

Robert C. Groman, Molly D. Allison, Cynthia L. Chandler, Nancy J. Copley, Stephen R. Gegg, David M. Glover, Danie B. Kinkade, Shannon M. Rauch, Adam Shepherd and Peter H. Wiebe Biological and Chemical Oceanography Data Management Office, Woods Hole Oceanographic Institution (WHOI), Woods Hole, MA USA

Abstract

Data are needed for analysis, synthesis, initial conditions and verifying i output, and advancing new theories on how the world works. The data r to be of high quality, collected and processed by established and comm accepted standards, discoverable, and accessible. There has been progr recently towards making data discoverable and accessible because of a bination of government requirements and higher expectations, but there still much to do. The Biological and Chemical Oceanography Data Mana ment Office (BCO-DMO) works in partnership with ocean science invest tors to serve data from research projects funded by the Biological and C ical Oceanography Sections and the Division of Polar Programs Antarcti Organisms & Ecosystems Program at the U.S. National Science Foundation. While BCO-DMO's data management model addresses many of the obstacles challenging the goal of data interoperability, there is still much to accomplish. This presentation targets where we are going and how we plan to get there.

Dialogue: April 2, 20xx

You: Show me the data that compares favorably with my MANTL version 7 model run from last night.

Computer: Can't because data aren't federated. Where should I look for the data? Anyone's data?

You: No. Only use data from trusted, authoritative sources.

Computer: But are temp1, air temperature and mytemp data comparable to your sea surface temperature?

You: Not sure. Check their ontologies.

Computer: Okay, but Smith's data comes from a new instrument with differences in acquisition and processing steps. Should I use these data along with all the others?

HOME

You: Why is this still so complicated?

BC - DM



Introduction to BCO-DMO



The Biological and Chemical Oceanography Data Management Office (BCO-DMO) staff members work with investigators to serve data online from research projects funded by the Biological and Chemical Oceanography Sections and the Division of Polar Programs Antarctic Organisms & Ecosystems Program at the U.S. National Science oundation

As of November 14, 2013 we have migrated our metadata database from ColdFusion to Drupal, a public domain content management system.

Figure 1: Current BCO-DMO home page providing text-based access to the metadata and data managed by BCO-DMO.

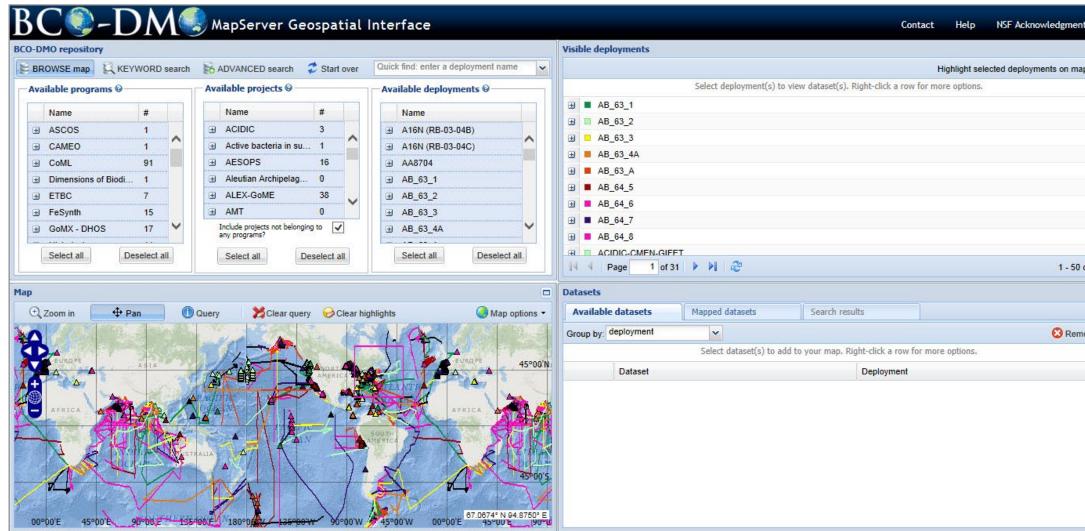


Figure 2: BCO-DMO's geospatial access is based on MapServer originally developed at the University of Minnesota.

