

# NODC Ocean Data Archives and Metadata

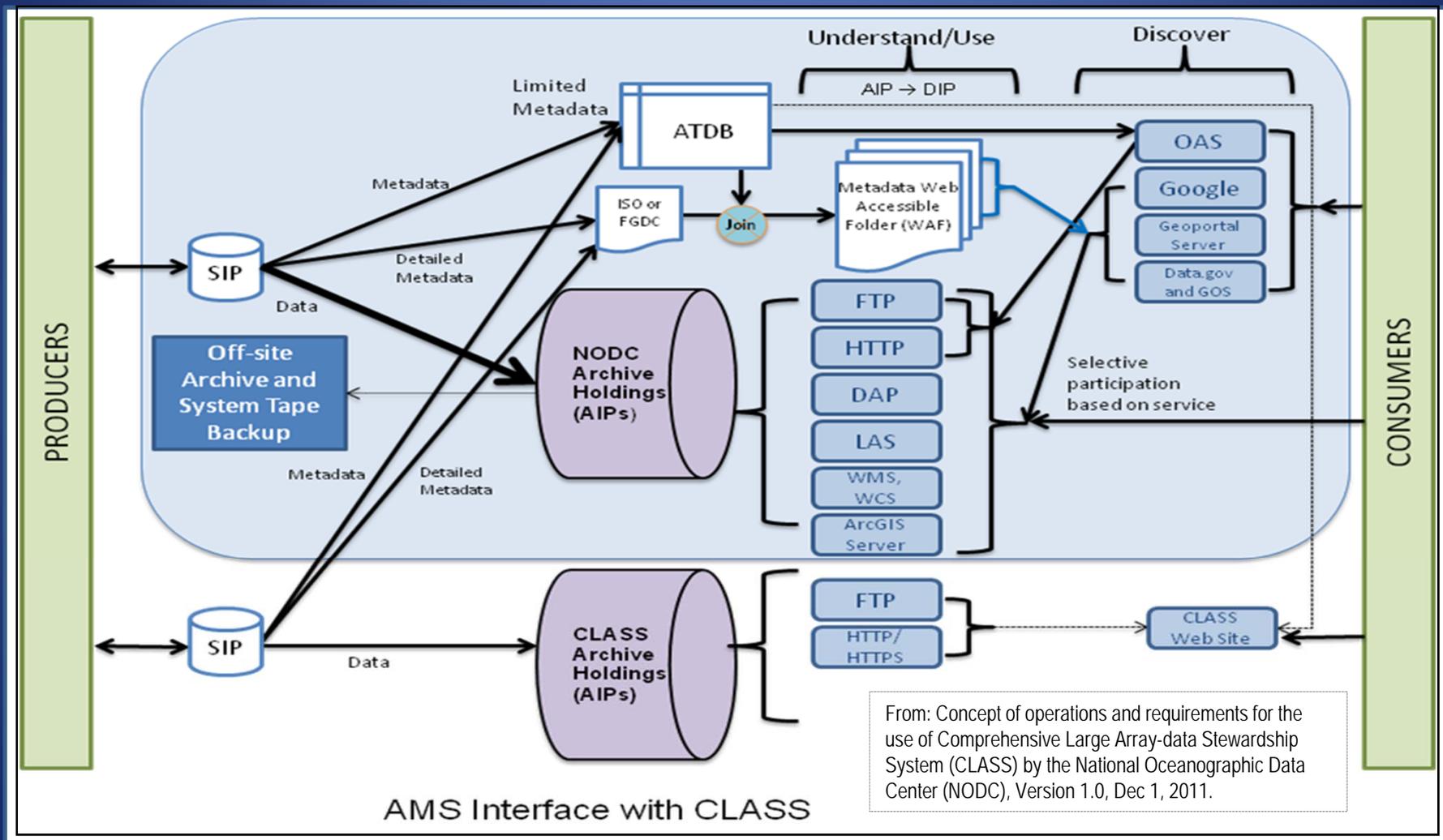
Don Collins

NOAA/National Oceanographic Data Center

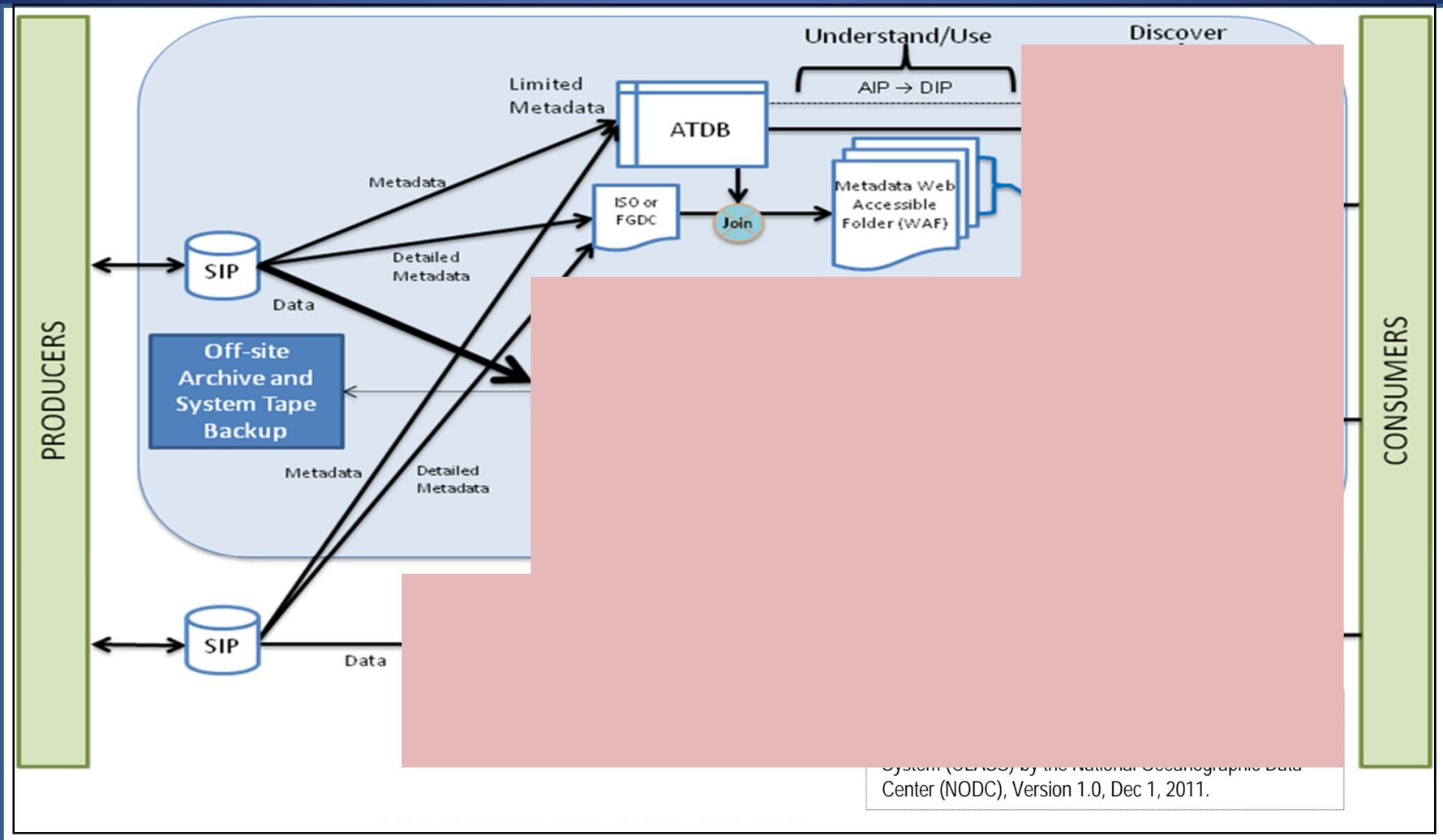
# NODC Archive Management System: Diverse holdings of ocean data and information

- NODC scope generally includes
  - Global deep ocean to inland coastal waters and Great Lakes
  - Physical, chemical, biological *in situ* observations
  - Space-based ocean observations and products
  - Ocean model outputs, ecosystems analyses...
  - Photo collections, video collections...
- “NOAA Procedure for Scientific Records Appraisal and Archival Approval: Guide for data managers”

