



NOAA EDM

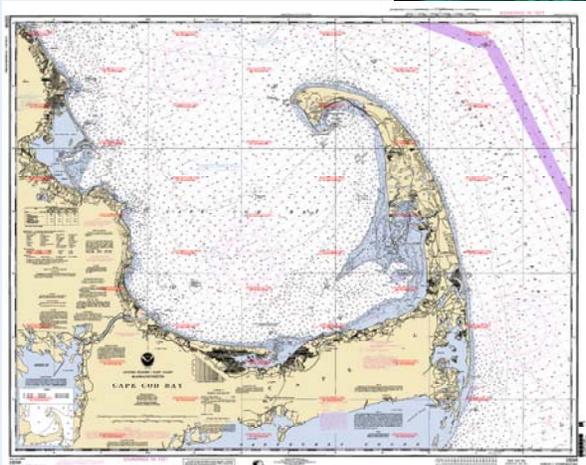
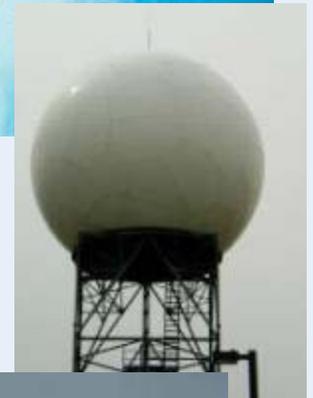
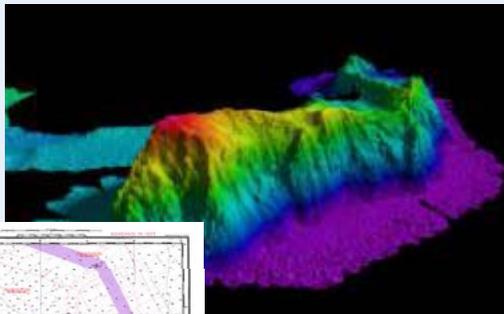
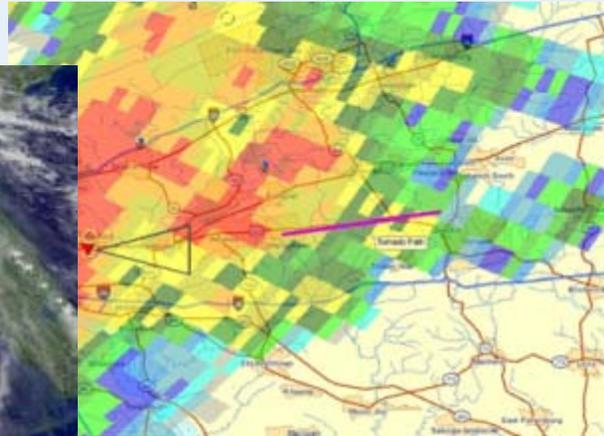
NOAA Environmental Data Management Conference

Joseph Klimavicz

NOAA Chief Information Officer & Director, HPCC

Implementing NOAA's Data Management Vision

NOAA's Data is a Global Resource



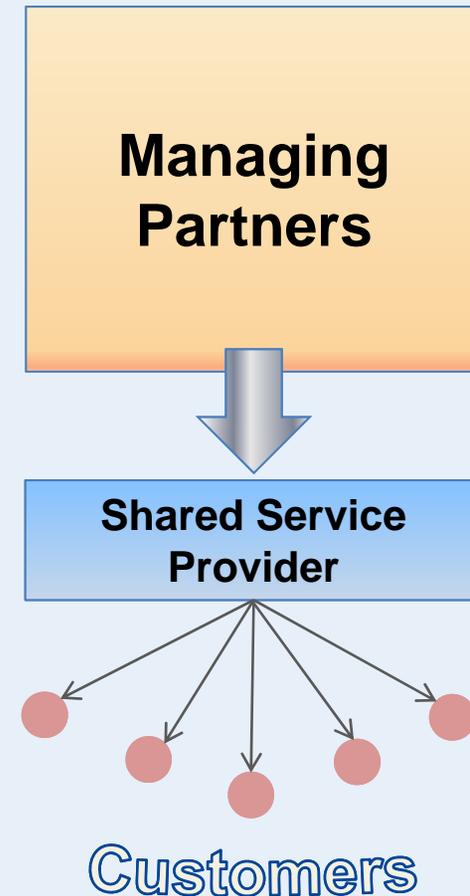
Role of NOAA OCIO

- Acquisition, management, and use of NOAA's information and information technology resources
- Policy, standards, and architecture
- Performance management
- Oversight of EDMC and GIS Committees



