DWH LTDM

- Goals:
  - To ensure the legacy of the science from the DWH disaster is accessible to researchers studying future disasters in the region and around the world
  - To improve the quality of science coming out of the funds from the DWH disaster and ensure the research products are well documented and in stable formats

Data Archives

In the field of data management, the terms "archive" and "repository" often are used interchangeably. Within the Federal government, however, the term "archive" is specific to the mission and activities of the National Archives and Records Administration (NARA). Only NARA, or a Federal entity officially delegated by NARA for the long-term curation of specific products, should be referred to as an "archive."

Federal Archival

- From Open Archival Information System (OAIS)

  …an archive, consisting of an organization, which may be part of a larger organization, of people and systems that has accepted the responsibility to preserve information and make it available for a designated community. It meets a set of such responsibilities as defined in this International Standard, and this allows an OAIS archive to be distinguished from other uses of the term "archive". (from iso.org)

- from NOAA and U.S. National Archives and Records Administration (NARA)

  The NOAA National Data Centers are tasked with storing environmental data and making this data available to researchers, scientists, and anyone else that has a need for it, as well as in support of NOAA's mission. Destroy/delete 75 years after cutoff upon approval by NOAA and NESDIS stakeholders. A longer retention may be necessary for research purposes.

Data Types

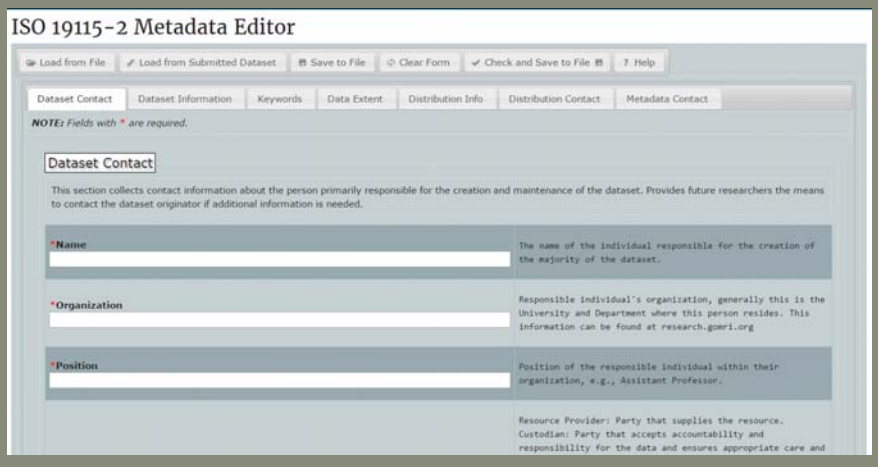
Baselines

- How do we know we are restoring to previous conditions?
  - What information is available to know that
- Comparable metrics and baselines for monitoring and restoration activities
- What is the new baseline for the GoM since DWH?
- Want to be able to look back at DWH data to answer questions for future spills

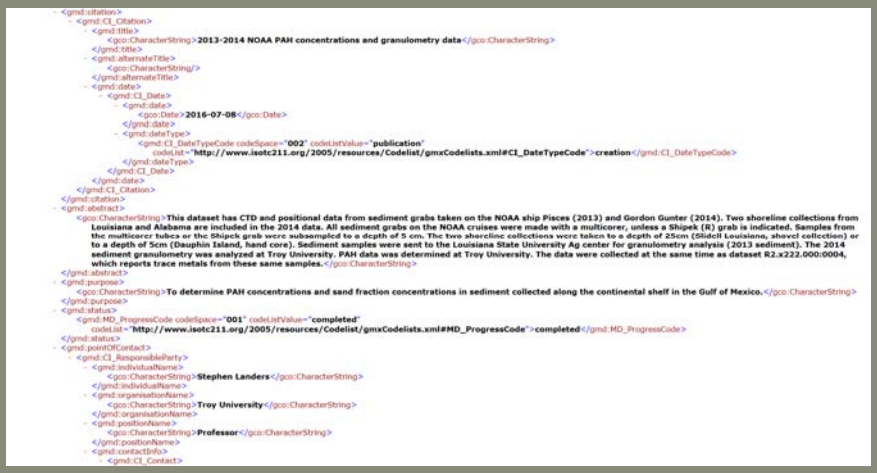
Metadata

- Documentation of data is essential to ensure that future users understand how the data was collected and who to contact with questions
- ISO 19115 standard- should be adopted as much as possible
  - This is what the federal government is using
  - Other standards should be able to be transformed into ISO
- Darwin Core could be considered for biological data, for specific repositories

