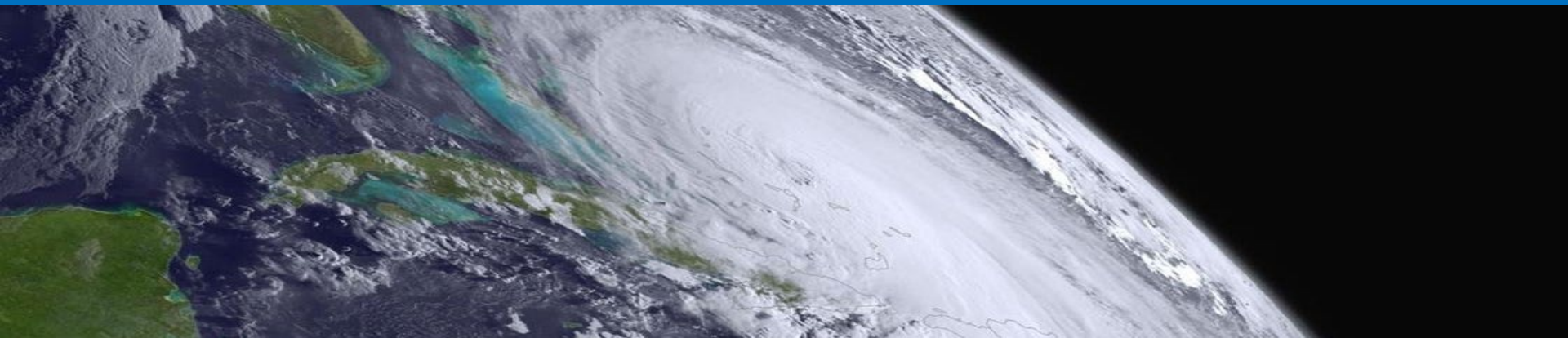# NOAA

Tony LaVoi
NOAA Chief Data Officer

April 27, 2022

NOAA Data Governance Supporting Open Science and Open Data
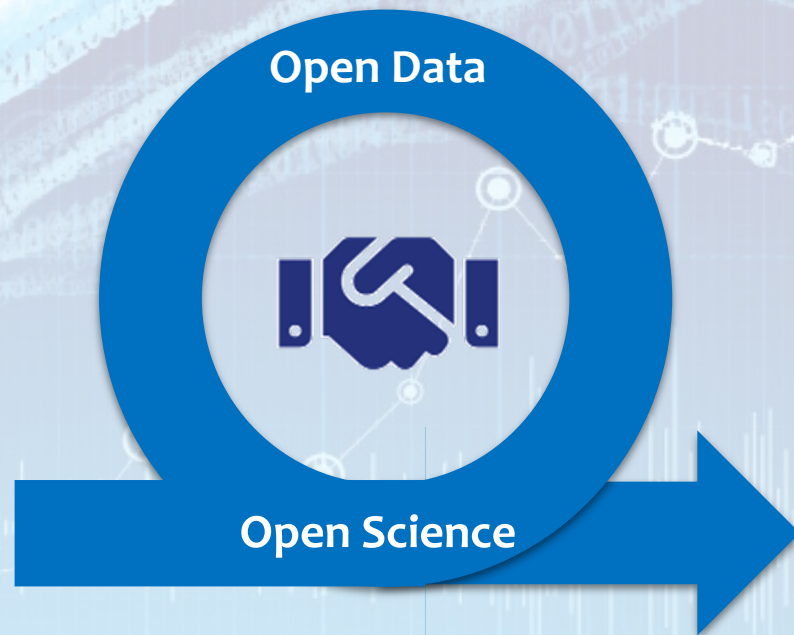
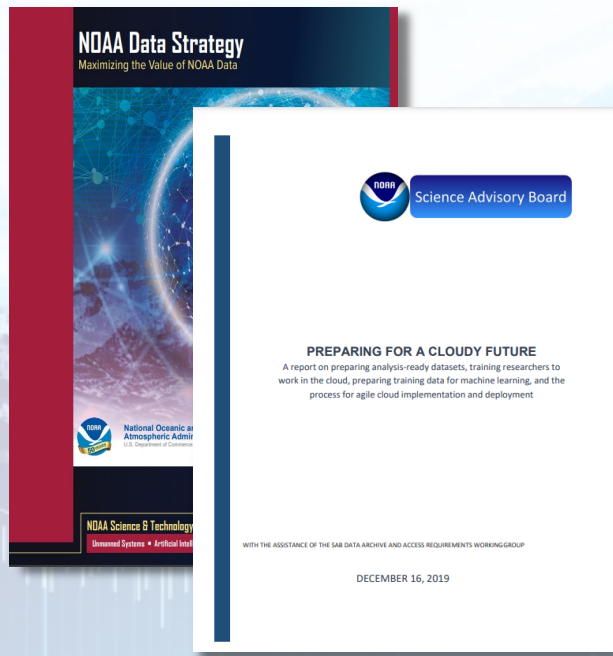**Response to DAARWG Recommendations on the NOAA Data Strategic Action Plan**
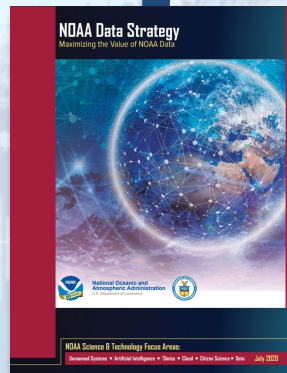
DAARWG Engagement & New Opportunities