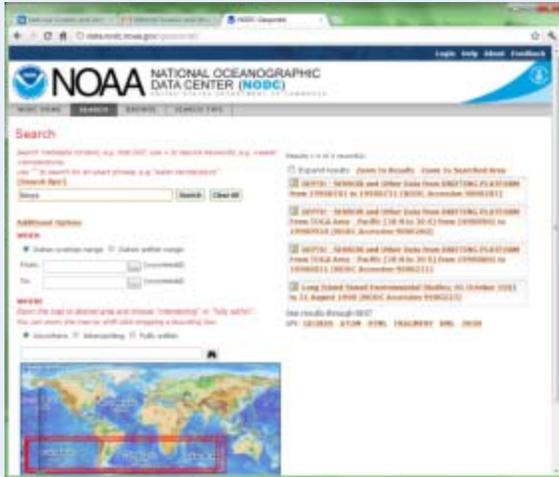
IT Consolidation / Shared Services

- Federal CIO Shared IT Services Strategy
- DOC IT Portfolio Management Policy
- NOAA opportunities
 - Cloud First Policy
 - Shared geospatial hosting (open source and Esri)
 - Enterprise Geoportal Server
 - Geospatial Enterprise License Agreements



Enhancing NOAA Data Access

Geoportal Server



NOAA GeoPlatform

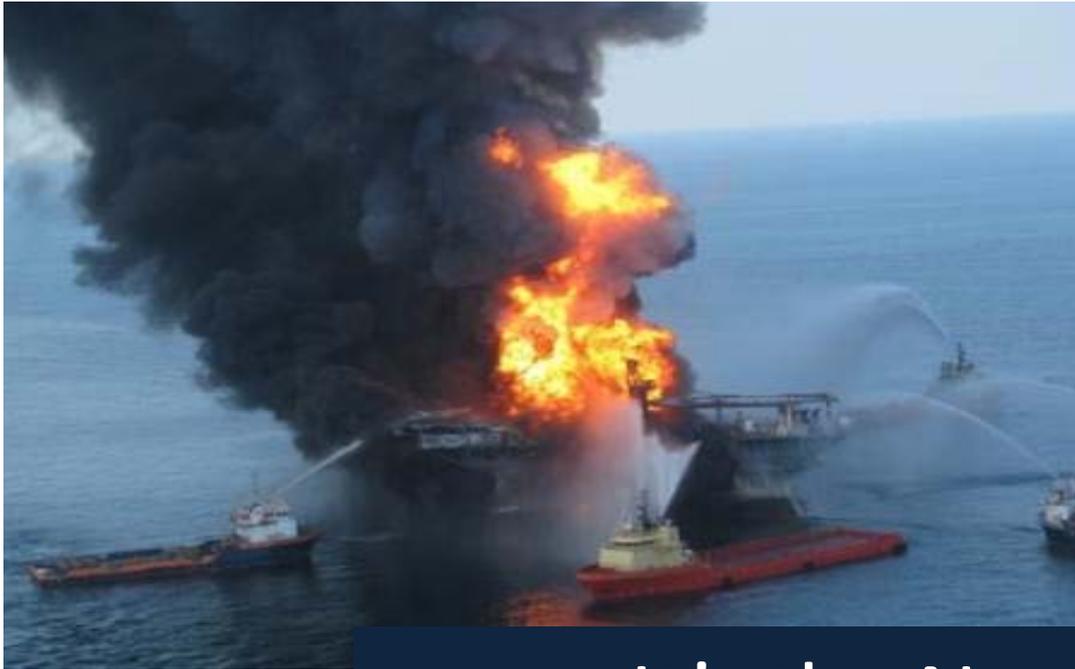


DWH Information Repository



Unified Access Framework (UAF)





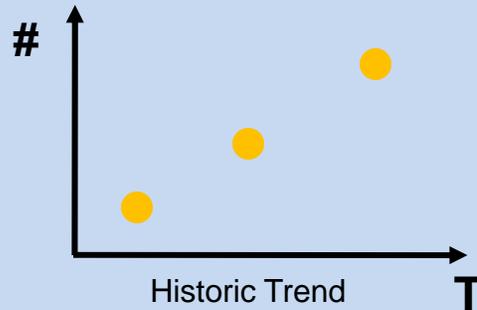
It's the Next Disaster.
Do we know what data we have?



