

Regional IOOS Catalog Workshop

Goal Statement

The primary goals of the workshop are as follows:

1. agreement on the content for a dataset of IOOS-relevant platform and sensor information,
2. agreement on a standardized method for making this information available, and
3. agreement on roles and responsibilities for updating and aggregating the dataset(s).

Secondary goals include the identification of additional uses of the dataset(s), including value-added efforts, and identification of tools and methods that will allow for an automated method or methods for collecting and maintaining the dataset(s).

Boundaries and Expectations

The primary goals of the workshop intend to address an immediate and narrowly defined need. An ability to identify and visualize *who* has *what* observational sensors in the water, and *where* those sensors or platforms are located is the need. This need has been expressed by a number of different entities, including Regional Associations, Ocean.US, the U.S. GOOS Steering Committee, and NOAA.

Currently there are many groups collecting this information in an *ad hoc* and often repetitive and disorganized manner, both on regional and national scales. This effort will attempt to bring a measure of minimum standardization to the process so that information is collected and shared in the same fashion, and thus becomes usable by anyone who cares to do so.

Also, please be aware that additional broad scale efforts are underway that will likely address similar issues in a much more comprehensive and systematic fashion. This particular effort is intended to address an immediate need, and can be completed with a modest investment. To the extent possible, we will make full use of past or ongoing efforts that can inform this process.

Confirmed Participants for Regional IOOS Catalog Workshop

Participant Name	Affiliation	Lodging
Matt Howard	GCOOS/Texas A&M	Falmouth Inn
Jeremy Cothran	Caro-COOPS/SEACOOS/University of South Carolina	Falmouth Inn
John Ulmer	NOAA CSC	Inn on the Square
Daniel Martin	NOAA CSC	Local
Jim Boyd	NOAA CSC	Inn on the Square
Otina Fox (Tuesday)	AOOS	Falmouth Inn
John Kelley	NOAA CSDL	?
Stuart Eddy	GLOS/Great Lakes Commission	Falmouth Inn
Dave Forrest	VIMS	Falmouth Inn
Anne Ball	NOAA CSC	Inn on the Square
Eric Bridger	GoMOOS	Falmouth Inn (driving in Monday AM, so Monday night only)
Vembu Subramanian	COMPS/University of South Florida	Inn on the Square
Josh Pederson	NOAA	Falmouth Inn
John Graybeal	MMI/MBARI	Nautilus Motor Inn
Julie Bosch (Tuesday)	NOAA NCDDC	Inn on the Square
Lisa Hazard	SCCOOS/Scripps	Inn on the Square
Janet Fredericks	MVCO/WHOI	Local
Charles Seaton	CORIE/Oregon Health and Sciences University	Falmouth Inn

***Regional IOOS Catalog Workshop
June 19-20, 2006
Woods Hole, MA
Draft Agenda***

Monday, June 19

8AM-9AM	Introduction <ul style="list-style-type: none"> • Goals, objectives, expectations, and boundaries • Catalog/registry with potentially multiple uses and users • Minimum content standard to support the main goal
9AM-10AM 10AM-11AM	Describe/Discuss Scope and Assumptions <ul style="list-style-type: none"> • List assumptions and describe scope • Discuss and refine project boundaries
10AM-10:15AM	BREAK
10:15AM-11AM	Describe/Discuss Scope and Assumptions (Continued)
11AM-Noon	Three Overviews <ul style="list-style-type: none"> • OceanObs – Josh Pederson • Inventory Database – Vembu Subramanian • Marine Metadata Initiative (MMI) – John Graybeal
Noon-1:30PM	LUNCH
1:30PM-3:30PM	Content Standard <ul style="list-style-type: none"> • List of platform/sensor characteristics that need to be handled • Controlled vocabulary • Define vocabulary terms
3:30PM-3:45PM	BREAK
3:45PM-5:30PM	Content Standard (Continued)
5:30PM	ADJOURN

Tuesday, June 20

8AM-8:15AM	Overview for the Day
8:15AM-10AM	Implementation <ul style="list-style-type: none">• Definition of local data feeds (immediate, mid-term, long-term)• Back end data feeds from local observatories to aggregation point• Front end map display
10AM-10:15AM	BREAK
10:15AM-11AM	Implementation (Continued)
11AM-Noon	Implementation Process – Roles and Responsibilities <ul style="list-style-type: none">• Short term manual process – Stage 1 is low tech• Mid to long term – move to automated process – Stage 2 web form w/ authentication to enable registration of a local obs catalog• Custodian – who runs the registry?
Noon-1:30PM	LUNCH
1:30PM-3:30PM	Implementation Process – Roles and Responsibilities (Continued)
3:30PM-3:45PM	BREAK
3:45PM-5:30PM	Future Planning <ul style="list-style-type: none">• Where does this go from here?• Continued coordination and communication?
5:30PM	ADJOURN WORKSHOP

Regional IOOS Catalog Workshop Draft Session Outlines

Session 2: Scope and Assumptions

Session Goal: *Reach consensus on design principles and boundaries.*

Prioritize on the IOOS core variable list to direct data acquisition efforts.
Prioritize long term fixed instruments and any real-time variable in core list.
Prioritize regional or sub-regional systems of platforms and/or sensors.
Utilize a feature-based approach, one record per logical sensor
Identify regional data steward.
Define an update rate and release schedule.
Seek "buy-in" from many providers.
Define a scope for the data set (e.g. federal or non-federal , RA, sub-regional, local).
Describe our authority or relationship to other efforts. Is this IOOS or an independent effort and how do we describe our participants and the scope of the system?

