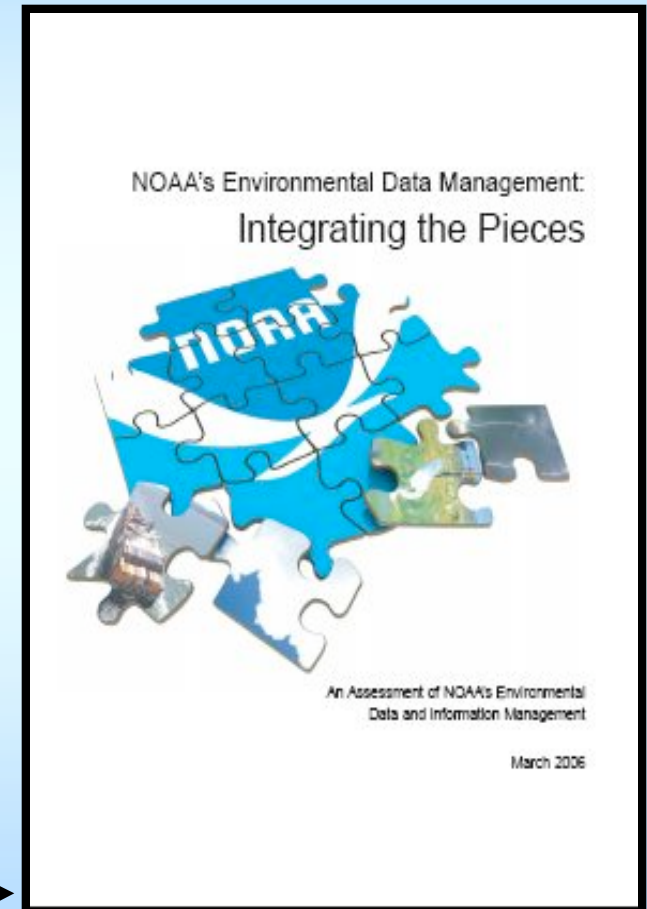
*Integrated Data & Information
in the 21st Century*

Data Management Report to Congress



NOAA Data Management Report to Congress

- **Based on Public Law 106-554 Sect. 515**
 - “...biennially...NOAA must complete and report to Congress an assessment of the adequacy of the environmental data and information systems...”
- **Assess capability to:**
 - acquire and transmit observations needed by the Program (NOAA and non-NOAA source)
 - validate and describe the observations, scientific data, and products
 - archive and provide discovery and access to these data and products in such a way as to enable integration across platforms and disciplines.
- **Corollary but more detailed report given to DAARWG (76 pp)**
 - Assesses capability by NOAA Mission Goals as opposed to just overall NOAA capability



Overall NOAA Assessment

2005 Summary Assessment Chart for all NOAA Goals

	Doing with current resources	End-to-End Environmental Data Management Functions										
		Observation Acquisition & Transmission			Scientific Data Management				Archive & Access			
		5-Year Plan	Maintain & Monitor	Collect & Rescue	5-Year Plan	Calibrate & Validate	Appropriate Formats	Complete Metadata	5-Year Plan	Long-term Preservation	Data Discovery	Access / Disseminate
NOAA Mission Goals	Ecosystems	P			P				N			
	Climate	F			F				F			
	Weather & Water	P			P				P			
	Commerce & Transportation	F			P				P			

- **Compared to 4 years ago, NOAA is doing a better job at:**
 - Collection and verification
 - Processing and transmission
 - Calibration, validation, and quality review
 - Corporate planning
- **NOAA is in the process of:**
 - Developing an integrated Earth observing and data management system