Monitoring Performance DM

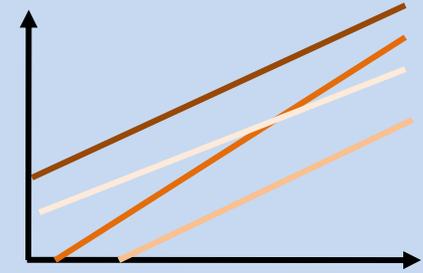
Quantity

of Catalogs



of Records

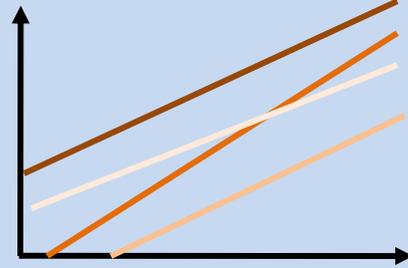
NOAA	25000
Catalog 1	10000
Catalog 2	5000
Catalog 3	10000



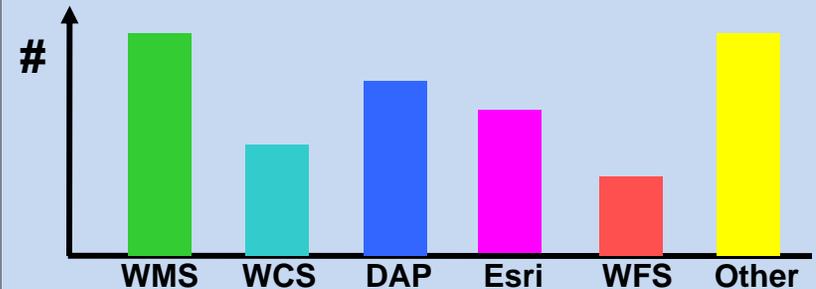
Accessibility

% Records with Online Service Link

NOAA	46%
Catalog 1	25%
Catalog 2	80%
Catalog 3	50%

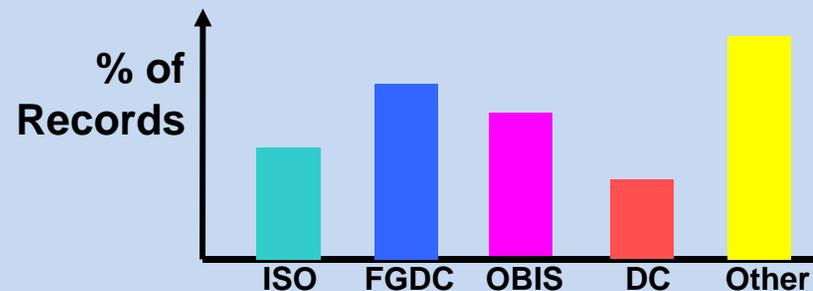


Service Types Offered



Documentation

Metadata Dialects Used



Metadata Completeness Scores

	Mean	σ	Min	Max
NOAA				
Catalog 1				
Catalog 2				



NOAA EDM

NOAA Environmental Data Management Conference

Logistics

Conference Overview

DAY 1 + Tuesday, May 15	
8:30	Registration Open <i>Gallery outside Chesapeake Ballroom</i>
9:00	Welcome Plenary <i>Chair: Jeff de La Beaujardière</i>
9:45	Joe Klimavicz <i>NOAA Chief Information Officer</i>
10:15	Announcements / Breakout Logistics
10:30	Break
10:45	Keynote: Mike Daconta, InCadence <i>Information Management and Semantics in the Cloud</i>
12:00	Lunch (on your own)
1:15	Day 1 Breakout Data Centers, Future of CLASS <i>Chair: Nancy Ritchey Chesapeake Ballroom</i> Data Citation <i>Chair: Dan Kowal Room 2103 (Second Floor)</i>
2:45	Break
3:00	Keynote: Dave Jones, StormCenter <i>Real-Time Cross-Agency Collaboration and Data Sharing for Decision Support - Serving Society with NOAA Data</i>
4:00	Lightning Talks <i>Chair: Jeff de La Beaujardière</i>
5:00	Networking Session <i>Room 2103 (Second Floor)</i>
7:00	End of Day 1