BC - DM Progress in Data Discovery and Access: The Next Steps

model
need
nunity
jress
a com-
e is
age-
tiga-
Chem-
ic
tion

ABOUT US RECENT DATASETS

RESOURCES

02/13/2014 CTD Log

01/31/2014

01/29/2014

01/24/2014

tracer filters

without mussels

FEATURED RESOURCES

01/24/2014

CTD cast logs from Tioga cruises

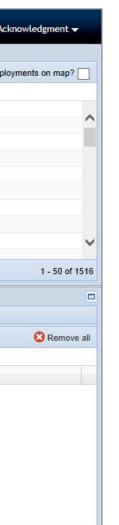
DATA

16S RRNA Gene And Particulate Monooxygenase 16S rRNA gene and particulate monooxygenase diversity

Palau Carbonate Chemistry Seawater carbonate chemistry from 9 sites in Palau, 2011-2012 Filter Tracer Content

Content of nitrogen and carbon in Fidepool Ammonium And

Ammonium in tidepools with and



Support investigators

- More support in preparation of Data Management Plan • Continue one to one interactions for the entire data life cycle, from proposal to collection to analysis to publication and data use and reuse (Higgins, 2012; Michener et al., 2011; Strasser et al., 2011)
- Uploading data (Figure 3)
- More data quality checking –range and bounds checking
- Help with dark data long tail of data
- More widespread use of Digital Object Identifier (DOI) Data citation credit

Improve data searching

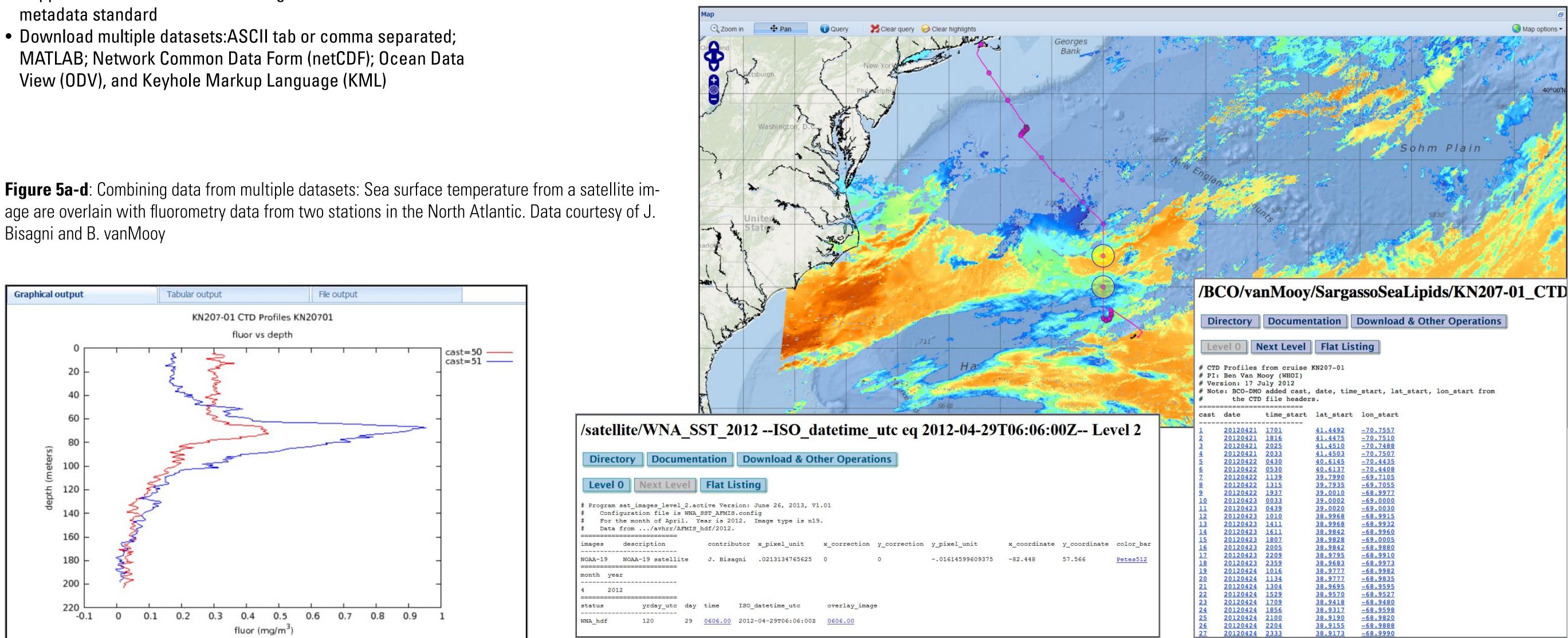
- Use of ontologies to help find data using common or similar terminology
- Linked Data start of federated data
- Interoperability of distributed systems combining Linked Data and Semantic Web technology
- Geospatial searching supported by faceted search features (Figure 4)

Assess fitness for purpose

• New instruments necessitate new ways to look at data • New data visualizations will be needed (Figure 5)

- **Access data and metadata** • Add Open-source Project for a Network Data Access
- Protocol (OPeNDAP) • Support International Standards Organization's ISO19115-2
- metadata standard • Download multiple datasets:ASCII tab or comma separated;
- View (ODV), and Keyhole Markup Language (KML)

Bisagni and B. vanMooy



References

Higgins, S. (2012). The lifecycle of data management. In Pryor, G. (ed.), Managing research data (pp. 17-46). London: Facet Publishing.

