



NOAA EDM

NOAA Environmental Data Management

Report to Unidata Policy Cmtee 2013-05-15

Jeff de La Beaujardière, PhD
NOAA Data Management Architect
jeff.deLaBeaujardiere@noaa.gov
+1 301-713-7175

Overview

- Vision for NOAA Enviro. Data Mgmt (EDM)
- EDM Framework (*SAB action*)
- EDM Dashboard
- EDM Virtual Workshop
- EDM Assessment of Systems of Record
- Data Citation pilot project
- Recent Presidential directives

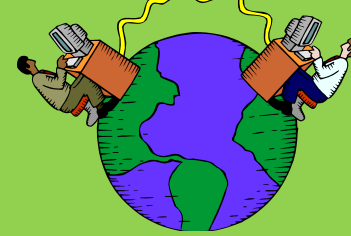
Vision for NOAA Data Management

All NOAA data will be:

• Discoverable



• Accessible



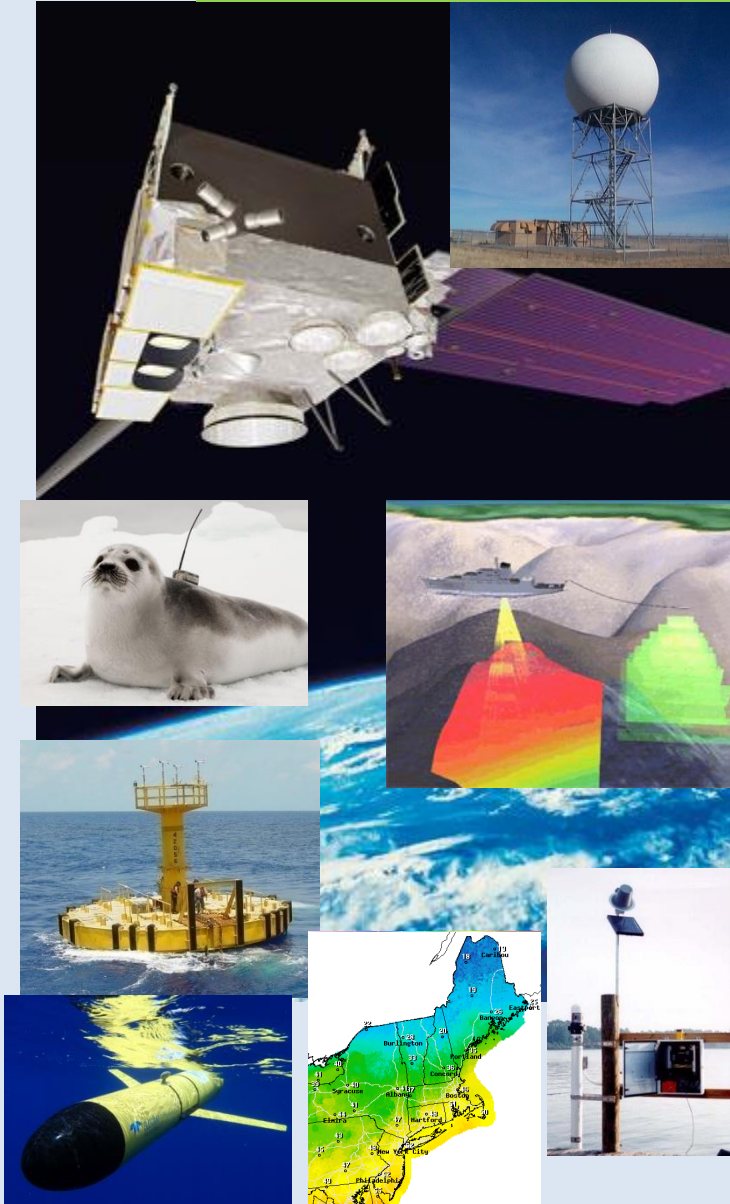
• Documented



• Preserved



for all types of users and applications.



