

*NOAA Science Advisory Board's
Data Archiving and Access Requirements Working Group
December 7-8, 2006*

CLASS and Enterprise Solutions
Rick Vizbulis



Agenda

- CLASS history
- What is an “archive?”
- Archive responsibilities
- What does definition of “Archive” mean for NOAA?
- Organizational changes needed to support archive
- Key roles and responsibilities – a starting point
- CLASS evolution to support archive
- Significant challenges

CLASS History

- SAA, GAA merged in 2001
- CLASS focused on a small number of data types (~ “magnificent seven”)
 - No formal, cross-cutting agreements regarding what archival is, roles and responsibilities, etc.
 - CLASS assumes some responsibilities that rightly belong to the archive(s)
 - OAIS-RM responsibilities are not addressed
- OAIS-RM takes hold and CLASS begins to adopt portions
- CLASS Level-1 Requirements identify
 - CLASS as “NOAA’s enterprise IT solution supporting archive”
 - CLASS will adopt OAIS-RM

What is an “archive?”

- Term is changing dramatically given the influences of digital data and their preservation

- The OAIS-RM provides the best modern definition:

“... an organization of people and systems ... that has accepted the responsibility to preserve information and make it available for a designated community.”

- Responsibilities, in abbreviated form:

- Negotiate for and accept information
- Obtain sufficient control of the information to ensure long-term preservation
- Determine the “designated community”
- Ensure “independent understandability”
- Follow documented policies and procedures that ensure preservation
- Make the preserved information accessible

What does the new definition of “Archive” mean for NOAA?

- *Information, understanding, and preservation* are inexplicably linked in a modern *archive*, thus ...
 - Archival is far more than simply preserving bits
 - Archival requires subject matter experts
 - Modern digital archives require substantial IT systems and expertise

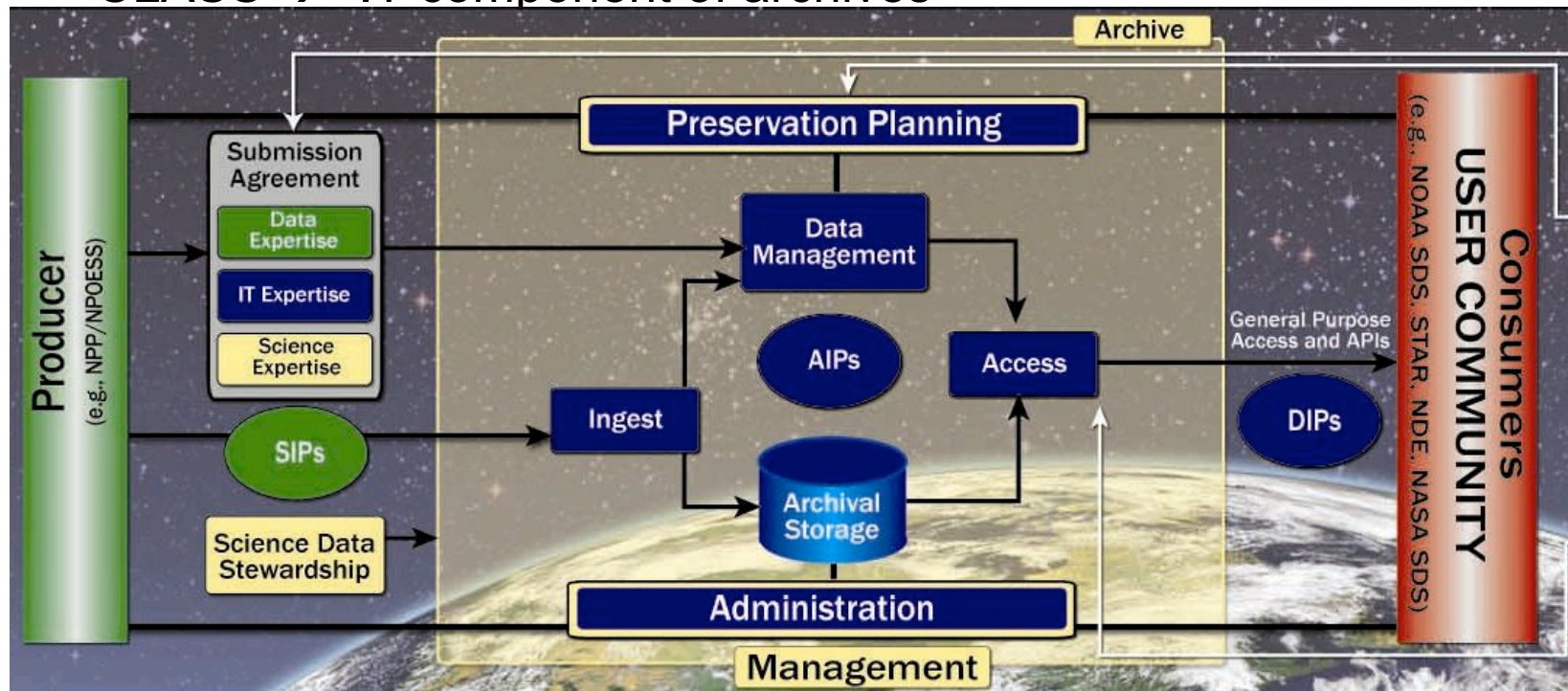
→ Archival is a cooperative, organizational effort