Metadata Creation

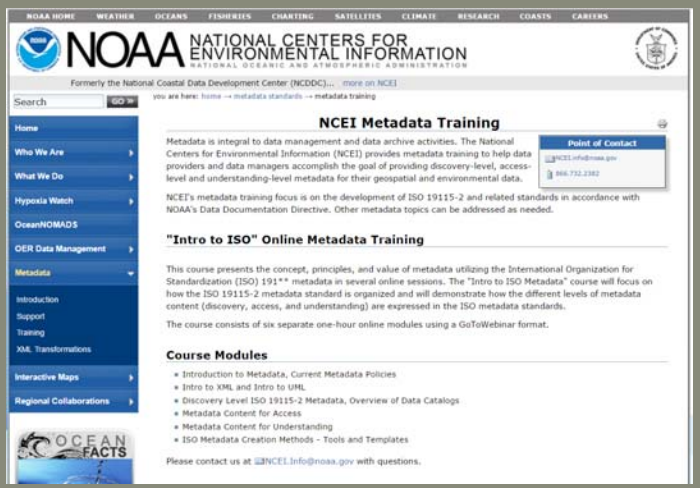


XML Metadata



```
- <gmd:citation>
  - <gmd:CI_Citation>
    - <gmd:title>
        <gco:CharacterString>2013-2014 NOAA PAH concentrations and granulometry data</gco:CharacterString>
      </gmd:title>
    - <gmd:alternateTitle>
        <gco:CharacterString/>
      </gmd:alternateTitle>
    - <gmd:date>
      - <gmd:CI_Date>
        - <gmd:date>
            <gco:Date>2016-07-08</gco:Date>
          </gmd:date>
        - <gmd:dateType>
            <gmd:CI_DateTypeCode codeSpace="002" codeListValue="publication"
              codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_DateTypeCode">creation</gmd:CI_DateTypeCode>
          </gmd:dateType>
        </gmd:CI_Date>
      </gmd:date>
    </gmd:CI_Citation>
</gmd:citation>
- <gmd:abstract>
    <gco:CharacterString>This dataset has CTD and positional data from sediment grabs taken on the NOAA ship Pisces (2013) and Gordon Gunter (2014). Two shoreline collections from
    Louisiana and Alabama are included in the 2014 data. All sediment grabs on the NOAA cruises were made with a multicorer, unless a Shipek (R) grab is indicated. Samples from
    the multicorer tubes or the Shipek grab were subsampled to a depth of 5 cm. The two shoreline collections were taken to a depth of 25cm (Slidell Louisiana, shovel collection) or
    to a depth of 5cm (Dauphin Island, hand core). Sediment samples were sent to the Louisiana State University Ag center for granulometry analysis (2013 sediment). The 2014
    sediment granulometry was analyzed at Troy University. PAH data was determined at Troy University. The data were collected at the same time as dataset R2.x222.000:0004,
    which reports trace metals from these same samples.</gco:CharacterString>
</gmd:abstract>
- <gmd:purpose>
    <gco:CharacterString>To determine PAH concentrations and sand fraction concentrations in sediment collected along the continental shelf in the Gulf of Mexico.</gco:CharacterString>
</gmd:purpose>
- <gmd:status>
    <gmd:MD_ProgressCode codeSpace="001" codeListValue="completed"
      codeList="http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#MD_ProgressCode">completed</gmd:MD_ProgressCode>
</gmd:status>
- <gmd:pointOfContact>
  - <gmd:CI_ResponsibleParty>
    - <gmd:individualName>
        <gco:CharacterString>Stephen Landers</gco:CharacterString>
      </gmd:individualName>
    - <gmd:organisationName>
        <gco:CharacterString>Troy University</gco:CharacterString>
      </gmd:organisationName>
    - <gmd:positionName>
        <gco:CharacterString>Professor</gco:CharacterString>
      </gmd:positionName>
    - <gmd:contactInfo>
      - <gmd:CI_Contact>
```

Metadata Training



NOAA HOME | WEATHER | OCEANS | FISHERIES | CHARTING | SATELLITES | CLIMATE | RESEARCH | COASTS | CAREERS

NOAA NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Formerly the National Coastal Data Development Center (NCDDC)... more on NCEI

Search GO

you are here: home → metadata standards → metadata training

- Home
- Who We Are
- What We Do
- Hypoxia Watch
- OceanNOMADS
- OER Data Management
- Metadata
  - Introduction
  - Support
  - Training
  - XML Transformations
- Interactive Maps
- Regional Collaborations

OCEAN FACTS

NCEI Metadata Training

Point of Contact
NCEI.Info@noaa.gov
866.732.2382

Metadata is integral to data management and data archive activities. The National Centers for Environmental Information (NCEI) provides metadata training to help data providers and data managers accomplish the goal of providing discovery-level, access-level and understanding-level metadata for their geospatial and environmental data.