NOAA Environmental Data Management Framework

Data Management Framework

Principles				
Governance				
Resources				
Standards				
Architecture				
Assessment				
Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle

- Purpose: To organize, guide and support NOAA environmental data management activities.
- Mandate: Science Advisory Board (SAB) recommendation to NOAA.
- <https://www.nosc.noaa.gov/EDMC/framework.php>

Data Management Framework

Data Management Framework

Principles				
Governance				
Resources				
Standards				
Architecture				
Assessment				
Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle

Principles

- Full and Open Access
 - except in very limited cases
- Data Preservation
 - for long-term usability
- Information Quality
 - known quality data, complete metadata
- Ease of Use
 - compatible services, formats, vocabularies

Data Management Framework

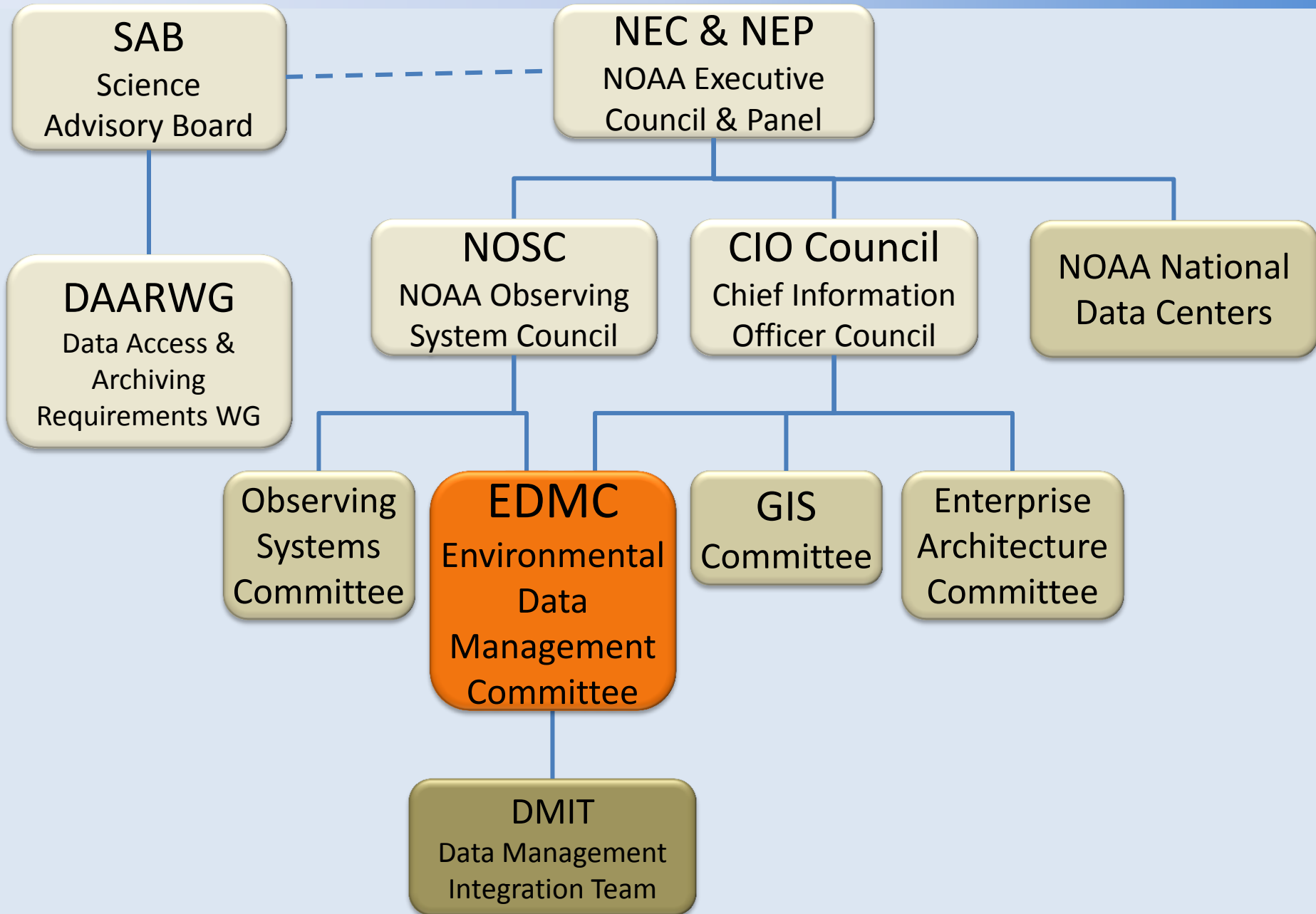
Data Management Framework

Principles				
Governance				
Resources				
Standards				
Architecture				
Assessment				
Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle

Governance

- NOAA Bodies
 - incl. EDMC
- NOAA Policies
 - incl. EDMC PDs
- US Policies
 - incl. OSTP PARR memo
- External Coordination

NOAA EDM Governance Bodies



EDMC Procedural Directives

(Environmental Data Management Committee)

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 Sharing by NOAA Grantees

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

Data Citation

Assign persistent identifiers to datasets and encourage citation.

Data Access

Establish & improve on-line services for data access

*in prep.
(2013)*

Public Access to Research Results (PARR)

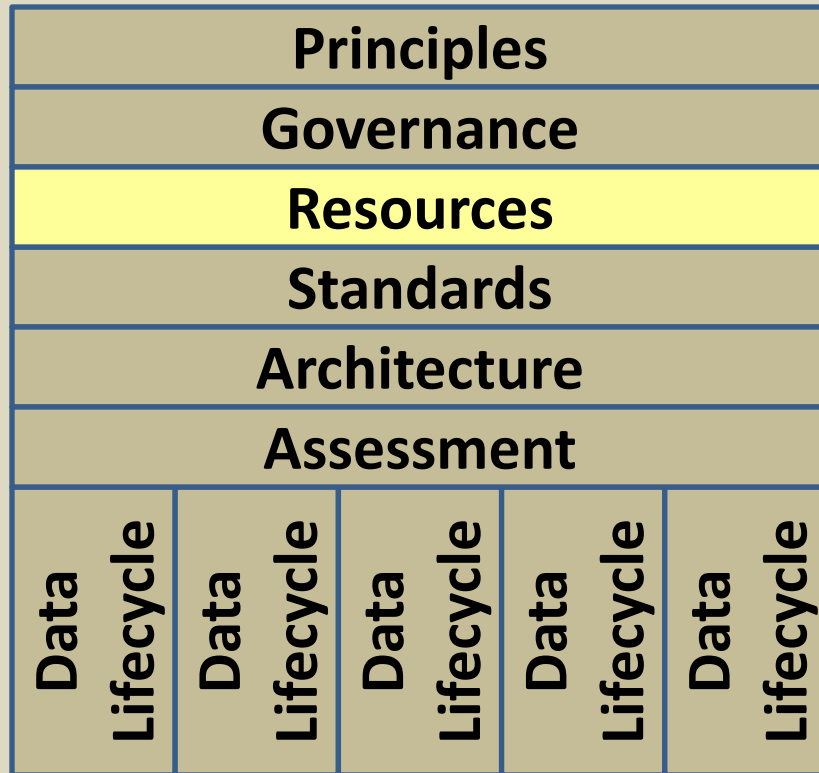
- Memo from White House Office of Science and Technology Policy (OSTP): "Increasing Access to the Results of Federally Funded Scientific Research"
 - http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf
 - Applies to "Publications" and "Digital Data"
 - Focus is more on policy than technology
 - Draft plans from each Agency due 2013 Aug 22
- Federal activity:
 - Interagency meetings hosted by OSTP
- NOAA activity:
 - PARR Cmtee established by NOAA Research Council to draft plan
 - Co-chairs: Jeff DLB (EDMC), Neal Kaske (NOAA Library)

Executive Order (May 9, 2013)

- **"Executive Order -- Making Open and Machine Readable the New Default for Government Information"**
 - <http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government->
 - Coupled with:
 - Open Data Policy -- Managing Information as an Asset
 - Implementation Guide
- Requires
 - Machine-readable inventory of agency data assets
 - Use of open standards and formats
 - Life-cycle data management planning
- Initial efforts by November 9, 2013

Data Management Framework

Data Management Framework



Resources

- Budget
 - Project-specific
 - NOAA-wide
- Personnel
 - Training
 - Recognition
 - Authority
- Other Resources
 - Annual Workshop
 - Teams
 - Wiki