# Format-agnostic data stewardship



# SIP Ingest and AIP Metadata



# Standards-oriented metadata for heterogeneous data

NODC  
National Oceanographic Data Center

You are here: [AMS](#) > Accession Preferences Logged on as:

**AMS: Search Accession Preferences:**

Accessions | People | Projects | Institutions | Countries | Platforms | Seanames | Datatypes | Instruments | Observations

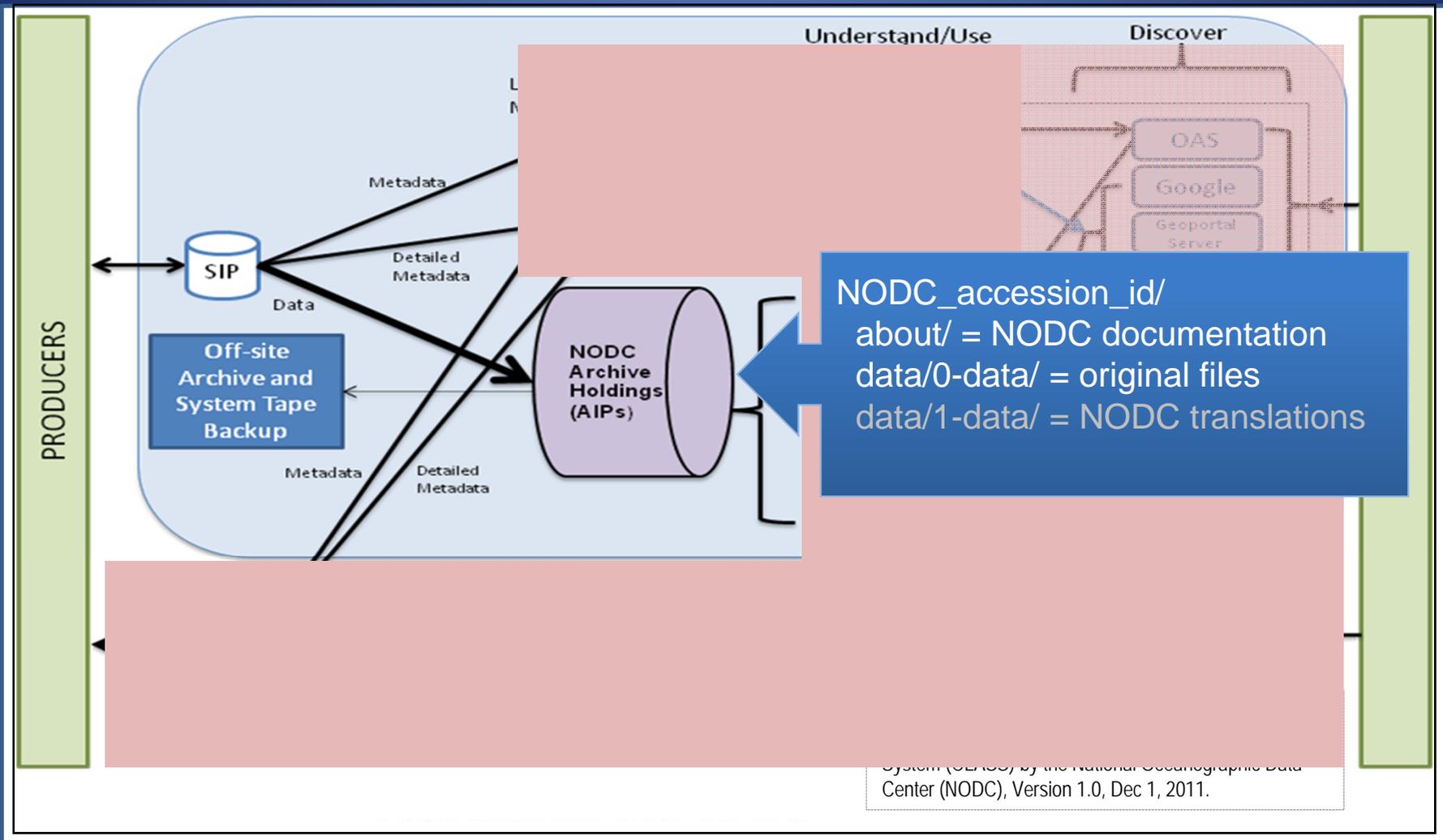
Preferences | Reports | Search

**Advanced Search: Choose the fields you would like to include in your search, and then press Continue below.**

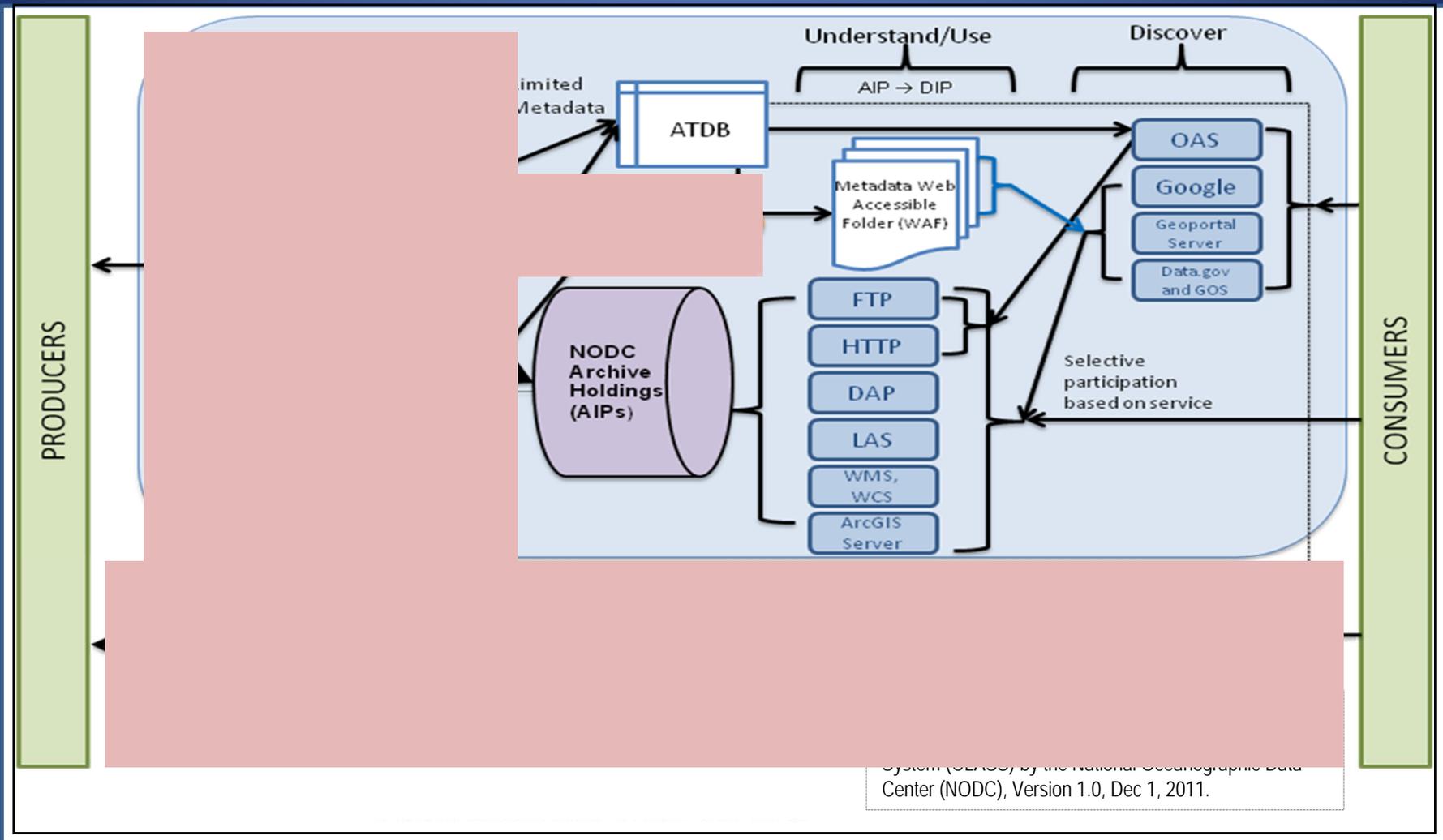
| data  | location   | time   |
|---|--|--|
| <input type="checkbox"/> accessions_id                      | <input checked="" type="checkbox"/> Seanames       | <input type="checkbox"/> Date received         |
| <input checked="" type="checkbox"/> Title                   | <input checked="" type="checkbox"/> West boundary  | <input checked="" type="checkbox"/> Start date |
| <input checked="" type="checkbox"/> Abstract                | <input checked="" type="checkbox"/> East boundary  | <input checked="" type="checkbox"/> End date   |
| <input checked="" type="checkbox"/> Observation types       | <input checked="" type="checkbox"/> North boundary | <input type="checkbox"/> Availability date     |
| <input checked="" type="checkbox"/> Instrument types        | <input checked="" type="checkbox"/> South boundary | <input type="checkbox"/> Keydate               |
| <input checked="" type="checkbox"/> Datatypes               |  | <input type="checkbox"/> Editdate              |
| <input checked="" type="checkbox"/> Submitting institution  |  |  |
| <input type="checkbox"/> Submitter                          |  |  |
| <input checked="" type="checkbox"/> Collecting institutions |  |  |
| <input checked="" type="checkbox"/> Contributing projects   |  |  |
| <input checked="" type="checkbox"/> Platforms               |  |  |
| <input type="checkbox"/> NODC contact                       |  |  |
| <input type="checkbox"/> Disposition                        |  |  |
| <input type="checkbox"/> Current status                     |  |  |
| <input type="checkbox"/> Custodian                          |  |  |
| <input type="checkbox"/> Metadata quality                   |  |  |
| <input type="checkbox"/> Metadata version                   |  |  |
| <input type="checkbox"/> Keyer                              |  |  |
| <input type="checkbox"/> Editor                             |  |  |