Start with a focus on the inventory database design (data content model and ERD).
Use mature database design principles, like code-lists and fixed domains.
Identify 1 or 2 individuals for data aggregation role.
Include a field(s) for data coverage.
Utilize ISO 19115 field types where possible.
Utilize comma or tab delimited format for stage 1 local release.
Provide an Excel template to assist in early data collection to create comma or tab delimited.
Plan for future automated harvest of XML content at sub-nodes.
Change control is a local client responsibility?

Deploy at an assembly center using the OGC WFS.
Develop a style sheet for the non-spatial on-line browse needs.
Select Browser, Excel and ArcGIS as primary client tools.
Develop one metadata record for the data set for GOS submittal.

What have we neglected/forgotten?

Session 4: Content standard

Session Goal: *Develop a list of core variables, their domains, and definitions and their relationships to each other.*

element	definition	data type	source	iso	dublin core
version	identifies inventory version	numeric			
source	platform owner		CSW		Y
sensor identifier	unique platform/sensor identifier				
subject		CharacterString	CSW	Y	Y
title		CharacterString	CSW	almost	Y
abstract		CharacterString	CSW	Y	Y
format		codelist	CSW		Y
identifier	unambiguous reference	identifier	CSW		Y
modified	date resource was last changed	date-8601	CSW		y
type		Code list	CSW		Y
location	x, y location in CRS units				
place name		CharacterString			
bounding box		boundingbox	CSW		
CRS	coordinate reference system (authority and ID) for the bounding box		CSW		Y
data url					
ioos variable(s)		code list			
collection frequency	period of data collection				
boundingbox type					
westboundlongitude	decimal degrees -positive east	numeric			
southbound latitude	decimal degrees -positive north	numeric			
eastboundlongitude	decimal degrees -positive east	numeric			
northboundlatitude	decimal degrees -positive north	numeric			

Session 5: Implementation (What are the players doing?)

Session Goal: *Develop a list of the activities required to implement a system in the short, medium, and long term.*

SHORT TERM Activities (0 - 6 months)

Provide for short term project maintenance/communication.
Provide for outreach to local observatories not present at the workshop.
Complete data model (resolve terms, domains, structure).
Identify local data manager (contacts) & sensor nodes (FTP or Web site holding data).
Define the scope and authority of this effort.
Develop a data flow diagram.
Settle on a data format (CSV, Tab delimited, XML, Excel??).
Develop a QC process for data exchange.
Develop a local publish process (email, FTP, HTML, secure/insecure).
Develop ingest and aggregation mechanism.
Write a metadata record describing the aggregated data set.
Develop a registry of sensor nodes (FTP, Web sites, or Web Services where data is available).
Develop default web service for publication of the aggregated data set.
Develop default web application for presentation.
Develop XSLT/Stylesheet to convert WFS to output for non-GIS users (e.g. browsers).
Consider synchronizing records with others.
Develop instructions for participation and technical outreach.
Contribute documentation to the CIR (give some longevity to the effort).

MID TERM Activities (6-18 months)

Provide for long term project maintenance/communication.
Develop local data record development tools (e.g. spreadsheet with pull downs and controlled lists).
Develop or borrow XML schema for local exchange.
Develop version 2 aggregation service to handle new local data exchange mechanisms.
Develop default web service for publication.
Develop synchronization plan.

LONG TERM Activities (18+ months)

Online data record development and management including transaction support.
?ML – Consider how we would (or, should we) converge with other relevant XML dialects.

Session 6: Implementation Process - Roles and Responsibilities (Who is doing it when?)

Session Goal: Attach resources to the tasks identified in session 5.

Implementation may be carried out in a staged pattern starting with simple, low cost mechanisms in the short term and progressing to more automated and sophisticated mechanisms in the mid term, and finally to a mature system in the long term.

Responsibilities:

- Project Management (periodic calls, emails, and reporting to the group)
- Data Provision (local observatories)
- Data Assembly/Aggregation
- Presentation Web Application development and hosting.
- Outreach: reporting to other local observatories
- Reporting to NFRA, and DMAC

Schedule: Plan session 5 tasks for short-term and mid-term work.

Session 7: Future Planning

Session Goal: Briefly consider how to carry this work into the future.

Where does this go from here?

Continued coordination and communication?

Convergence with other efforts?

Contributions to IOOS, CIR, and DTL?

Support from CIR and DTL?

