# **IOOS** Catalog

Presented by Luke Campbell and Micah Wengren



Micah Wengren

• Coming to IOOS from NOS' Office of Coast Survey Development Laboratory (NOAA nautical charting)

• Supported NOAA Data Catalog (data.noaa.gov) implementation

•FOSS4G GIS/Geodata management background

• Early NOAA work on weather/ocean observation and forecast portal, nowCOAST (nowcoast.noaa.gov)



# Background

Previous catalog:

- NGDC Geoportal
- Github and Emails
- System Monitoring
  - <u>catalog.ioos.us</u>

				5									Lond
Collec	tion Source Lis	t											
Filter by:	Record Set Name:	NOAMIOO	S/MARACOOS	¥ ÷	Component:	- IA-			Servi	ce Type:	-Al -		•
	Status:	-All -		0	Service URL:						List Collection Se	LUCCER	
	Record Set	Component	Svc Type	Svc Status	Sve Uri			Update Date	Submit Date	Notes			
Show	NOAA/IOOS/MARACOOS/	Dave Foster - pointOlContact - Applied	THREDDS	REMOVED	http://sos.maracoos.org/stable/	agg_catalogs/weatherflow_agg_	catalog.xml	2015- 09-13 00:00:00 MDT	2014- 03-12 00:00:00 MDT	updated sv	type to THREDDS	- rbr	
Show	NOAA/IOOB/MARACOOB/	Dave Foster - pointOfContact - Applied	THREDDS	REMOVED	http://sos.maracoos.org/stable/	lagg_catalogs/sldmb_agg_catalo	g.xml	2015- 09-13 00:00:00 MDT	2014- 03-12 00:00:00 MDT	updated sv	type to THREDDS	- rbr	
Show	NOAA/IOOS/MARACOOS/	Dave Foster - pointOlContact - Applied	THREDOS	REMOVED	http://sos.maracoos.org/thredd	s/catalog.xml		2015- 09-13 00:00:00 MDT	2014- 03-12 00:00:00 MDT	manual test failed:HTTP http://sos.m	returned "opendap. /1.1 502 Proxy Error aracoos.org/thredd	dap.DAP2Ex on URL= s/dodsC/agg	ception: Method /navhycom/agg.nc
Show	NOAA/IOOS/MARACOOS/	Dave Foster - pointOlContact - Applied	909	REMOVED	http://sos.maracoos.org/stable/	isos/wflow1127-agg.nomi		2015- 09-13 00:00:00 MDT	2014- 04-03 00:00:00 MDT	NULL datas	et response. (am)		
Show	NOAA1005/MARACOOS/	Dave Foster - pointOfContact - Applied	SOS	REMOVED	http://sos.maracoos.org/stable/	'sas/wflaw307-sgg.ncml		2015- 09-13 00:00:00 MDT	2014- 04-03 00:00:00 MDT	NULL datas	et response. (am)		
Show	NOAA/IOOS/MARACOOS/	Dave Foster - pointOlContact - Applied	SOS	REMOVED	http://sos.maracoos.org/stable/	'sos/wflow44654-agg.nomi		2011- 09-13 00:00:00 MDT	2014- 04-03 00:00:00 MDT	NULL datas	et response. (am)		
Show	NOAA/IOOS/MARACOOS/	Dave Foster - pointOlContact - Applied	SOS	REMOVED	http://sos.maracoos.org/stable/	isos/wflow46182-agg.nomi		2015- 09-13 00:00:00 MDT	2014- 04-03 00:00:00 MDT	NULL datas	et response. (am)		
Show	NOAA/DOS/MARACOOS/	Nickitas Georgas - pointOlContact	THREDDS	REMOVED	http://colossus.dl.stevens-tech.	edu/thredds/complete_latest.xm	Ê.	2015- 09-11 00:00:00	2014- 04-18 00:00:00				

- 1.Data Provider sent an email or created an issue on github for a new data source.
- 2. Anna would get notified.
- 3.Rob or Anna would review the source for simple issues
- 4. Anna would create a harvest job for NGDC's harvesters to synchronize source
- 5.NGDC Geoportal would get updated.



# Defining Goals for the Future



# **Defining Goals for Future**

# Identifying what didn't work for us:

### Synchronization:

- Partners and RA's spent too much energy and resources attempting to synchronize datasets with the IOOS Catalog
- Data is fluid but the catalog is mostly static



### Our Goals:

- Smooth Synchronization
- Low Effort Maintenance
- High Visibility
- Tools

### Where we are





### What is it?

### **Feature Overview**



Complete catalog system with easy to use web interface and a powerful API



Strong integration with third-party CMS's like Drupal and WordPress



Data visualization and analytics



Workflow support lets departments or groups manage their own data publishing



### Fine-grained access control

Ē	
_	

Integrated data storage and full data API



Federated structure: easily set up new instances with common search





### Why do we use it?

- Python based web development project
- Highly Extensible
- Wide community support
- Very active project