NCEI's metadata training focus is on the development of ISO 19115-2 and related standards in accordance with NOAA's Data Documentation Directive. Other metadata topics can be addressed as needed.

"Intro to ISO" Online Metadata Training

This course presents the concept, principles, and value of metadata utilizing the International Organization for Standardization (ISO) 191** metadata in several online sessions. The "Intro to ISO Metadata" course will focus on how the ISO 19115-2 metadata standard is organized and will demonstrate how the different levels of metadata content (discovery, access, and understanding) are expressed in the ISO metadata standards.

The course consists of six separate one-hour online modules using a GoToWebinar format.

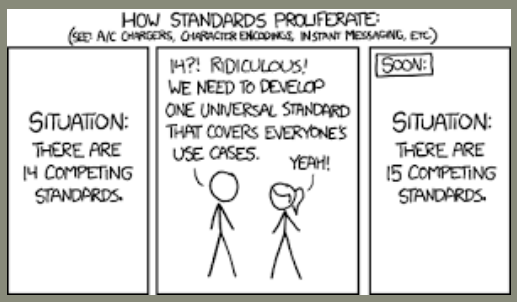
Course Modules

- Introduction to Metadata, Current Metadata Policies
- Intro to XML and Intro to UML
- Discovery Level ISO 19115-2 Metadata, Overview of Data Catalogs
- Metadata Content for Access
- Metadata Content for Understanding
- ISO Metadata Creation Methods - Tools and Templates

Please contact us at NCEI.Info@noaa.gov with questions.

Data Standards

- Identification of standards early in the process and get community buy in
- Standards need to be adequately communicated to data collectors



Federal Mandates

- The Digital Government Strategy and Open Data Policy were developed for the Government to better deliver information (data) and services.
- Federal agencies are under certain mandates that could inhibit data from being accepted if not properly formatted and documented
- *Common Framework for Earth-Observation Data*, March 2016, Office of Science and Technology Policy

- These standards need to be properly communicated to the data collectors

Data Sharing

- Length of time from collection to sharing
  - Real time data – data that is shared as soon as it is collected
    - Cruise data
    - Satellite data
    - Buoy data
  - Other data is shared depending on:
    - Funder/publisher requirements
    - Federal or state mandates
    - Requirements of other collaborators (foreign, private, industry, etc.)

Data Holds

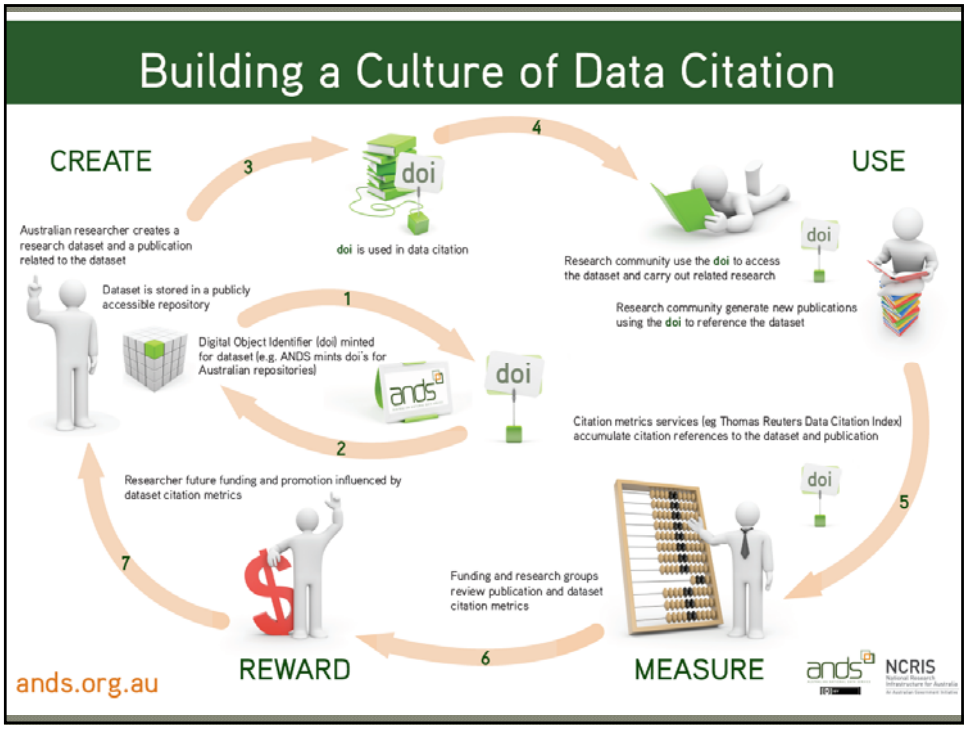
- If data is to be held for any reason the documentation of that data should begin before it is submitted for public access.
- Groups that start documentation before the data is collected have a leg up when the data is ready for publication
- Tracking of data from project onset is essential