EDM Virtual Workshop

- NOAA-wide Virtual Workshop
 - All participants connecting remotely via webinar software
- June 25-27, 13:00-16:30 EDT
- Theme: NOAA EDM: current state, target state, next steps
- Six 90-minute sessions:
 - Intro & Overview
 - Catalog & Search
 - Data Access
 - Data Usability
 - Preservation & Citation
 - Wrap-up & Final Discussion

(tentative)



Data Management Framework

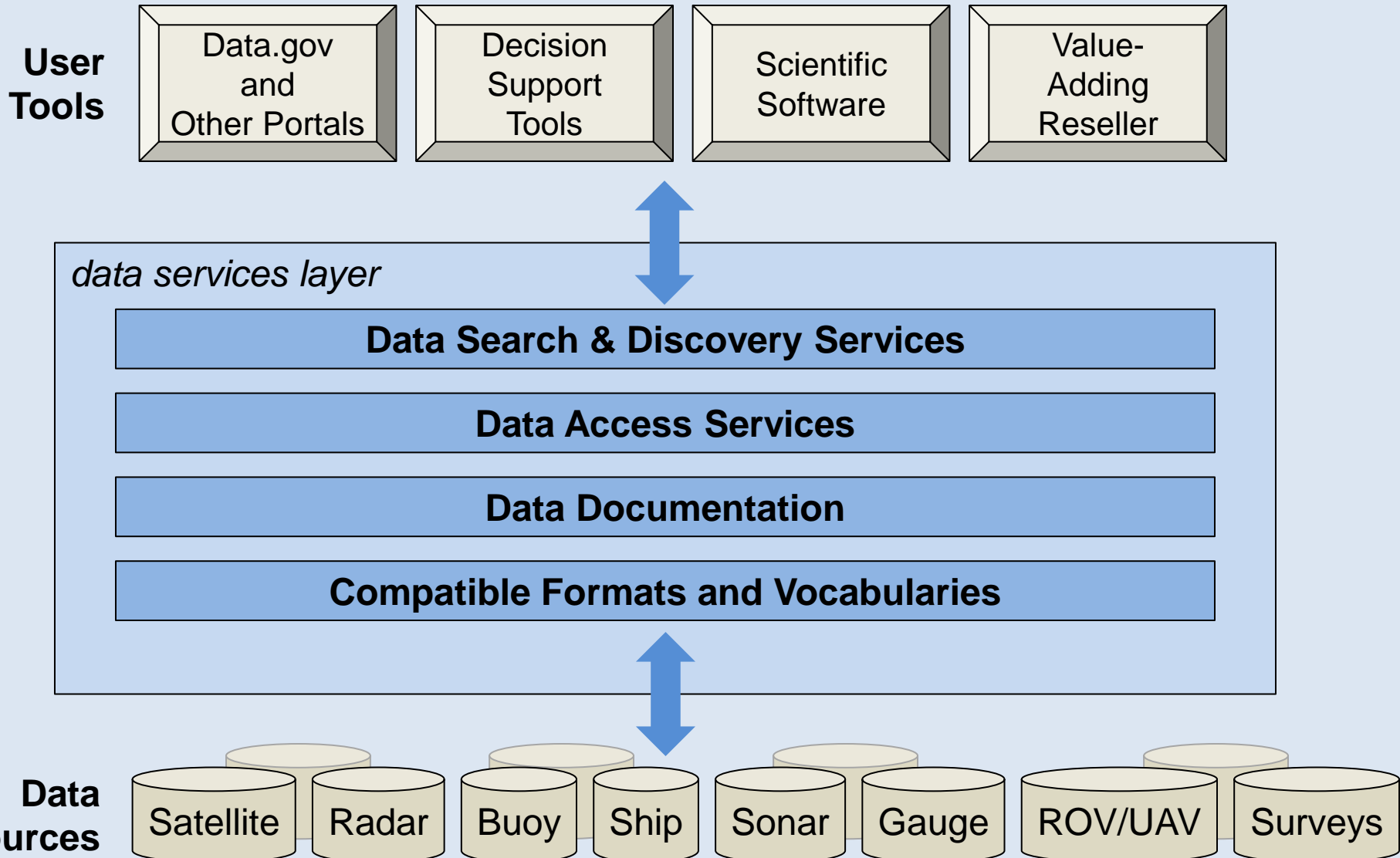
Data Management Framework

Principles				
Governance				
Resources				
Standards				
Architecture				
Assessment				
Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle

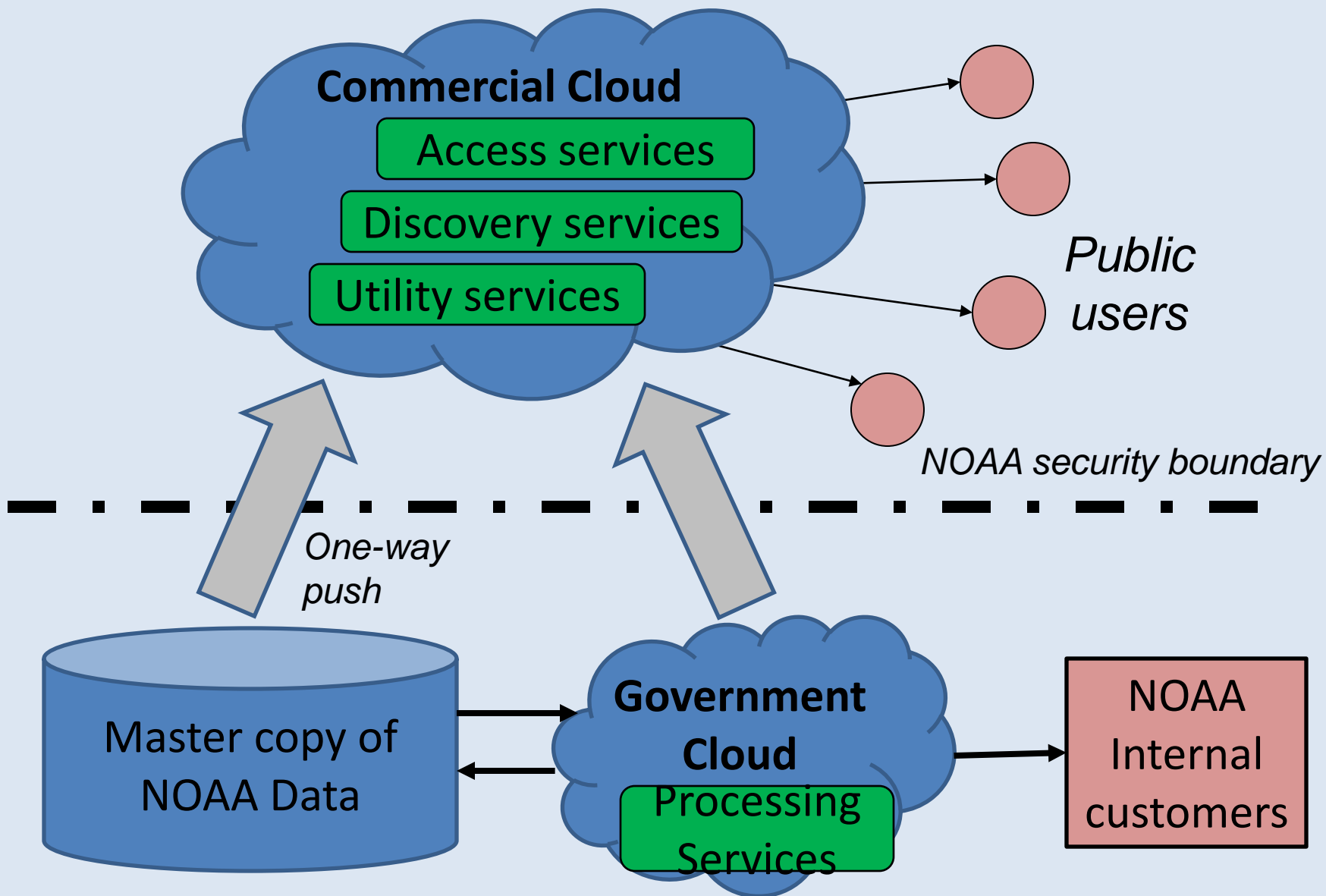
Architecture

- Service-based approach
- Designing for flexibility
 - ability to leverage Cloud & other technologies
- National Data Centers
- Legacy systems and agreements

