Data Integration within US IOOS

Derrick Snowden, US IOOS
EDMC/Data Integration Projects Session
16-May-2012
U.S. IOOS: Program Overview

WHO

WHAT

- Observations
- Data Management
- Modeling and Analysis

WHY: 7 Goals, 1 System
- Weather and climate change
- Maritime operations
- Natural hazards
- Homeland security
- Public health risks
- Healthy coastal ecosystems
- Sustain Living Marine Resources

Enhances science and improves decision making
What do we mean by integration?

http://nvs.nanoos.org shown above, is one of eleven RAs and integrates 167 assets from Federal, Tribal, State, County, University and Commercial organizations as well as international partners.
Data Integration supports frameworks we can build useful services upon.

Illustrates the Publish, Find, Bind Pattern. OGC Reference Model (OGC 08-062r7)
Technologies applied to Data Services

- netCDF – Climate and Forecast conventions – OPeNDAP
- THREDDS (data cataloging and distribution with a flexible plugin environment)
- Hyrax (.opendap.org)
- Sensor Web Enablement (Framework/family of services, encoding standard, family of services)
- SOS (OGC Sensor Observation Service)
- ERDDAP
- W*S (OGC Web * Service, primarily WMS)
Technologies applied to Data Services: the short list

OPeNDAP

OGC®
Making location count.
Technologies applied to Data Registries/Catalogs

- ISO 191** Geospatial metadata standards
- ncISO – Harvester supporting netCDF ➔ ISO 191** translation
- ESRI Geoportal Server
- GI-CAT, Geonetwork
- Opensearch, OGC CS/W
- ERDDAP
- THREDDS
Technologies applied to Service Clients

- nctoolbox
- Environmental Data Connector
- ERDDAP
- Javascript library for SOS parsing
- Python library for SOS parsing
Clients in action

```python
>> links = opensearch(q)
>> nc = ncugrid(links.dap{1})
>> z = nc.data('zeta',…)
>> grid = nc.grid('zeta',…)
>> plot(grid)
```

Built on open source software supported by NSF/Unidata, NOAA, USGS, and several industrial partners.
Putting it all together.
Software as the deliverable, advances the technologies faster than pilot projects.

• nctoolbox - [http://code.google.com/p/nctoolbox](http://code.google.com/p/nctoolbox)
• 52North SOS - [https://github.com/axiomalaska/52north-sos](https://github.com/axiomalaska/52north-sos)
• ncSOS - [https://github.com/asascience-open/ncSOS](https://github.com/asascience-open/ncSOS)
• ncISO - [http://www.ngdc.noaa.gov/eds/tds/](http://www.ngdc.noaa.gov/eds/tds/)
• Environmental Data Connector - [http://www.pfeg.noaa.gov/products/edc/](http://www.pfeg.noaa.gov/products/edc/)
• pyoos - [https://github.com/asascience-open/pyoos](https://github.com/asascience-open/pyoos)
• NetCDF Java Library – Unstructured - [https://github.com/asascience-open/NetCDF-Java-UGRID](https://github.com/asascience-open/NetCDF-Java-UGRID)

NOTE: IOOS did not fully support any one of these software development efforts, but by leveraging open source tools some of the integration objectives have been achieved.
Challenges

• Open Source!!! We need to encourage developers working on NOAA issues to utilize Open Source tools and technologies as they develop, not just as they consume.

• Data Catalog standards/software lag behind service registry standards. We want more sophisticated queries that reveal resources about platforms, not just high level data set information.

• Standards landscape is more sophisticated than any one group or developer can effectively track. Code sharing is great but is it enough?