Open Data

Open Science

"Evidence Act"

NOAA Data Strategy

Federal Data Strategy & Administration Priorities

DOC Data Strategy

NOAA Strategic Plan

CLIMATE

NEW BLUE ECONOMY

EQUITY

NOAA Data Strategic Action Plan
Maximizing the Value of NOAA Data

National Oceanic and Atmospheric Administration
U.S. Department of commerce

NOAA Science & Technology Focus Areas:
Uncrewed Systems • Artificial Intelligence • 'Omics • Cloud • Citizen Science • Data   June 2021

High Priority FY22-FY23 Actions

Data Governance

Data Policy Framework

Data Workforce

Open Data

NDSAP FY22 - 23 High Priority Actions

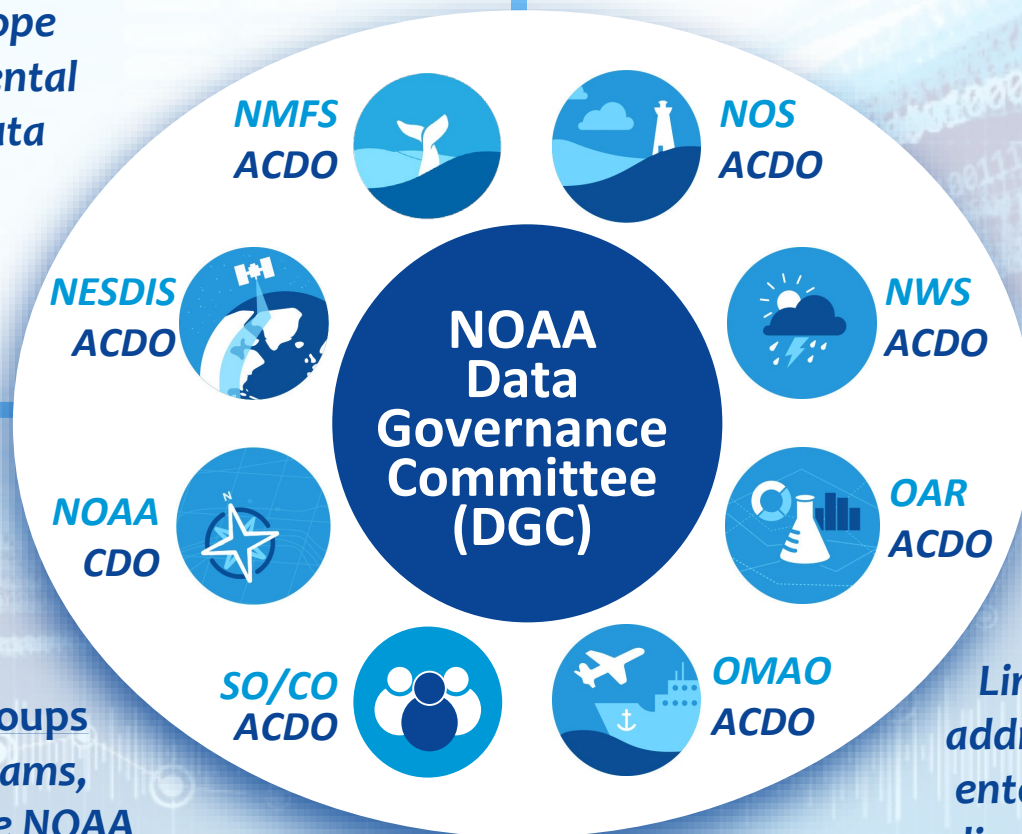| Grouping	Priority Actions	Status
Governance	Action 1: Establish the NOAA CDO structure	In Progress / NOAA CDO and ACDOs
	Action 3: Formalize LO data governance roles	In Progress / ACDOs
	Action 4: Establish a NOAA Data Governance Body	Complete / Data Governance Committee
	Action 5: Socialize and promote the NOAA Data Governance	In Progress / Comms Strategy
Policy	Action 6: Update NOAA EDMC Data Directives	In Progress / NAO and Handbook
	Action 11: Develop NOAA data licensing guidance	In Progress / Data Licensing Task Team
Data Workforce	Action 13: Assess and report NOAA's data program maturity	In Progress / Commerce Coordination
	Action 16: Assess and increase staff data literacy and skills	In Progress / Workforce Task Team & OHCS
Open Data	Action 8: Develop a NOAA Open Data Plan	Pending OMB Guidance
	Action 9: Assess baseline compliance with Open Data rqmt's	Pending OMB Guidance
	Action 10: Maintain a comprehensive NOAA data inventory	In Progress / Catalog Refresh
Metrics	Action 22: Track the use of NOAA data	In Progress / NODD and Chief Economist