Data Services Layer



Potential Cloud Deployment Scenario



Data Management Framework

Data Management Framework

Principles				
Governance				
Resources				
Standards				
Architecture				
Assessment				
Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle

Assessment

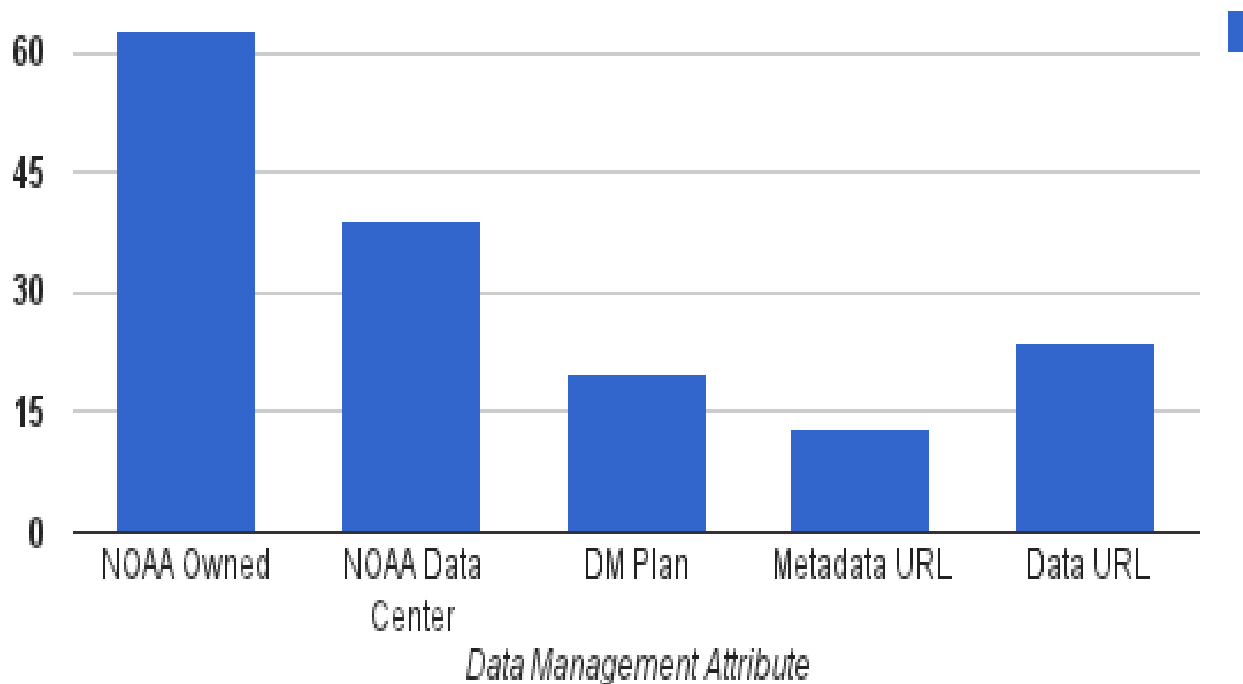
- Current state
 - Observing System of Record EDM study
- Progress measurement
 - EDMC Reporting
 - EDM Dashboard
- Feedback from users & implementers

Obs. System of Record EDM Assessment

- Goal: For NOAA-owned Observing Systems of Record ($63 \leq N \leq 86$), determine
 - Data Management plan existence/location
 - Data Center used for long-term preservation
 - Metadata location & format
 - Data access services offered

**CORL=Consolidated Observing Requirements List.
NOSA=NOAA Observing Systems Architecture.*

EDM Assessment of Observing Systems of Record



EDM Dashboard



NOAA EDM

NOAA Environmental Data Management

EDM Dashboard

[Project Description](#)

[The Team](#)

[Project Calendar](#)

[Tasks & Actions](#)

[Documents](#)

[Issues and Comments](#)

[Sitemap](#)

Recent site activity

[EDM Dashboard](#)

edited by Jeff de La Beaujardiere - NOAA Federal

[Issues and Comments](#)

comment from Denise Blake - NOAA Federal

comment from Jason Chasse - NOAA Federal

comment from Jennifer Bosch - NOAA Affiliate

comment from Jennifer Bosch - NOAA Affiliate

comment from Jennifer Bosch - NOAA Affiliate

[The Team](#)

edited by Martin Yapur - NOAA Federal

[Issues and Comments](#)

edited by Jeff de La Beaujardiere - NOAA Federal

- NOAA Federal

EDM Dashboard

<http://sites.google.com/a/noaa.gov/edm-dashboard/>

(Under Development)

(internal access only)

NOTE: Explanations of each chart may be found on the [Project Description](#) page.