DAY 2 + Wednesday, May 16	
8:30	Registration Open <i>Gallery outside Chesapeake Ballroom</i>
9:00	Day 2 Morning Breakout Data Access and Discovery, Part 1 <i>Chair: Ted Habermann Room 1105 (Concourse Level)</i> Data Preservation <i>Chair: Nancy Ritchey Room 0109 (Lower Level)</i> Data Integration Projects <i>Chair: Carmel Ortiz Room 0105 (Lower Level)</i> Ocean Biological & Chemical Data <i>Chairs: Herman Garcia, Krisa Arzayus Chasen Family Room (Concourse Level)</i>
Break	
12:00	Lunch (on your own)
1:30	Day 2 Afternoon Breakout Data Access and Discovery, Part 2 <i>Chair: Derrick Snowden Room 1105 (Concourse Level)</i> Cloud Computing <i>Chair: Scott Hausman Room 0105 (Lower Level)</i> Data Management Planning <i>Chair: Ruth Duerr Chasen Family Room (Concourse Level)</i> Data Documentation <i>Chair: Ted Habermann Room 0109 (Lower Level)</i>
Break	
4:30	End of Day 2

DAY 3 + Thursday, May 17	
8:30	Registration Open <i>Gallery outside Chesapeake Ballroom</i>
9:00	Intro to Geoportals & GeoPlatforms <i>Chairs: Kim Jenkins, Tony Lavoie</i>
9:25	Discuss Data Portal & Catalog Relationships
9:45	Day 3 Breakout Geoportals Servers, Catalogs NOAA GeoPlatform/Technical <i>Room 1105 (Concourse Level)</i>
10:30	Break
10:45	Keynote: Peter Fox, RPI <i>Interagency GeoData 2012 opportunities: Broadening the Community and the Discussion</i>
11:30	Closing Plenary <i>Chair: Jeff de La Beaujardière Awards, wrap-up, and next steps</i>
12:30	End of Conference

Where to go...