Michener, W. K., Porter, J., Servilla, M., & Vanderbilt, K. (2011). Long term ecological research and information management. Ecological Informatics, 6(1), 13-24. doi: http://dx.doi. org/10.1016/j.ecoinf.2010.11.005

Strasser, C., Cook, R., Michener, W., Budden, A., & Koskela, R. (2011). DataONE promoting data stewardship through best practices. In Jones, M.B. & Gries, C. (eds.), Proceedings of the Envi-

Figure 3a-b: Originally contributed spreadsheets can take on an infinite number of organizations and column naming conventions (figure a). In this case, the data columns are labeled clearly, but the data should be re-organized (figure b) before serving.

Figure 4: The beta version of the BCO-DMO faceted search option allows selection criteria to be chosen in any order.

	A	B	C	D	E	F	G	H
1	Multinet (100 µm mesh size)							
2	Stat.	Date	Time	Position Latitude	- Longitude	Water depth	Sampling depth	filtr. Vol.
3			(UTC)	<u>s</u>	W	<u>(m)</u>	<u>(m)</u>	(m ³)
4	16	12/1/2004	8:40	68°10,407'	54°52,735'	1897	1000-500	126
5							500-200	70
6							200-100	22
7							100-50	10
8							<u>50-0</u>	12
9	31	12/6/2004	13:55	68°05,977'	55°16,025'	1524	1000-500	125
10							500-200	75
11							200-100	21
12							100-50	12.5
13							<u>50-0</u>	12.5
14	46	12/9/2004	7:45	68°02,785'	55°12,653'	1592	1000-500	125
15							500-200	75
16							200-100	25
17							100-50	12.5
18							50-0	12.5
19	58	12/13/2004	7:50	67°53,869'	55°24,418'	1366	1000-500	126
20							500-200	72
21							200-100	24
22							100-50	12
23							50-0	12

A	B	C	D	E	F	G	H	I	J	K	L	M
1 Zooplankton Station Data												
2 Sigi Schiel, Principle Investigato	r											
3 Multinet (100 μm mesh size)									optional	optional	optional	
4 cruise_id	station	date_utc	time_utc	lat(S)	lon(W)	r -	net	e	n –	×	d	d
5 (notice no spaces or special char	acters)	local)	local)	decimal	decimal	meters		meters	meters	meters	meters	m³
6 ANTXII_2	16	12/1/2004	8:40	6810.407	5452.735	1897	1	1000-500	500	1000	750	126
7 ANTXII_2	16	12/1/2004	8:40	6810.407	5452.735	1897	2	500-200	200	500	350	70
8 ANTXII_2	16	12/1/2004	8:40	6810.407	5452.735	1897	3	200-100	100	200	150	22
9 ANTXII_2	16	12/1/2004	8:40	6810.407	5452.735	1897	4	100-50	50	100	75	10
0 ANTXII_2	16	12/1/2004	8:40	6810.407	5452.735	1897	5	50-0	0	50	25	12
1 ANTXIL2	31	12/6/2004	13:55	6805.977	5516.025	1524	1	1000-500	500	1000	750	125
2 ANTXII_2	31	12/6/2004	13:55	6805.977	5516.025	1524	2	500-200	200	500	350	75
13 ANTXII_2	31	12/6/2004	13:55	6805.977	5516.025	1524	3	200-100	100	200	150	21
4 ANTXII_2	31	12/6/2004	13:55	6805.977	5516.025	1524	4	100-50	50	100	75	12.5
15 ANTXII_2	31	12/6/2004	13:55	6805.977	5516.025	1524	5	50-0	0	50	25	12.5
6 ANTXII_2	46	12/9/2004	7:45	6802.785	5512.653	1592	1	1000-500	500	1000	750	125
7 ANTXII_2	46	12/9/2004	7:45	6802.785	5512.653	1592	2	500-200	200	500	350	75
8 ANTXII_2	46	12/9/2004	7:45	6802.785	5512.653	1592	3	200-100	100	200	150	25
9 ANTXII_2	46	12/9/2004	7:45	6802.785	5512.653	1592	4	100-50	50	100	75	12.5
0 ANTXII_2	46	12/9/2004	7:45	6802.785	5512.653	1592	5	50-0	0	50	25	12.5
21 ANTXIL2	58	12/13/2004	7:50	6753.869	5524.418	1366	1	1000-500	500	1000	750	126
2 ANTXII_2	58	12/13/2004	7:50	6753.869	5524.418	1366	2	500-200	200	500	350	72
23 ANTXII_2	58	12/13/2004	7:50	6753.869	5524.418	1366	3	200-100	100	200	150	24
24 ANTXII_2	58	12/13/2004	7:50	6753.869	5524.418	1366	4	100-50	50	100	75	12
25 ANTXII_2	58	12/13/2004	7:50	6753.869	5524.418	1366	5	50-0	0	50	25	12
26 ANTXII_2	77	12/17/2004	8:00	6746.24	5519.18	1473	1	1000-500	500	1000	750	126
27 ANTXIL2	77	12/17/2004	8:00	6746.24	5519.18	1473	2	500-200	200	500	350	72
28 ANTXII_2	77	12/17/2004	8:00	6746.24	5519.18	1473	3	200-100	100	200	150	23
9 ANTXII_2	77	12/17/2004	8:00	6746.24	5519.18	1473	4	100-50	50	100	75	9
0 ANTXII_2	77	12/17/2004	8:00	6746.24	5519.18	1473	5	50-0	0	50	25	11
ANTXII 2	97	12/21/2004	7:40	6749.98	5530.11	1259	1	1000-500	500	1000	750	140