NOAA Ongoing Challenges

2005 Summary for all NOAA Goals

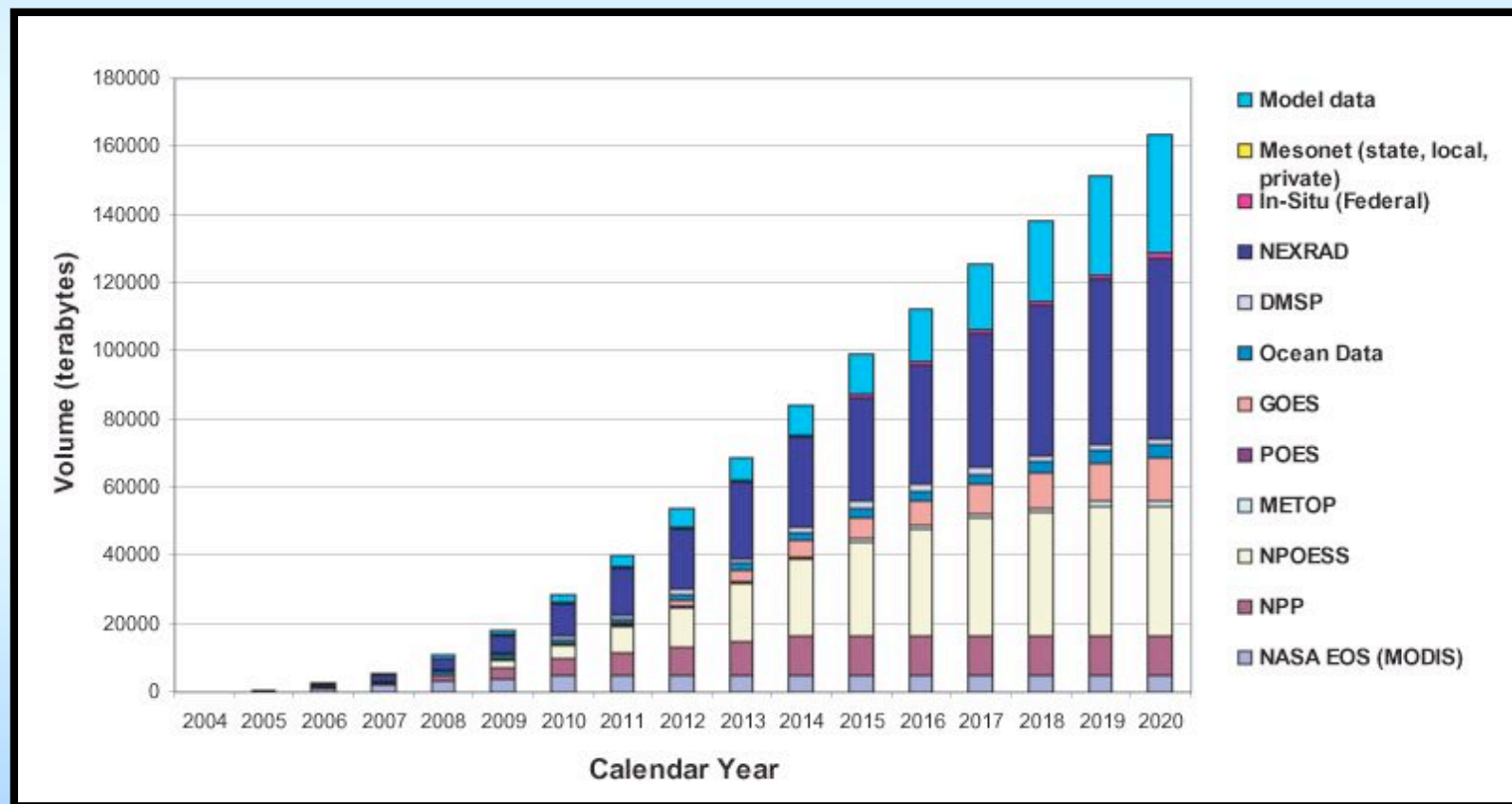
- **Size**
 - Managing exponentially growing data volumes
- **Metadata**
 - Appropriately describing data to ensure long-term utility
- **Integration**
 - Providing data in standard formats and protocols to enable integration
- **Access**
 - Providing clear and easy discovery of and access to data and products
- **At Risk Data Sets**
 - Collecting data at risk to extend the environmental record



Highlighted Challenge

Data Volume

NOAA must acquire, transmit, process, quality assure, document, develop & deliver products, & archive greatly increasing data volumes



From 2005 Report to Congress

Data Archive & Access Requirements Working Group
Washington, DC



Agenda for Remainder of Workshop

- Session 1: NOAA Data Center presentations
- Session 2: NOAA Goal presentations
- Session 3: Cross cutting data management presentations
- Session 4: National Research Council report presentations
- Session 5: Discussion and next steps

