# Evolving Requirements for Data and Information Management

David McKinnie Office of Ocean Exploration and Research

Sharon Mesick National Centers for Environmental Information

Megan Cromwell Northern Gulf Institute/Mississippi State University National Centers for Environmental Information



Ocean Exploration and Research

# **OER** Mission and the Data Imperative

- OER's mission is to explore the deep ocean:
  - map to modern standards
  - conduct baseline characterization
  - make that data and information available and accessible to whomever needs it in a form that is useful for science, for decisions, for general information about the ocean
- If ocean exploration results are not available, accessible, and reusable, we have not succeeded in our mission



# **Data Management:** *Evolution in a changing landscape*

2000: OE / NOAA Data **Center Agreement** 

2001: Data Quality Act

2002: OE Data IPT Seminal Documents

2007: OE / NURP Merger

2009: Ocean Exploration Act

2015: NCEI: NOAA Data **Center Consolidation** 

2013: Open Data Access Policy

> 2015: NOAA Public **Access Policy (PARR)**

2019: Open Government Data Act



Expedition Portfolio by Year: Partnerships / NOAA Ship Okeanos Explorer / Competitive Grants

# **Evolving View of Characterization**

- 2008-2015
  - Sonar (multibeam/mapping operations)
  - Video (ROV operations)
  - Some oceanographic data
- 2015-2018
  - Water Column?
  - Acoustics?
  - Other measurements?
- 2019
  - Processes
  - Connectivity
  - Rapid evolution

EVOLVING BASELINE
INITIAL CHARACTERIZATION
SONAR
VIDEO OPTICAL
ENUIRONMENTAL CONDITIONS



on obtaining spect

the prince like the set of a structure of

200

1. 1

O

0

# OER's conceptual model for an expedition

- We explore a 3-D polygon
- We need to think in terms of:
  - How all data collected by the expedition in that polygon might relate
  - Developing new CONOPS using gliders, other platforms and sensors (Deep Argo? Saildrone? Silixia?)
  - Integrating data collected in that polygon in an extended time domain (e.g. moored arrays, Argo floats).



# **OER's Data Access and Use Requirement**

- NOAA Ship Okeanos Explorer retires in 2023-24. There is no dedicated replacement; OER will begin operate from multiple platforms soon
- By 2025, OER must be heavily invested in autonomy
- The new CONOPS that will result mean much more complex and diverse data work flows, more diverse data types, and radically increased data volume









SCIENTIFIC PAPERS

# DATA AVAILABILITY: 2+ YEARS

# **THE PRESENT**

REAL-TIME SCIENCE ON SHORE

S NOAA

#### OPEN ACCESS DATA

----

# DATA AVAILABILITY: 6 WEEKS



National Centers for Environmental Information: Making Ocean Exploration Data Available, Accessible, and Useful

# **Outward-facing Data Management**



OceanExplorer.NOAA.gov





# **Community of Practice**



# Learn: Creating and Archiving Born-Digital Video

# Federal Agency Digitization Guidelines Initiative Audio-Visual Working Group

- Library of Congress
- National Archives and Records Administration
- Smithsonian Institution

#### Fiscal '15 Report

- Federal Case Studies (OER one of eight case studies)
- Recommendations
- Resources

http://www.digitizationguidelines.gov/guidelines/video\_bornDigital.html







## Learn: National Science Foundation

#### 2016 Video Workshop: Establishing Community Standards for Underwater Video Acquisition, Tagging, Archiving, Access

- Recording
- File Naming Conventions
- Video Formats
- Compression
- Metadata Priorities
- Timecode embedding
- Audio Channels
- Annotation / Event Logs
- Archive / Open Access





#### *Innovate*: Video Data Management



### *Innovate*: Video Data Management



OceanExplorer.NOAA.gov



Ocean Exploration and Research

#### Innovate: Video Data Management



ideo Details	Dive Summary	Feedback

## Innovate: Sharing Video Management Approaches

- Lessons learned
  - Video guidance shared across NOAA
  - Best practices and templates
- Legacy data integration
  - Career collections and key expeditions available online
  - Recovery from disparate media
  - Data now available in standard formats
- Benefits for Future
  - Legacy data recovery has implications for machine learning
  - Reinforces need for data architecture in advance