arch			Visible deployments	
Return to classic map 🧣	Start over			Highlight selected deployments on map?
Categories	Category : Instruments	You are looking at	Select deployment(s) to view dat	aset(s). Right-click a row for more options.
Select a category.	Select or search for a value. 💉 💈 Creset	View or modify your current search.	🕀 🔲 KN199-04 (GT10)	
Instruments	CTD TRIAXUS	Instruments 🥖 😒		
Programs	CTD-FRRfluorometer	Dissolved Oxygen		
Projects	Deep Submersible Incubation Device	Sensor		
Deployments	Dissolved Oxygen Sensor			
Platforms	Drifter Buoy			
Instruments by	Druck PDCR 4020 pressure sensor			
type	Echo sounder - single-beam			
Parameters V	Electronic Jig Machine			
Parameters by			🛛 🕅 🖣 Page 🚺 of 1 🕨 🕅 😂	1-
				1-
		· · ·	Datasets	
-	an 🕕 Query 🄀 Clear query 😔 Clear	· · ·	Datasets Available datasets Mapped datasets	
€ Zoom in 🛛 🕂 P	an Duery Sclear query Sclear	· · ·	Datasets	Remo
Ð Zoom in 🛛 🕂 P	A Carton March	r highlights O Map options	Datasets Available datasets Group by: deployment Select dataset(s) to add to your	
€ Zoom in	A Carton March	· · ·	Datasets Available datasets Group by: deployment Select dataset(s) to add to your	Remo
€ Zoom in		r highlights O Map options	Datasets Available datasets Group by: deployment Select dataset(s) to add to your	Remo r map. Right-click a row for more options.
€ Zoom in		r highlights O Map options	Datasets Available datasets Group by: deployment Select dataset(s) to add to your	Remo r map. Right-click a row for more options.
€ Zoom in		r highlights O Map options	Datasets Available datasets Group by: deployment Select dataset(s) to add to your	Remo r map. Right-click a row for more options.
€ Zoom in	North American Basin	r highlights O Map options	Datasets Available datasets Group by: deployment Select dataset(s) to add to your	Remo r map. Right-click a row for more options.
P	North American Basin Sargasso	r highlights Map options	Datasets Available datasets Group by: deployment Select dataset(s) to add to your Dataset	Remo r map. Right-click a row for more options.
€ Zoom in	North American Basin Sargasso	r highlights Map options	Datasets Available datasets Group by: deployment Select dataset(s) to add to your Dataset	Remo r map. Right-click a row for more options.

ronmental Information Management Conference 2011 (EIM 2011) (pp. 126-131). University of California. Retrieved from https://eim.ecoinformatics.org/eim2011/eim-proceedings-2011/view Acknowledgments

BCO-DMO is funded by the National Science Foundation. We acknowledge the work done by all of the investigators who contribute their data to BCO-DMO. The user interfaces to the BCO-DMO data system were developed in collaboration with Julie Allen and Katherine Joyce (WHOI). The geospatial interface to the BCO-DMO data system was developed in collaboration with Charlton Galvarino (Second Creek Consulting, LLC).

Poster: 1707 Session: 071 Ocean Sciences Meeting February 24-28, 2014