EDMC Work Plan Status

<u>Work Plan Item</u>	<u>Due</u>	<u>On Track?</u>	<u>Status</u>	<u>Data</u>
Develop EDM Framework (SAB action)	2013 Mar	DONE	DONE. Version 1.0 posted on EDMC web site.	https://www.nosc.noaa.gov/EDMC/framework
Develop External Data Usage Recommended Practice (SAB action)	2013 Mar	DONE	DONE. Version 1.0 posted on EDMC web site.	https://www.nosc.noaa.gov/EDMC/external
Host the 2013 EDM Virtual Workshop (June 25-27, 1-4pm)	2013 Jun	Green	In-person conference with travel not authorized. Planning Virtual Workshop instead.	http://goo.gl/L26Lr
Implement and monitor approved Procedural Directives	ongoing	Green	Reporting mechanism established via Google Drive. Monthly input from LO reps.	http://goo.gl/5KEe0
Write new Data Citation PD	2013 Sep	Yellow	Small team meeting by telecon. Working on dataset requirements and landing page creation mechanism.	http://goo.gl/ENkck
Write new Data Access PD	2013 Sep	Yellow	Outline of PD drafted. Increasing priority in light of OSTP PARR memo and Bg Earth Data Initiative.	http://goo.gl/0SXMS
Perform EDM Assessment of Observing Systems of Record	2013 May	Yellow	EDMC issued data call 2013-03-14 at NOSC direction. Responses due 2013-04-26. ~16% received so far.	http://goo.gl/Auztv
Establish prototype EDM Dashboard	2013 Feb	DONE	DONE. Demo to CIO Council Feb 12, NOSC Mar 7	https://sites.google.com/a/noaa.gov/edm-dashboard/

Data Management Framework

Data Management Framework

Principles				
Governance				
Resources				
Standards				
Architecture				
Assessment				
Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle	Data Lifecycle

Data Lifecycle

*Planning and
Production
Activities*

Requirements Definition
Planning
Development
Deployment
Operations

*Data
Management
Activities*

Collection
Processing
Quality Control
Documentation
Cataloging
Dissemination
Preservation
Stewardship
Usage Tracking
Final Disposition

*Usage
Activities*

Discovery
Reception
Understanding
Analysis
Value-Added Products
User Feedback
Citation
Tagging
Gap Assessment

Data Lifecycle Activities

Data Lifecycle

Planning and Production Activities

Requirements Definition
Planning
Development
Deployment
Operations

Applicability of EDMC Directives

DM Planning

Data Management Activities

Collection
Processing
Quality Control
Documentation
Cataloging
Dissemination
Preservation
Stewardship
Usage Tracking
Final Disposition

Data Documentation

Data Sharing by Grantees

Data Access

Archive Procedure

Usage Activities

Discovery
Reception
Understanding
Analysis
Value-Added Products
User Feedback
Citation
Tagging
Gap Assessment

Data Citation

Data Lifecycle

*Planning and
Production
Activities*

Requirements Definition
Planning
Development
Deployment
Operations

*Data
Management
Activities*

Collection
Processing
Quality Control
Documentation
Cataloging
Dissemination
Preservation
Stewardship
Usage Tracking
Final Disposition

*Usage
Activities*

Discovery
Reception
Understanding
Analysis
Value-Added Products
User Feedback
Citation
Tagging
Gap Assessment

**Focus of NOAA
Data Citation
pilot project**

NOAA Data Citation Pilot Project

- Goals:
 - assign persistent identifiers to archival datasets
 - enable citation of datasets used in results
 - encourage archival submission & complete metadata
 - enable usage tracking
- Status:
 - Have license to mint DOIs
 - Established team of Data Center reps + DM Architect
 - Working out technical details
 - metadata reqmts, landing page creation, dataset granularity
 - Hope to have first DOIs assigned by June