#### Develop: End-to-end Sample Data Management



OceanExplorer.NOAA.gov



Ocean Exploration and Research

### Develop: End-to-end Sample Data Management

#### Specimens Collected on Okeanos Explorer



Calendar Year



#### *Develop*: End-to-end Sample Data Management



#### *Develop*: Sample Data Access

		Victorgorgia sp. : Anthot	nelidae : Gorgonacea : Anthozoa : Cnidaria 🛛 🗙
Smithsonian		Specimen/Lot	
National Museum of Natural History   Visit Exhibits   NMNH Home NMNH Research & Collections   Invert   Search the Department of Invert	vents   About   Join Us tebrate Zoology   Collections Vertebrate Zoolo	Catalog Number: Scientific Name: Identified By: Date Identified: Classification: Common Name: Collection Name: Specimen Count: Preparation: Date Collected: Ocean:	USNM 1424234 Victorgorgia sp. Kelley, Christopher, University of Hawaii at Manoa (UNITED STATES) 2016 Animalia, Cnidaria, Anthozoa, Octocorallia, Gorgonacea, Scleraxonia, Anthothelidae Octocorals NOAA Collections 1 Ethanol - 95% 16 Aug 2016 (UTC_DateTime 20160816T004623) North Pacific Ocean
Keyword Search Results - Grid View		Country: Precise Locality:	United States Southwestern most seamount from Wake
Catalog# Kind of Object Scientific Name	Family Phylum	Centroid Latitude:	16.5591
☐ 1424234 Specimen/Lot Victorgorgia sp.	Anthothelidae Cnidaria	Centroid Longitude:	165.345 Deepwater Wonders of Wake
H 1453702 Specimen/Lot Victorgorgia sp.	Anthothelidae Cnidaria	Vessel:	Deep Discoverer ROV; Okeanos Explorer R/V
	Euplectellidae Porifera	Cruise:	EX1606
	Euplectellidae Porifera	Depth (m): Notes:	1215.48 Genomic DNA from this specimen has been deposited in the Ocean
	Euplectellidae Porifera		Genome Legacy (OGL) collection at Northeastern University. See "Other Numbers" for OGL Extract ID(s)
∃ 1453775 Specimen/Lot Walteria sp.	Euplectellidae Porifera	Other Numbers (Type	Value):
	Euplectellidae Porifera	Specimen Number	EX1606_20160816T004623_D2_DIVE14_SPEC03BIO
	Aphanipathidae Cnidaria	Extract ID (DNA)	E26091
H 1292604 Specimen/Lot cf. Bathycrinus sp.	Bathycrinidae Echinodermata	Primary Collector #:	EX1606 20160816T004623 D2 DIVE14 SPEC03BIO
H 🔲 1424212 Specimen/Lot cf. Jasonisis sp.	Isididae Cnidaria	Accession Number:	2078414
A Page 158 of 159 V A Clear Selections E	Export as KML Export Selected Res	EZID:	http://n2t.net/ark:/65665/339470450-4010-4941-9b24-c676ff43b531
Invertebrate Zoology Collections Keyword Search Search by	Field Help Feedback		



#### Develop: Sample Data Access



NMNH Home | NMNH Research & Collections | Invertebrate Zoology | Collect

#### Search the Department of Invertebrate Zoc

Ke	ywor	d Search R	esults - Grid Vi	ew		
		Catalog#	Kind of Object	Scientific Name	Family	Phylum
		1424234	Specimen/Lot	Victorgorgia sp.	Anthothelidae	Cnidaria
Ð		1453702	Specimen/Lot	Victorgorgia sp.	Anthothelidae	Cnidaria
Ð		1453729	Specimen/Lot	Walteria sp.	Euplectellidae	Porifera
±		1453785	Specimen/Lot	Walteria sp.	Euplectellidae	Porifera
Ð		1453738	Specimen/Lot	Walteria sp.	Euplectellidae	Porifera
Ð		1453775	Specimen/Lot	Walteria sp.	Euplectellidae	Porifera
±		1465275	Specimen/Lot	Walteria sp.	Euplectellidae	Porifera
Ð		1507320	Specimen/Lot	cf. Aphanipathes sp.	Aphanipathidae	Cnidaria
±		1292604	Specimen/Lot	cf. Bathycrinus sp.	Bathycrinidae	Echinoderma
Ð		1424212	Specimen/Lot	cf. Jasonisis sp.	Isididae	Cnidaria
M	4	Page 1	58 of 159 🕨 🕨	🕅   🍣   Clear Selections   E	Export as KML   Ex	oport Selected
In	verte	brate Zoolog	y Collections	Keyword Search Search by	Field Help	Feedback

Museum samples are only as valuable as the data associated with them.

The *Okeanos* collected material has extremely high scientific value because of the finely detailed electronic data associated with it.

SODA ensures that all this data remains associated with the proper samples, and streamlines importing that information to the Smithsonian Natural History Museum catalog with minimal (if any) errors, where that data is publicly available to anyone interested.

As a result, the *Okeanos* samples remain high quality for scientific studies performed now and well into the future.

Dr. Abigail Reft, Smithsonian National Museum of Natural History

OceanExplorer.NOAA.gov

#### Develop: Sample Data Analysis



https://oceanexplorer.noaa.gov/okeanos/animal\_guide/animal\_guide.html

### **Benthic Deepwater Animal Identification Guide**



# The Value of Collaboration

OER is a key partner for the DSCRTP. As the major NOAA funder of deep-sea exploration and research, OER's expertise, research, and information products are central to the DSCRTP mission. As a management-oriented research program, the DSCRTP complements OER's work and directly links it to resource managers' needs.