Scope
Broadening scope from environmental to all NOAA data

Roles & Responsibilities
Expanded roles, responsibilities and relationships

NMFS ACDO

NOS ACDO

NESDIS ACDO

NWS ACDO

NOAA CDO

OAR ACDO

SO/CO ACDO

OMAO ACDO

NOAA Data Governance Committee (DGC)

Working Groups
Persistent teams, addressing core NOAA enterprise data topics

Task Teams
Limited-duration teams addressing a specific NOAA enterprise data topic with direct NOAA Data Strategy linkages

Expanding the NOAA Data Community

Working Groups (WG)	Task Teams (TT)
1. NOAA GIS WG	1. Data Digital Object Identifier (DOI) TT
2. NOAA Enterprise Metadata WG	2. Data Licensing TT
3. Government Information Services WG	3. Environmental Data Policy TT
4. Data Innovation WG	4. Data Catalog TT
5. Environmental Data Management Workshop WG	5. Workforce Development TT
6. Data Communications WG	

NOAA Data Dissemination Report

Key question: Is there more we can do, as a leading Federal data agency, to provide better access to data & information for our users?
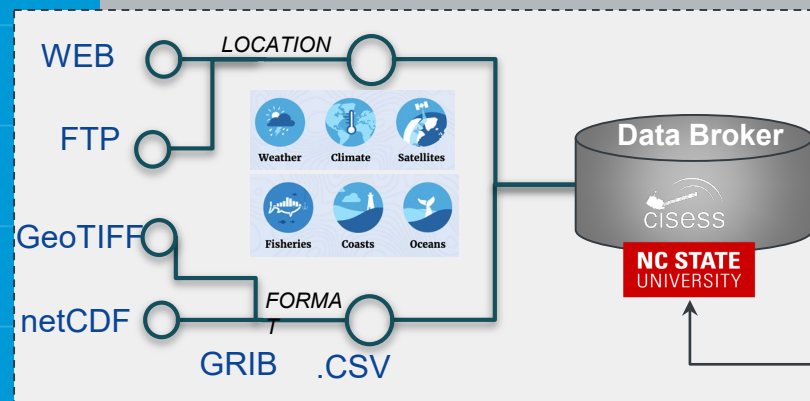
1. **Need**: a future-looking vision and plan to align data dissemination programs at the enterprise and mission-area levels
2. **Foundation**: NOAA Data & Cloud Strategies, Evidence Act, FAIR Principles, Open Science/Open Data initiatives, and existing NOAA systems (NODD, CLASS, IDP, Digital Coast)
3. **Informed by**: user feedback on dissemination methods & systems; key challenges faced by NOAA application owners and data managers in disseminating data
4. **Result**: report that characterizes current data dissemination methods, documents ongoing improvements, and provides requirements, recommendations, and resource estimates needed for future dissemination approaches
5. **When**: draft report by the end of September 2022

When data is in diverse formats and from different locations, we need to establish methods to tackle big data components such as portability, reusability, extensibility, analytics and infrastructure in order to get to interoperability.