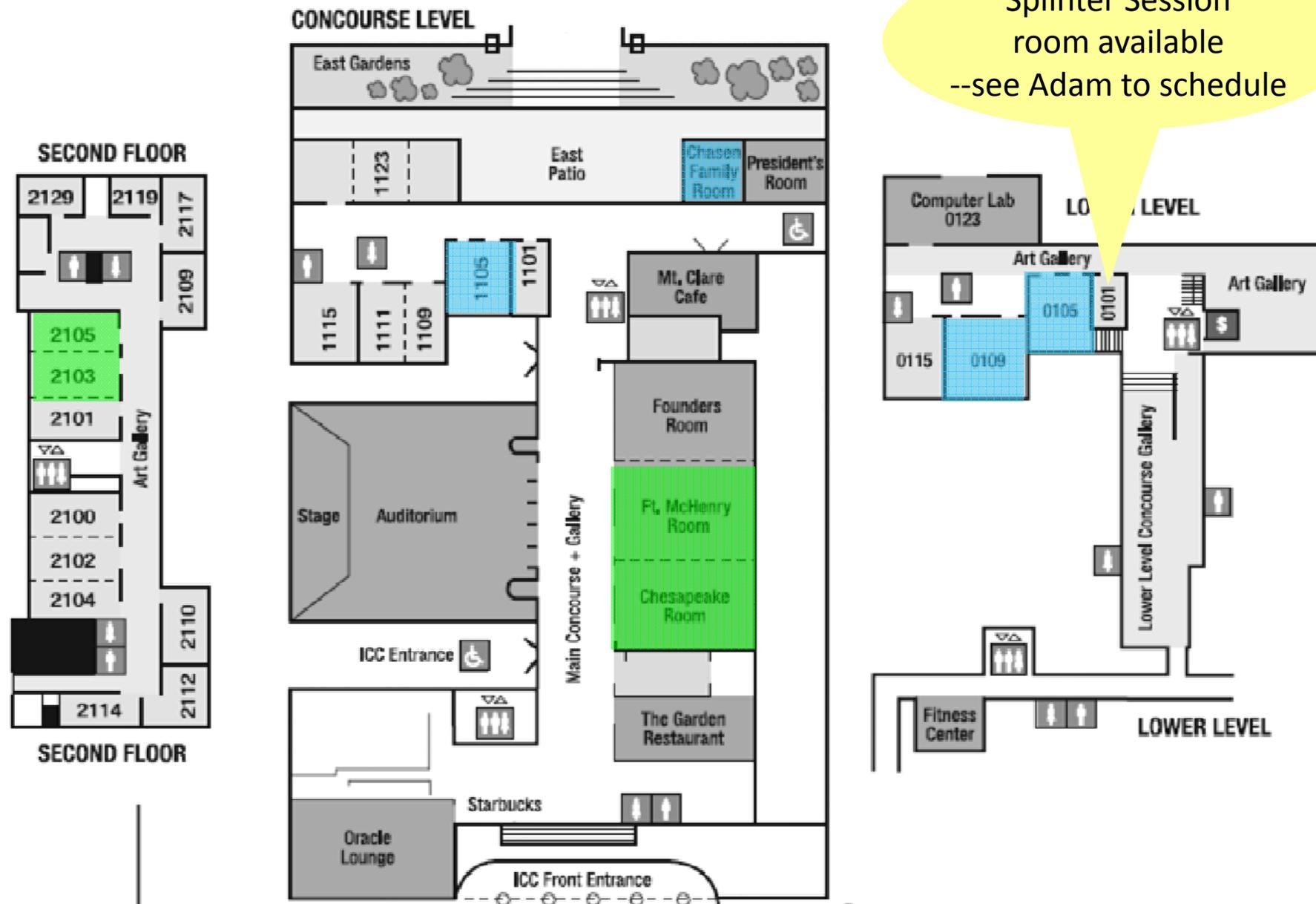
 All plenary sessions are located in the Chesapeake Ballroom on the Concourse Level

 Refer to schedule for locations of breakouts and open sessions

Conference Overview

Day 1 - Tuesday	Day 2 - Wednesday	Day 3 - Thursday
Opening Plenary	Breakouts (4)	Portals & Catalogs
Keynote: Mike Daconta, <i>InCadence</i>		Keynote: Peter Fox, <i>RPI</i>
<i>Lunch</i>	<i>Lunch</i>	Wrap-up/closing
Breakouts (2)	Breakouts (4)	
Keynote: Dave Jones, <i>StormCenter</i>		
Lightning Talks		
Evening Networking Session		

Plenary & Breakout Rooms



Online conference materials:

<http://www.nosc.noaa.gov/edmcon>

The screenshot displays the NOAA 2012 EDM Conference website. At the top, there is a header with the NOAA logo on the left, the text "NOAA 2012 EDM Conference" in the center, and "Your Conference Materials Online" below it. On the right is a "go green" logo. Below the header is a section titled "Conference Overview Materials" containing a grid of links: "Conference Overview", "Schedule At-a-Glance", "Collaborative Notepads", "Twitter Hashtag #noaaedmc", "Marriott MAP & Where to Go List", and "Remote Attendance". Below this is a row of three colored boxes representing the conference days: "Day 1 Tuesday, May 15" (green), "Day 2 Wednesday, May 16" (blue), and "Day 3 Thursday, May 17" (purple). Underneath is a "Speaker Bios" section with a grid of speaker names: Michael C Daconta, Dave Jones, Peter Fox, Joe Klimavicz, and Jeff de La Beaujardière. The final section is "Other conference materials" with a grid of links: "List of Lightning Talks", "Register of Attendees", "Comments and Feedback", "Rate the Lightning Talks", "Print All Materials", and "Contact Us". Several links are circled in red: "Collaborative Notepads", "Twitter Hashtag #noaaedmc", the three day boxes, "Rate the Lightning Talks", and "Comments and Feedback". A yellow speech bubble points to "Collaborative Notepads" with the text "password: edmc".