Work in progress...  
XML\_blob element to manage  
additional metadata that does  
not have an existing ATDB  
element. III, 14-16 Aug

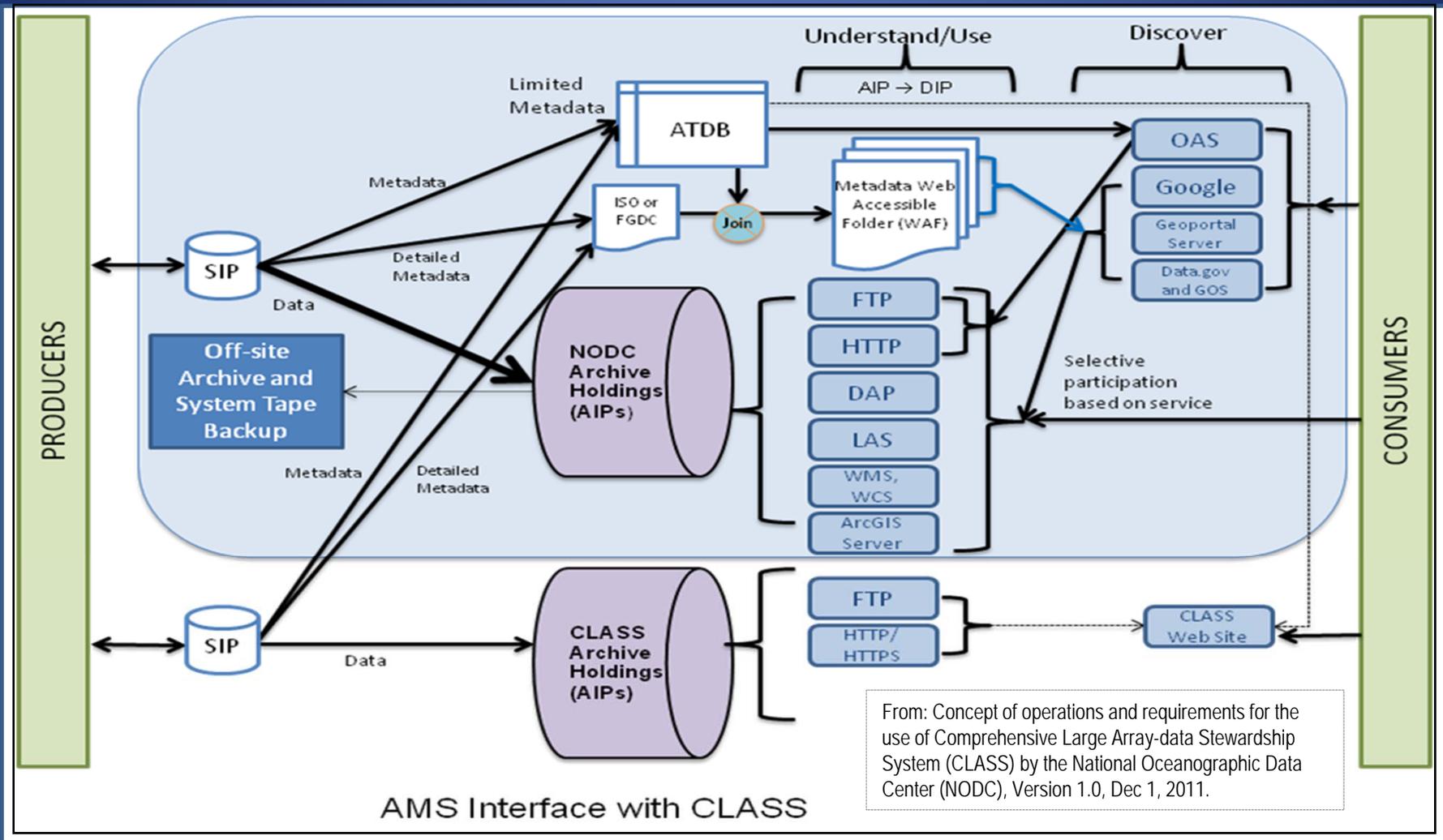
# AIPs in *canonical form*



# DIP Discovery and Access

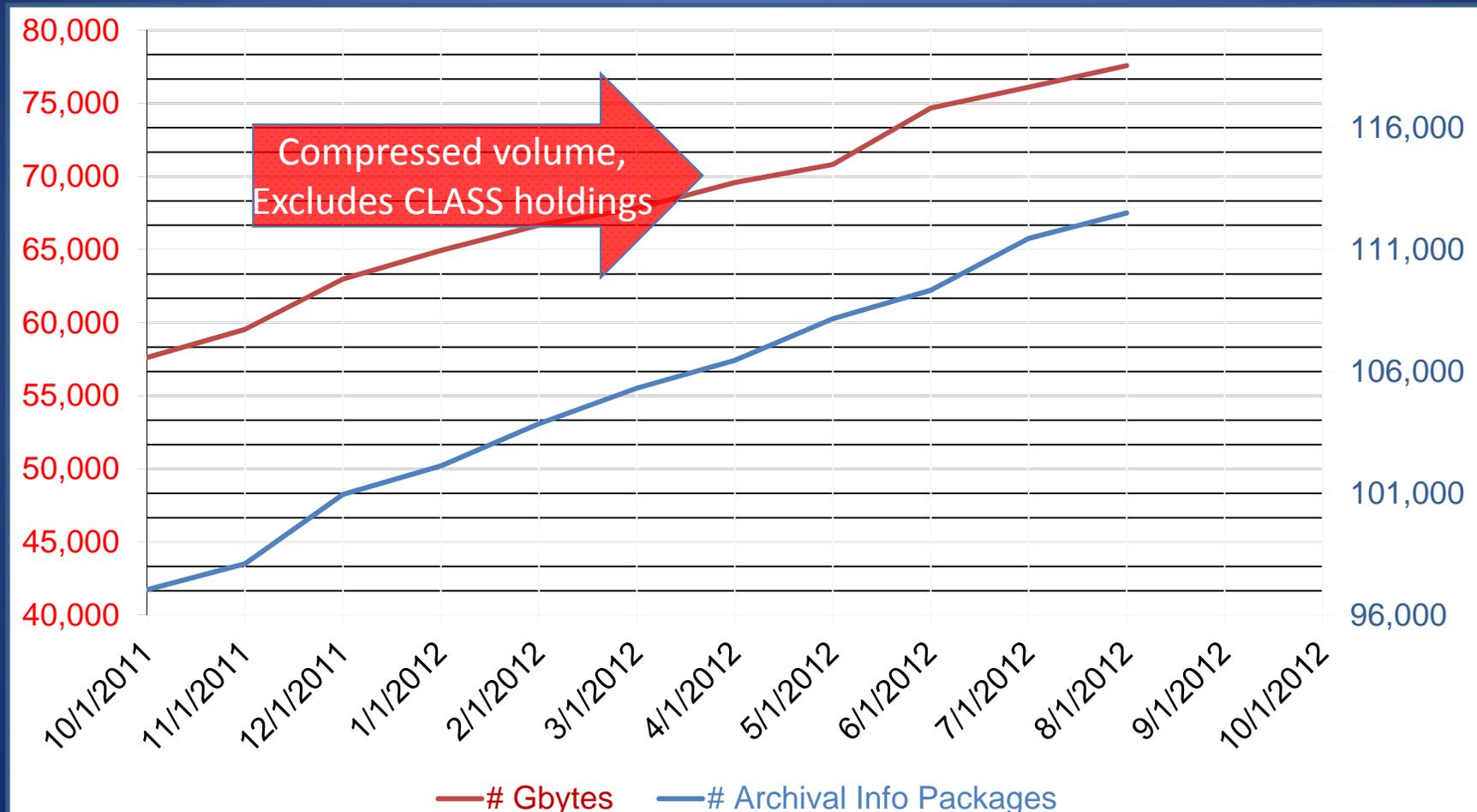


# Ongoing NODC-CLASS coordination



# Diverse holdings continue to grow

## NODC Archival Holdings Trend -- FY2012



# Automated ingest: Higher volume, lower diversity

- NODC Submission Information Form describes *automated, recurring* data submissions
  - Average **about 17,200 new AIPs per year**
  - Require sufficient details to enable automated acquisition and ingest
  - At least 15 operational automated ingest processes
    - BCO-DMO, NCEP WAVEWATCH III<sup>®</sup> and COFS/ROFS , GHRSSST
  - *In prep* data acquisition automation efforts include