Regional IOOS Catalog Workshop Notes

Day 1

Scope and Assumptions session:

- Add in the idea that should be available at the most modest shops -- Low barrier for participation
- Staged approach – short term, mid, long – focus on short
- Simple data model and how to make that data available (format)
- Do we keep it to the core variable list or not? Maybe not a small list of variables... Whatever the system wants to report – may need to do some filtering and/or tweaking to accommodate.
- Is it sensors or variables? Variables is what we are concerned with.
- Requests should come through RAs
- Difference in the complexity of information needed – RAs need the more complex info (sensor resolution, quality, etc.), and we just need the parameters measured where
- Need to define everything (i.e., so that “long term” means the same timeframe for everyone)

Data Content session:

What variables are you measuring and where?

- Show funding agencies what we have or need
- Let regional folks make resource investment decisions

This is a project management system and not data management system

See Excel file (Data Content_Inventory) for details of data content...

Day 2

Implementation Session:

- Web form for low level users
- XML or CVS format that will be spit out of the Web form or developed directly
- Local files that will be pulled, not pushed
- Work through the NFRA and RAs – PIs and Coordinators
- Registry information could be kept at RA level
- Solicitation should include the info that this is the first step in a process of beginning to collect various types of information
- Aggregator issues

- Aggregator keeps history of what the individual systems look like through time?
- QC
- Flat file/public
- Response – status feedback
- Synchronized/duplicate records
- One national FGDC record
- CSC to keep talking with Josh, Vembu, and John K. about who and how to aggregate and/or publish
- Publish
 - Visual
 - National map
 - Viz at local level immediately to QC points and get a “reward” for participating – KML file?
 - Broker – draw bounding box, get list of data URLs, use some service to GetLatest data (Matt H)
 - GIF compatible for resource managers, etc... - WFS
 - NowCOAST will use this as publish
 - API
 - CSW
 - WFS
 - KML development/demonstration
 - SOAP

Roles and Responsibilities Session:

1. Project coordination – CSC, Matt Howard, Stuart Eddy
2. Finalize content standard and file format
 - John Ulmer, Daniel Martin, Anne Ball, Julie Bosch, John Graybeal, Jeremy Cothran
2. Host/maintain registry (URL, contact, email)
 - CSC
3. Devise observation identifier - arbitrary key, ability to map to other systems
 - Dave Gilhousen, John Kelley, Jeremy Cothran, John Graybeal
3. Host aggregator – maintain history, QC data, flat file output, cut/paste validator
 - CSC, and talking w/ Josh Pederson and Vembu Subramanian
4. Build/host web-based file builder – linked to the data checker/plotter/gift
 - Josh Pederson, Jeremy Cothran
4. Data checker/plotter for local data sets – Google Maps/Earth
 - Lisa Hazard
 - Jeremy Cothran
4. Host national map view – WMS/WFS, support basic queries
 - John Kelley, Eric Bridger, Josh Pederson
5. Outreach/Letter to RAs – warning and guidance document
 - Matt Howard
6. Write/post one FGDC metadata record for entire dataset

- Josh Pederson, CSC
- Spreadsheet record builder?
 - CSC

Standing Issues/Parking Lot:

- Local burden
- Data file format
- Field names blessed
- Attribute domains
- Domain for measured parameters
- Desktop file validator to give immediate ability to proof the file structure and base content
- Need to make provision for the HF radar and ADCP – bounding box with the z measurement

Future Directions Session:

- ACT Interoperable Sensors mtg – SensorML – Tom Shyka
- MMI in Oct/Nov – J. Graybeal
- OOS Tech in the fall
- DMAC regional caucus has \$20K for meeting

Regional IOOS Catalog Content Model

PHASE 1				
	element	definition	data type	notes
?	author	email?		
r	version	version of data model used	numeric	
r	operator	owner/operator of the (platform/sensor) collecting a variable(s)	codelist	
r	regional association	name of regional association in which the operator participates.		
?	sensor package identifier	unique platform/sensor id		tabled - we know we need a unique identifier for each observation but have not settled on precisely how to achieve that effect.
r	observation identifier	observation identifier:resolves multiple instances of a variable at a single logical sensor.	identifier	
r	modified	date this record was last changed	date-8601	
r	local platform name	local observatory's human readable name for the platform	charString	the label to go on plots of this data.
?	location key word	place name from a controlled list	charString	from geographic name info service
r	latitude	decimal degrees, WGS84		location is logically 1 entry of latitude and longitude
r	longitude	decimal degrees, WGS84		
r	z	altitude with positive is up		need units and length
r	plane of reference	baseline for altitude reference	codelist	rename to vertical datum?
r	bounding box	coverage of measurements	boundingbox	use GML standard
r	planned start date	planned or actual launch date		
r	planned end date	planned or actual end date		
o	operator url	organization/operator URL		
o	platform url	online description of the platform/sensor		
o	data url	online access address		
r	observation name	variable name of observation	codelist	add observation type?
r	observation status		codelist	operational planned offline
?	comments	free text entry	charString	rename to abstract to align with existing standards?

r = required

o = optional

element sequence in this table is arbitrary