NOAA 2012 EDM Conference
Your Conference Materials Online

Conference Overview Materials

Conference Overview	Schedule At-a-Glance	Collaborative Notepads
Twitter Hashtag #noaaedmc	Marriott MAP & Where to Go List	Remote Attendance
Day 1 Tuesday, May 15	Day 2 Wednesday, May 16	Day 3 Thursday, May 17

Speaker Bios

Michael C Daconta	Dave Jones	Peter Fox
Joe Klimavicz	Jeff de La Beaujardière	

Other conference materials

List of Lightning Talks	Register of Attendees	Comments and Feedback
Rate the Lightning Talks	Print All Materials	Contact Us

password:
edmc

Coffee Break
15 minutes



Keynote: Mike Daconta

10:45

Information Management in the Cloud

NOAA 2012 Environmental Data Management Conference

May 15 to 17 + Marriott Inn and Conference Center at UMD College Park

Keynote Speaker (Tuesday morning)

Information Management and Semantics in the Cloud



Michael C. Daconta

Vice President
Advanced Technology for InCadence Strategic Solutions

Mr. Daconta is the Vice President of Advanced Technology for InCadence Strategic Solutions (<http://www.incadencecorp.com>) where he currently guides multiple advanced technology projects for government and commercial customers. Mr. Daconta is a well-known author, lecturer and columnist having authored or co-authored 11 technical books, numerous magazine articles and online columns. He also writes the monthly "Reality Check" column for Government Computer News. Previously, Mr.

Daconta was the Metadata Program Manager for the Department of Homeland Security where he spearheaded data standardization, stewardship, and metadata registration. He was selected by the Office of Management and Budget and the Federal CIO Council to lead the Federal Enterprise Architecture (FEA) Data Reference Model (DRM) Working group which successfully delivered DRM V2.0 in December 2005.

In conjunction with the Department of Justice he launched the National Information Exchange Model (NIEM) to provide a reusable set of core XML components for building exchange packages. For his work at DHS, Mr. Daconta was selected to the prestigious "Fed 100" by Federal Computer Week magazine.

Other past assignments include the Chief Architect of the Defense Intelligence Agency's Virtual Knowledge Base Project and designer of the electronic mortgage XML standard for Fannie Mae. Mr. Daconta recently was awarded patent #7299408 by the USPTO for his electronic validator he invented for Fannie Mae. His most recent book is entitled, "Information As Product: How to deliver the right information to the right person at the right time" His other books cover the Semantic Web, XML, XUL, Java, C++ and C.

He earned his Masters degree in Computer Science from Nova Southeastern University and his bachelor's degree in Computer Science from New York University.

Abstract: *Information Management and Semantics in the Cloud*

In this presentation, Mr. Daconta will examine the emergence of cloud computing on federal IT,

NOAA Environmental Data Management Conference

Michael C. Daconta

VP Advanced Technology, InCadence Strategic Solutions

NOAA 2012 Environmental Data Management Conference

May 15 to 17 + Marriott Inn and Conference Center at UMD College Park

Keynote Speaker (Tuesday morning)

Information Management and Semantics in the Cloud



Michael C. Daconta

Vice President
Advanced Technology for InCadence Strategic Solutions

Mr. Daconta is the Vice President of Advanced Technology for InCadence Strategic Solutions (<http://www.incadencecorp.com>) where he currently guides multiple advanced technology projects for government and commercial customers. Mr. Daconta is a well-known author, lecturer and columnist having authored or co-authored 11 technical books, numerous magazine articles and online columns. He also writes the monthly "Reality Check" column for Government Computer News. Previously, Mr.

Daconta was the Metadata Program Manager for the Department of Homeland Security where he spearheaded data standardization, stewardship, and metadata registration. He was selected by the

Keynote:

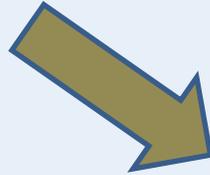
Information Management and Semantics in the Cloud

He earned his Masters degree in Computer Science from Nova Southeastern University and his bachelor's degree in Computer Science from New York University.