    - UC Davis Bodega Marine Lab
    - NWS/National Data Buoy Center

# Non-automated ingest: lower volume, higher diversity

- Non-automated SIPs require *different levels of manual scrutiny* and ingest processing
  - Non-repeating, single time data submissions, highly variable types of data, metadata, formats, etc.
  - Average about **240 new AIPs per year**
  - Recommend metadata documentation guidance
    - NODC/NCDDC-developed MERMAid metadata editor and ISO transforms, ISO training courses
    - GEO-IDE and NGDC wiki guidance
    - NODC Electronic Data Documentation Form – elements map to FGDC elements
    - **Online Submission System** is in prep for Fall 2012 release

# File-level metadata and data standards

- Encourage data providers to **use netCDF for data**
  - NODC-developed netCDF templates for various types of geospatial data
  - <http://www.nodc.noaa.gov/data/formats/netcdf/>
- Encourage data providers to **use CF and ACDD metadata conventions in netCDF files**
  - Use ncISO to extract ISO metadata from netCDF files in AIPs
- Encourage FGDC or ISO format metadata for all SIPs
- **Provide** an FGDC formatted metadata record for all AIPs in Web Accessible Folders for discovery and access

***Metadata Golden Rule: Provide as much metadata for others as you wish others would provide for you!***

Thanks! Questions?

Donald.Collins@noaa.gov  
301.713.3272 x154

Archive data at NODC

See <http://www.nodc.noaa.gov/General/NODC-Submit/>  
or  
Send email to [NODC.DataOfficer@noaa.gov](mailto:NODC.DataOfficer@noaa.gov)

Find data at NODC

See <http://data.nodc.noaa.gov/geoportal/>  
or  
Send email to [NODC.Services@noaa.gov](mailto:NODC.Services@noaa.gov)