NOAA data access spans disparate formats and locations

WEB

FTP

GeoTIFF

netCDF

LOCATION

FORMAT

GRIB   .CSV

Weather   Climate   Satellites

Fisheries   Coasts   Oceans

Data Broker

cisess

NC STATE UNIVERSITY

DATA CONVERSION

Google Cloud Platform

Microsoft Azure

PANGEO

amazon web services

DATA TRANSFER

- Centralized location
- Documentation
- Performance

End users' ability to have accelerated analytical capabilities (via optimized access and formats) build new tools and capabilities that serves the public

NOAA Open Data Dissemination
Cloud Data Analytics Summary
Analytics Team: CISESS-NC
Version 0.4.0 | April 14, 2022

19.81 PB
Total Storage

NODD-Wide Monthly Metrics (Last 31 Days)

Week | Month

Combined metrics from Amazon Web Services (AWS), Microsoft Azure (Azure), and Google Cloud Platform (GCP).

Accessions ⓘ   Ratio

5.55 PB

Interactions ⓘ

18496 M

Uploads ⓘ   Ratio

902 TB

By Cloud Service Provider   By NOAA Line Office

NOAA is Leading Development of an AI-Ready Data Standard

Data Preparation

Data Quality

AI-Ready Data

Data Documentation
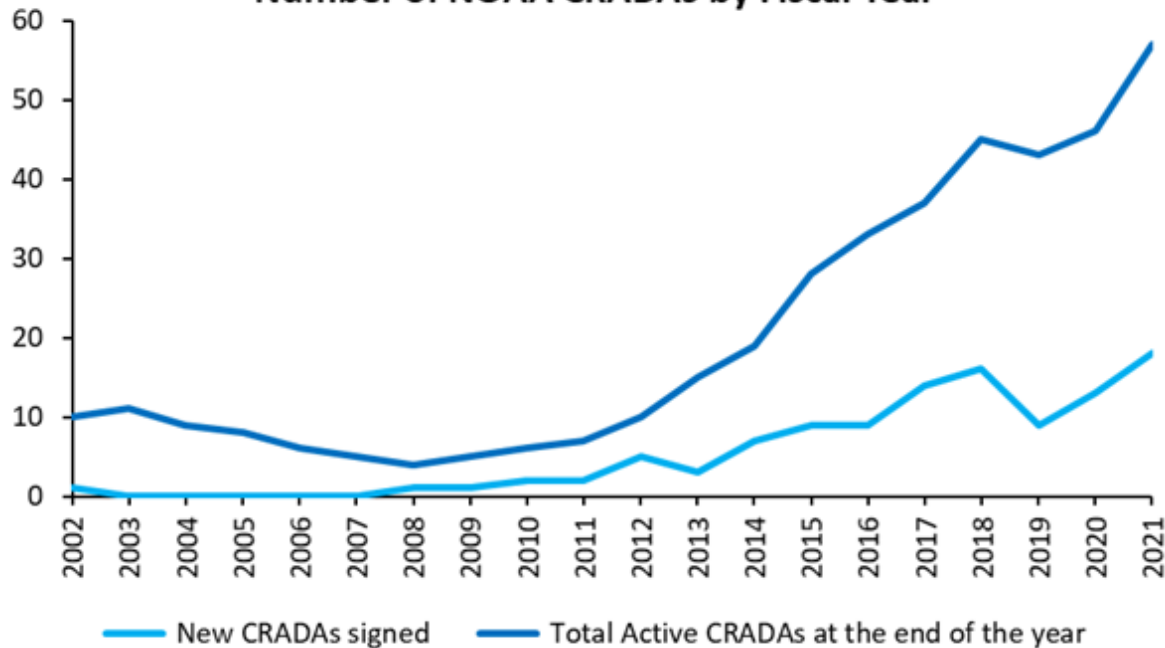
Data Access

Collaboration under the Earth Science Information Partners (ESIP)

An AI-Ready Data standard is critical to ensure NOAA Data are usable for AI/ML applications

Partnerships Fueling Innovation

Number of NOAA CRADAs by Fiscal Year

New CRADAs signed — Total Active CRADAs at the end of the year

NOAA

NOAA

Graph: https://techpartnerships.noaa.gov/2022/01/11/public-private-research-partnerships-are-fueling-noaa-innovation/

Images: https://www.noaa.gov/stories/these-5-technologies-are-helping-save-our-ocean
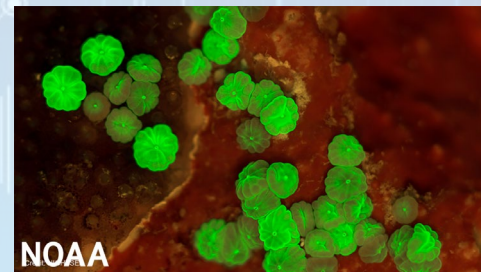
Take Home Message

| NOAA Data Governance | Open Science Open Data | Partnerships Innovation |

CLIMATE

NEW BLUE ECONOMY

EQUITY

BACKUP SLIDES

Importance of Data Governance

- Process of **setting and implementing priorities** and policies for managing data as a strategic asset in service of an agency's mission.

