Federated Discovery in the Earth Sciences

R. Duerr
NSIDC affiliations and sponsorship

Cooperative Institute for Research in Environmental Sciences

University of Colorado Boulder

World Data System

Main sponsors:

- NASA
- National Science Foundation
- National Oceanographic and Atmospheric Administration

National Snow & Ice Data Center

1540 30th St
Boulder, CO 80303
(303) 492-6028
nsidc.org

Federated Discovery in the Earth Sciences
The National Snow and Ice Data Center...

- Provides tools for data access
- Manages and distributes scientific data
- Researches the cryosphere and data science
- Supports local and traditional knowledge
- Supports data users
- Educates the public about the cryosphere
The Early Days: NASA and the Version 0 System

- Early 1990's NASA developed a system that allowed users to search and order data across all of its DAACs
NASA's Version 0 System

- Catalog queries went to the GCMD
- Inventory queries went to all repositories in parallel based on:
  - Source
  - Sensor
  - Parameter
  - Geographic region
- Users could request browse images of selected items
- Users could request items selected from the inventory results
NASA's Version 0 System

Federated Discovery in the Earth Sciences
NASA's Current System (Reverb/ECHO)

- Catalog queries still go to the GCMD
- Inventory queries go to a central inventory
- Orders go to the DAACs
**NASA's Reverb/ECHO System**

Federated Discovery in the Earth Sciences

Diagram showing the integration of GCMD, ECHO, Reverb, and DAAC A and DAAC B with inventory order and catalog query/response processes.
Why does this work for NASA?

- They have control over the metadata!
- Only two metadata standards (catalog, inventory) to support and even those are being harmonized.
Beyond NASA

- A variety of Earth science data aggregation portals are out there
- Most make folks provide metadata in their standard format

from https://cn.dataone.org/onemercury/
CNR's Broker

- CNR's broker uses accessors to translate heterogenous data into other forms for discovery and access.
- It too uses a common metadata model under the hood.

from [http://essi-lab.eu/do/view/GIcat/GIcatDocumentation](http://essi-lab.eu/do/view/GIcat/GIcatDocumentation)
How well does this work?

- Manual configuration required
- You need to know about the resource to be able to add it

Federated Discovery in the Earth Sciences
How well does this work?

Without significant configuration, only simple capabilities can be provided.
How well does this work?

### Sea Ice Index
- **Temporal Coverage**: 1979-01-01 to continuous
- **Parameter**: Ice Extent, Ice Growth/Melt, Reflectance, Sea Ice Concentration
- **Data Format**: ASCII (txt), ESRI Shapefiles (.shp, .shx, .dbf, and .prj), PNG
- **Summary**: The Sea Ice Index provides a quick look at Arctic- and Antarctic-wide changes in sea ice. It is a source for consistent, up-to-date sea ice extent and concentration images and data. [More Detail](#)

### Multisensor Analyzed Sea Ice Extent - Northern Hemisphere (MASIE-NH)
- **Temporal Coverage**: 2006-10-01 to continuous
- **Parameter**: Ice Edges, Ice Extent, Ice Growth/Melt
- **Data Format**: ASCII (txt), ESRI Shapefiles (.shp, .shx, .dbf, and .prj), Geographic Tagged Image File Format (GeoTIFF) (.geotiff), Keyhole Markup Language (.kml), Microsoft Excel, NetCDF, PNG
- **Summary**: NOTICE: The data providers have released a new version of the input data for MASIE. We are working to incorporate this new version into our processing system and plan to do so... [More Detail](#)

### Sea Ice Trends and Climatologies from SMMR and SSM/I-SSMIS
- **Temporal Coverage**: 1978-10-26 to 2013-12-31
- **Parameter**: Ice Extent, Sea Ice Concentration
- **Data Format**: ASCII (txt), Binary, Joint Photographic Expert Group (JPEG) (.jpg), PNG
- **Summary**: Notice Regarding Data Corrections 04 September 2013 These data sets were recently revised. Please refer to the Documentation link provided above for details regarding data... [More Detail](#)

### Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data
- **Temporal Coverage**: 1978-10-01 to 2013-12-31
- **Parameter**: Sea Ice Concentration

---

Federated Discovery in the Earth Sciences
How well does this work?

It doesn't scale!!!
What's next?

- Adding web crawling to automatically find resources
- Attempting to automatically configure brokers with found resources
- Replacement of a static metadata store with a semantic store containing whatever information can be deduced from the data and services detected
Questions?