Data Citation

- support proper attribution and credit
- support collaboration and reuse of data
- enable reproducibility of findings
- foster faster and more efficient research progress
- provide the means to share data with future researchers

Persistent Identifiers

- A Digital Object Identifier (DOI) is a commonly used type of identifier that is used to link to digital objects.
- Use of a persistent identifier makes data search and accessibility easier for future users
- Open Researcher and Contributor ID (ORCHID) is a persistent identifier for researchers



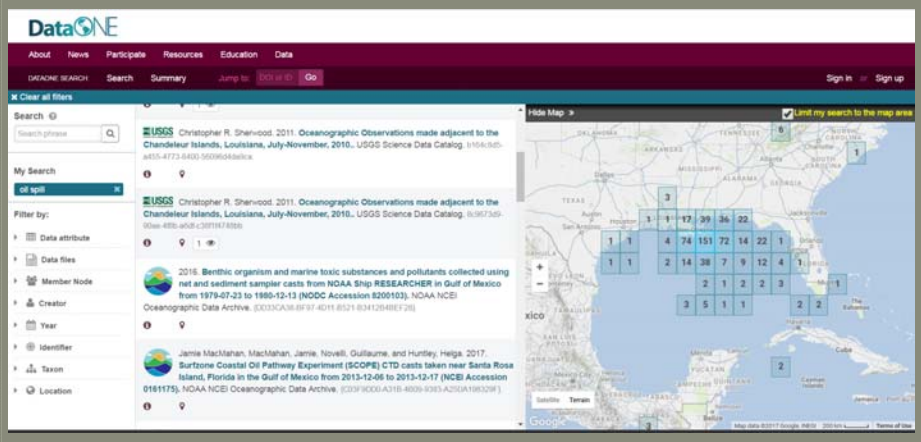
Data Access

- Providing services that allow users to retrieve data for exploration, analysis, or decision making
- Rely on sets of common standards and protocol (e.g. OPeNDAP, WMS, WCS, ERRDAP, FTP, SOS)
- Often community-driven
- Need for both human access and machine-to-machine access

Interoperability

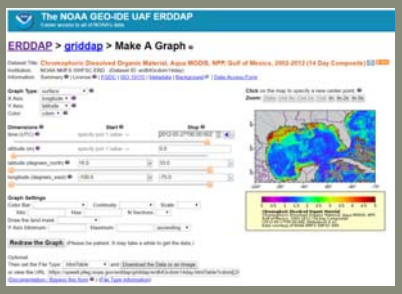
- Technology methods (examples):
  - OpenDAP
  - THREDDS
- Machine to machine data tools
- This allows for better and easier data synthesis

Distributed Data Frameworks

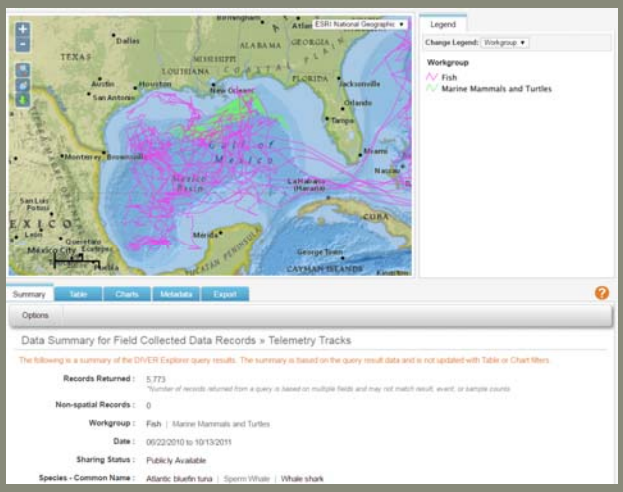


Data Visualization

- Good data management practices allow data to be easily reused and synthesized to develop useful products
- Display and manipulation of integrated data
  - ERDAPP
  - Cesium
  - ESRI
  - And many more



Telemetry Data



Visualization examples

- https://story.maps.arcgis.com/apps/MapSeries/index.html?appid=597d573e58514bdbbeb53ba2179d2359
- https://storm.pps.eosdis.nasa.gov/storm/cesium/EventViewer.html?position=-84.50,20.44,1199998&view=6.28,-1.00,6.28&fname=2A.GPM.DPR.V620160118.20170524-S141109-E144108.V04A.RT-H5