### Why do we use it?

- PyCSW
- Spatial Awareness
- API



### How it works





# How it works

### **IOOS** Catalog







### **IOOS** Service Monitor





### **IOOS** Service Monitor

- Internal Tool
- Checks service availability and uptime
- Future: metadata compliance checking



sos.maracoos.org			Services Available: 854/854				
MARACOOS	DAP	HRECOS Aggregated Station HRTVBM Data	Harvest Successful	8 hours ago	19/20		
MARACOOS	DAP	HRECOS Aggregated Station HRWSTPTH Data	Harvest Successful	8 hours ago	19/20		
MARACOOS	DAP	HRECOS Aggregated Station HRPMNTM Data	Harvest Successful	8 hours ago	19/20		
MARACOOS	DAP	HRECOS Aggregated Station HRPMNTH Data	Harvest Successful	8 hours ago	19/20		
MARACOOS	DAP	HRECOS Aggregated Station HRPIER84 Data	Harvest Successful	8 hours ago	19/20		



### Central WAF

#### Database CKAN CSW Database Database OPeNDAP WAF Process or Function CKAN Central WAF CKAN **PyCSW** WMS Harvesters Spatial Plugin User Interface > CSW SOS CKAN Web Interface Data Source Central WAF IOOS Catalog RA Hosted Web Registration Harvest Service Monitoring Interface Configuration Database IOOS Service IOOS Service Monitor Web Monitor Interface IOOS Registry IOOS Service Monitor



### Central WAF

- Stores ALL ISO
   records for IOOS
- Synchronized through harvesters
- Configurable
- Future: Hierarchical



### **IOOS** Registry





IOOS Registry (Future)

- Web interface for data providers to set up harvesting
- Receive notifications for errors and harvesting issues
- Metadata Checker



### **RA** Hosted





# How it works

RA and Partner Hosted WAF

- Expose datasets and services through a WAF
- Simple to create, simple to manage, simple to harvest
- CSW for the more advanced solutions



### Harvesting Process Now:

- Manual curation of ISO records in central WAF at <u>http://data.ioos.us/waf/</u>
- CKAN harvesting of all records in WAF
- Changes to WAF are done by hand



# Where we're heading

The short term



# Where we're heading

- Providers Responsibilities
  - Create and publish a WAF
  - Hosting a CSW
- IOOS Catalog Responsibilities
  - Automated Harvesting from RA provided WAFs
  - Better metadata controls
    - Improving CKAN-Spatial for our ISO records

- Pursue parallel development with NOAA Catalog, open source on github
- Better representation of our data within CKAN

sp025-20150108T1921	🚠 Dataset 🔮 Groups	⊙ Activity Stream							
Organization	sp025-20150108T1921								
GliderDAC IOOS National Glider Data Assembly Center read more	Spray glider profile data from Scripps Institution of Oceanography Instrument Development Group (supported by NOAA).         Access       Description         Coverage       Keywords       Constraints								
	Online Access	Download This URL provides a standard OPeNDAP html interface for selecting data from this dataset. Change the extension to .info for a description of the dataset. OPeNDAP (DAP Client Access) These data are available for access with an OPeNDAP client OGC:SOS (OGC Sensor Observation Service) These data are available for access with an SOS client							
	Distribution Service Type	THREDDS OPENDAP							
	Distributor	Scripps Institution of Oceanography							
	Dataset Point of Contact	Scripps Institution of Oceanography							

### Where we're heading

### • Service Monitoring from CKAN sources

sos.maracoos	s.org		Services Available: 854/854			
MARACOOS	DAP	HRECOS Aggregated Station HRTVBM Data	Harvest Successful	8 hours ago	19/20	
MARACOOS	DAP	HRECOS Aggregated Station HRWSTPTH Data	Harvest Successful	8 hours ago	19/20	
MARACOOS	DAP	HRECOS Aggregated Station HRPMNTM Data	Harvest Successful	8 hours ago	19/20	
MARACOOS	DAP	HRECOS Aggregated Station HRPMNTH Data	Harvest Successful	8 hours ago	19/20	
MARACOOS	DAP	HRECOS Aggregated Station HRPIER84 Data	Harvest Successful	8 hours ago	19/20	



### To the future and beyond

Longer term solutions



Registry tools

- Web tool for managing registered endpoints
- Harvest notifications
- Automated compliance checking

DIOOS | Integrated Ocean Observing System

 Compliance Checker

 Your dataset scored 223 out of 241 points

 During the cf check

 For dataset 20160505.nc

 Scoring Breakdown:

 High Priority (a)

 Name
 Score

 2.2 Valid netCDF data types
 14/15

 2.3 Legal variable names
 15/15



### To the future and beyond

### Catalog Integrations



# To the future and beyond

### The Big Picture





### Walkthrough and Questions