Abstract: *Information Management and Semantics in the Cloud*

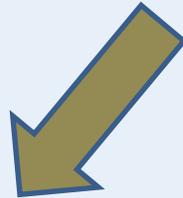
In this presentation, Mr. Daconta will examine the emergence of cloud computing on federal IT,

Lunch
until 1:15pm



Breakout Sessions 1:15-2:45 pm

- *Ballroom: Data Centers & CLASS*
- *Room 2103: Data Citation*



Plenary 3:00-5:00 pm

- *Keynote: Dave Jones, Real-Time Cross-Agency Collaboration and Data Sharing*
- *Lightning Talks*

Networking Session 5-7pm

NOAA 2012 Environmental Data Management Conference

May 15 to 17 + Marriott Inn and Conference Center at UMD College Park

Keynote Speaker (Tuesday afternoon)

Real-Time Cross-Agency Collaboration and Data Sharing for Decision Support through a Common Operating Platform (COP) - Serving Society with NOAA Data



Dave Jones

Founder, President and CEO
StormCenter Communications, Inc.

Principal Investigator
NASA Applied Sciences Program

Dubbed an "Applications Futurist" by NASA, Dave Jones combines years of experience in meteorology, broadcasting and Earth observation with a vision to enable real-time scientific data visualization and collaboration to enhance decision making. Dave worked at NASA's Goddard Space Flight Center (GSFC) as a research support meteorologist then went on to develop television weather graphics systems (which became WSI's Weather Producer) and traveled worldwide training TV meteorologists and weathercasters on how to use these systems to forecast and present weathercasts.

NOAA Environmental Data Management Conference

Dave Jones

Founder, President & CEO, StormCenter Communications, Inc.

NOAA 2012 Environmental Data Management Conference

May 15 to 17 + Marriott Inn and Conference Center at UMD College Park

Keynote Speaker (Tuesday afternoon)

Real-Time Cross-Agency Collaboration and Data Sharing for Decision Support through a Common Operating Platform (COP) - Serving Society with NOAA Data



Dave Jones

Founder, President and CEO
StormCenter Communications, Inc.

Principal Investigator
NASA Applied Sciences Program

Dubbed an "Applications Futurist" by NASA, Dave Jones combines years of experience in meteorology, broadcasting and Earth observation with a vision to enable real-time scientific data visualization and collaboration to enhance decision making. Dave worked at NASA's Goddard Space Flight Center (GSFC) as a research support meteorologist then went on to develop television weather graphics systems (which became WSI's Weather Producer) and traveled the

Keynote:

Real-Time Cross-Agency Collaboration and Data Sharing for Decision Support through a Common Operating Platform - Serving Society with NOAA Data

Lightning Talks



5 minutes



Rate the talks--award Thursday:
<https://www.nosc.noaa.gov/edmcon/>

Becky Baltes, IOOS	Establishing a Standards-based approach to Glider Data
Margarita Gregg, NODC	NODC: Implementing a progressive model for environmental data stewardship
Ted Habermann, NGDC	New Documentation Conventions for netCDF and HDF
Scott Hausman, NCDC	The Future of CLASS
Chris MacDermaid, CIRA/OAR ESRL	EarthCube Layered Architecture Concept Group
David E. McClure, Jr., Data.gov & NOAA OCIO	Open Government Data - A Service for the Government and Public
Kevin O'Brien, UW/NOAA JISAO	The Unified Access Framework (UAF)
Bob Simons, NMFS SWFSC ERD	Guidelines for Data Distribution Systems
Jebb Stewart, CIRA/OAR, ESRL	NOAA Earth Information System (NEIS)
Karen Stocks, UCSD/NSF OOI CI	Data Management in the Ocean Observatories Cyberinfrastructure

Networking Session **5-7 pm** *Room 2103*

Wednesday: Breakouts 9am-noon, 1:30-4:30

Morning Sessions	Room	Level
Breakout: Data Access and Discovery, Part 1	Room 1105	Concourse
Breakout: Data Preservation	Room 109	Lower Level
Breakout: Data Integration Projects	Room 105	Lower Level
Breakout: Ocean Biological & Chemical Data	Chasen Family Room	Concourse
Afternoon Sessions	Room	Level
Breakout: Data Access and Discovery, Part 2	Room 1105	Concourse
Breakout: Cloud Computing	Room 105	Lower
Breakout: Data Management Planning	Chasen Family Room	Concourse
Breakout: Data Documentation	Room 109	Lower

NOAA Data Citation Procedural Directive

EDMC Procedural Directives

Data Management Planning PD

Plan, in advance, how you will preserve, document and distribute your data.