Together, our programs have been able to leverage additional ship time, make connections between exploration and management needs, and achieve more together than either program could alone.

Thomas Hourigan, Ph.D.

Chief Scientist, Deep Sea Coral Research and Technology Program





#### https://noaa.maps.arcgis.com/home/index.html



OceanExplorer.NOAA.gov

#### New geospatial web services



Shaded relief imagery, 3D visualization



Near real-time updates, automatically updated when ship is at sea

#### Web Services: Okeanos Explorer Bathymetric Grids





OceanExplorer.NOAA.gov







Ocean Exploration and Research

# The Future of Data: Direct-to-Cloud-to-Users with Telepresence X.0 and Data X.0

### Future: Telepresence X.0 and Data X.0





### Science Benefits: Telepresence X.0 and Data X.0







Ocean Exploration and Research

#### Interactive Science Tools from Web Services



VIRTUAL WHITEBOARD



#### Combining output from multiple web services to:

- drive operations;
- verify data quality; and,
- develop intelligent data in near-real time

	A 8 + 0 1	
		party production press, mar.
A		
	B managed	
	Contraction (M) -	O weater
Store of the second		0
<b>Stat</b>	9 mil -	0



#### ENVIRONMENTAL VARIABLES



etay Erines My Monhatore		the second se						_	
Jun Des (UTC) End Des (UTC)		Corner	Latuda	Longhues	Opt	.99	Actor	1	
Nov-2010 19:09:45 17:Nov-2018 19:09	to sergeptic		12 \$425	47 0000 1	700 4	Testorite	43		
Non-2012 10-14 20 17-14-0028 10-14	and an and a second sec		17.0425	37.55405	100.0	Sectorie.	1.		
New Gotte tel to 24 th few 2018 reito	of sergence		17.04222	07.86800	700.0	Serburne	55		
Nov-2018 19:18:47 17-Nov-2518 19:16	Langement & services 17		17.54242	-87.00510	790.2	SaxScritte	5		
Nov-2016 (9:10.42 17-Nov-2016 (9:16	C physical adirection		17.94242	-01 602/1	720.5	Gastrice	45		
Ami-2018 /9 12:42 17 Nov-2018 19:19	i aly		(7.942)7	-87.88812	700.4	Gasleria	4		ANNOTATION
New-2018 19:20 12 17/Nov-2018 19:20	2 ayranome		17.04240	477.56873	100.0	EMILIO DA	8		/
Nov-2018 19/22 07 17-Nov-2018 19/221	51 pyrtebra		11 (#28	1-87.55514	100.3	GAIGEBA	5		
Nov-2018 10:25:22 17-Nov-2018 19:25	22 colphynema		17 64245	-67.00000	750.4	Sectors	31		LUG
1000 1100 00171 11 10 01 100 001	() symmetry		17.04245	-67.38602	200.4	Tas[res	4		
nev-data télahék intervidaté telati	H M		17.04250	101,99924	792.1	inticia	5		
HO-2018 16/26/10 17-NOV2018 19/29	to aprende the		17.84253	-27 85499	100.4	14130151	5		
Noi-2018 (9:29-18 17-Nov-2018 19:29	15 radiolatian		17.94253	107 86459	705.2	designede	5		
Amil 2016 10/2018 17. Nev 2018 10/20	C mysel shrink		17 \$1244	417 5387.0	702 A	Sastria	6		
April 18 18 31 18 17 Nov 3018 19 31	S sayaayaanaa aanaayaana		17.04348	107.83817	702.8	Sastuta	D.		
No-3018 1831 17 17 No+2018 1831	17. stutoy pyresome		11 84203	47 38503	702.5	345018	24		
Non-2018 19:35-13 17-Non-2018 19:35	5 calyoophoran elemenophore		17.04246	-07.56612	700.1	Serbria	3		
ADD-3018 19:54:32 17-1404-3218 19:54	12		17.64222	-87 84500	2(0.)	featone	32		

#### Science Benefits: the Virtual Whiteboard



#### Envisioning the Virtual Whiteboard for collaborative scientific analysis

- Support annotations and future data accessibility and usability
- Contributes to scientific understanding

### **Operational Guidance: 3D Chemistry Model**



OceanExplorer.NOAA.gov



### **Integrated GIS Analysis Tools**





Ocean Exploration and Research

OceanExplorer.NOAA.gov

### **Data Management Evolution**

#### **EXPERIENCE:**

Π

2010: Access Best practices implemented and refined aboard NOAA Ship Okeanos Explorer

#### **UNDERSTAND:**

2019: Reuse

**Ocean Exploration** data readily accessible and usable

#### **EMPOWER USERS:**

2024: Service Economy Telepresence X.0 and Data X.0 will empower users to build their own experience

**EXPLORE:** 

A

2002: Archive Define and document best data management practices



Ocean Exploration and Research



National Centers for Environmental Information