Organizational changes needed to support archive

- Key initial organizational steps
 - Organizational adoption of the OAIS-RM as NOAA’s archive model
 - Establishment of archive ConOps
 - Clarification of archive roles and responsibilities
 - Document requirements for key participants, systems
 - Establishment of “What to Archive” & “Where to Archive” policies and decision processes
 - Establishment of clear, consistent, archive requirements process
- Key initial technical steps
 - Establishment of organizational process for standards, interoperability

Key roles and responsibilities – a starting point

- NNDCs → “Archives”
 - Overall responsibility for OAIS-RM “mandatory responsibilities”
 - Key responsibility → Preservation planning
- CLASS → “IT component of archives”



Evolution to support 'enterprise' archive

- Technical
 - Transition to architecture based on OAIS-RM
 - Interoperability transitions as developed in conjunction with GEO-IDE (service-oriented architecture, standards, specific interfaces)
 - Incorporate mechanisms to mitigate “data specificity” issues
Candidates include: (Plug-ins, Data models, Data Definition Languages...)
 - Incorporate mechanisms (primarily ingest and access extensions) in support of archive data stewardship activities
 - Develop extensible metadata model
 - Enhance geospatial capabilities

Standards, standards, and more standards!

Enterprise Transition Issues

- Need 'enterprise' archive requirements, standards, processes & procedures
 - Documented organizational roles and responsibilities
 - Published organizational process, evaluation criteria & priority
 - Detailed enterprise archive requirements
 - Specific Standards (data formats, metadata, interfaces)

Benefit Data Producers, Archives & IT system
Developers



Thank You

Background

Long Term Architecture

- Long term System Architecture Overview - complete
- “To-Be” System Architecture Overview – in progress
- Transition Plan
- System Architecture Reference Manual

GEO-IDE & CLASS

- **CLASS shall initiate pilot programs with the GEO-IDE project to support risk reducing development and phased integration of standards for metadata, machine-to-machine interfaces, and archive**
 - Identify how systems will interoperate with CLASS
 - Enumerate CLASS services and APIs for:

- Ingest
- Discovery/search
- Metadata management
- Dissemination
 - Notification
 - Subscription
 - Order
 - Delivery
 - Media transfer

- Geospatial capabilities
- Data Operations
 - Metadata extraction
 - Subsetting
 - Visualization
 - Aggregation
 - Format translation
 - ...

NGDC & CLASS

- NGDC Application Programming Interface
 - Integrates selected CLASS & NGDC capabilities
 - Demonstrates ease of integration using SOA technologies (Web Services in this case)
 - Shows benefits and feasibility of integrating NOAA systems
 - Provides “lessons learned” for production CLASS API development and collaboration with GEO-IDE

NODC and CLASS

- NODC Pilot Project documentation and definition:
 - Initial Operational Capability (IOC) Requirements for National Oceanographic Data Center (NODC) Archival Data Transfer
 - Operational Threads of Initial Operational Capability (IOC) for National Oceanographic Data Center (NODC) Archival Data Transfer
- Implementation February 2007

NCDC and CLASS

- Implement National Operational Model Archive & Distribution System (NOMADS) front-end to CLASS back-end data management services
- Incorporate NEXRAD National Mosaic and QPE products in CLASS
- Incorporate Integrated Surface Data (in NetCDF) in CLASS

Geospatial Efforts

- **The CLASS access tool and application program interface shall support geospatial discovery, browse and access**
- Feedback from CLASS User Workshop include potential Use Cases:
 - Inventory Search
 - Metadata Ingest
 - Metadata Presentation
 - Data Subscription
 - Metadata Subscription
 - Data sub-setting
 - Access
- Work with DMC Archive Requirements Working Group to validate & prioritize
- Implement through GEO-IDE risk reduction