Archive Procedure

What to archive, how to submit to archive.

Data Documentation

How to apply ISO 19115 metadata for discovery, use & understanding.

Data Access & Discovery

Provide on-line services so your data can be found and retrieved.

Data Sharing by NOAA Grantees

State how you will share data, and share within 2 years.

Data Citation

Use unique identifiers to allow data to be referenced and tracked.

NOAA Data Citation PD Plans

- Notional Schedule:
 - FY2012 Q4: First draft.
 - FY2013: EDMC Approval.
 - Late 2013: Begin assigning IDs.
 - 2014: Start requiring data owners to assign IDs.
 - 2015: Recommend users cite with ID.
 - Government probably cannot directly require citation.
 - 2015: Work with scientific journals to mandate data citation in publications.

NOAA Use of Digital Object Identifiers

- Planning to purchase NOAA-wide license to mint DOIs
- What level of granularity to assign DOIs?
 - May assign UUIDs at finer levels
- What is DOI target? → Metadata record?
- Allow continued use of existing unique IDs

Data Access & Discovery PD

EDMC Procedural Directives

Data Management Planning PD

Plan, in advance, how you will preserve, document and distribute your data.

Archive Procedure

What to archive, how to submit to archive.

Data Documentation

How to apply ISO 19115 metadata for discovery, use & understanding.

Data Access & Discovery

Provide on-line services so your data can be found and retrieved.

Data Sharing by NOAA Grantees

State how you will share data, and share within 2 years.

Data Citation

Use unique identifiers to allow data to be referenced and tracked.

*First draft: FY2012 Q3
Approval: FY2013*

Data Access & Discovery PD Plans

- Notional Schedule:
 - FY2012 Q3: First draft.
 - FY2013: EDMC Approval.
 - FY2014 Q1: Begin implementation.

Data Access & Discovery Background

- NOAA Administrative Order (NAO) 212-15 says:
 - Environmental **data will be visible, accessible** and independently understandable to users, except where limited by law, regulation, policy ... or by security requirements.
 - Management of NOAA environmental data [includes] Providing for **data access and dissemination**
- NAO 212-15 does not say how
- NAO 212-15 only hints at discovery ("visible")

Typical Problems

- Data in unusual formats
- Datasets not made available by owner
- Data service available but not registered in catalog
- Data accessible only via a web-based user interface
- Data only in prepackaged data files via FTP or HTTP
- Data only distributed via GTS
- Proprietary or custom data services
- Offline data with ordering mechanism
 - *What is an appropriate cutoff for recent/online vs older/offline data?*
- Services with incomplete or insufficient metadata
- Catalogs that cannot be federated (no distributed search)

Data Access & Discovery PD Scope?

Data Access & Discovery PD:

Appropriate middle ground



NAO 212-15:

Conceptual -

"Thou shalt make
thine data
accessible"

EDM Wiki:

Detailed guidance regarding
specific service versions,
software, best practices, profiles,
terminology, formats

Potential Topics in PD

- Policy background & rationale
- Data Access
 - Overview of service types: DAP, OGC, Esri, ...
 - Overview of data formats: CF/NetCDF, HDF, ASCII, XML, GML, Shapefile, ...
 - Recommended formats & service types for classes of data:
 - Remote obs, in situ obs, model outputs, cadastral, ...
 - Regular grids, irregular grids, point, profile, trajectory, polygon, imagery, visualization
 - Pull (user request) vs Push (subscribe, alert) services
 - Service metadata content
 - Registration of service in catalog
- Discovery
 - Approved NOAA Catalog instances
 - catalog search interfaces
- Requirement to seek existing solutions before implementing new ones
- Shared services (hosting, catalog)
- Links to EDM Wiki

Seeking

- Advice on scope of PD
- Specific guidance to include
- Case studies/successes/warnings
- Contributors of text
- Reviewers