- Includes the **people, processes, and systems** needed to maximize and protect the value of NOAA's data assets.

- **Improves decision-making** and provides a **foundation for communication and collaboration** across disparate business systems and processes.

- Ensures that NOAA has **sufficient human and financial resources** available to meet data-driven challenges.

- Three interrelated components: **DGC, ACDOs, and Data Policy Framework**

Data Governance Portfolios

Data Governance and Policy *(Data Strategy Goals 1, 2, & 3)*

- Govern and manage data strategically throughout the policy process to most effectively steward the US taxpayers' investment.

Data Lifecycle Management and Integration *(Data Strategy Goals 2, 3 & 4)*

- Coordinate and support implementation of data management responsibilities across the agency while ensuring alignment of data management life-cycle architectures, applications, and technology innovations.

Workforce Capabilities, Capacity, and Innovation *(Data Strategy Goal 4)*

- Build and maintain the necessary technical tools and staff skills to support data driven science and decision making while cultivating creativity and innovation by developing and supporting a workforce with diverse backgrounds, experiences, and expertise in an inclusive environment.

Stakeholder Engagement and Collaboration *(Data Strategy Goal 5)*

- Identify and engage internal and external stakeholders and leverage partnerships to maximize the value of NOAA data to the Nation.

DAARWG Recommendations 1-4

#	DAARWG Recommendation	NOAA Response
1	The Plan should state who ultimately owns and evolves the Data Strategy for NOAA - is it the CDO, or another individual or body?	The NOAA CDO is the senior executive responsible for the execution of the Data Strategic Action Plan. This is now included in the NDSAP introduction.
2	The Plan should clarify the source of the CDOs budget and how it can be protected from cuts in future years.	A discussion of the CDO budget was not included in the NDSAP as it was considered out of scope for the document.
3	Carefully consider which groups should actually have CDO representation, and ensure all groups follow existing and future data-related policies.	The CDO is an official member of the CIO and Observing System Councils, as well as a participant in other NOAA councils and committees. Additionally, Line Office Assistant Chief Data Officer (ACDO) positions are being established. The LO ACDOs will have a matrixed reporting relationship to the NOAA CDO.
4	Either (a) very clearly delineate the scope and relationship of DGB & EDMC, or (b) rename EDMC as DGB, or (c) recharter EDMC to include the functions and authority intended for DGB.	The EDMC has been rechartered as the NOAA Data Governance Committee (DGC). A new Terms of Reference was approved to clarify the scope, roles and responsibilities of the DGC. The transition from the EDMC to the DGC, as well as the establishment of LO ACDOs, is taking place in FY22Q3-4.

DAARWG Recommendations 5-8

#	DAARWG Recommendation	NOAA Response
5	If a separate (non-EDMC) body is to be created, consider following established NOAA nomenclature and having it be called the Data Council; EDMC could then report to the Data Council instead of to both the NOAA Observing Systems Council and CIO Council as currently done.	The EDMC has been rechartered as the NOAA Data Governance Committee (DGC). A new Terms of Reference was approved to clarify the scope, roles and responsibilities of the DGC. The DGC will continue to report to the NOSC and the CIO Council.
6	Consider whether the CDO should be Chair ex officio of the Data Council, DGB, or EDMC, as appropriate.	The new Terms of Reference for the DGC establishes the CDO as the DGC Chair.
7	The Plan should avoid duplicating or rewriting past policies [with the NOAA Open Data Plan] unless there is an identified shortcoming with the prior documents	OMB guidance on agency Open Data Plans is still pending, as well as the subsequent Commerce guidance that will be issued after OMB; NOAA's Action 3.2 to publish an Open Data Plan is on hold pending this guidance.
8	The Plan should avoid duplicating past actions or reinventing past methods [with the data inventory] unless there is an identified shortcoming with the prior approach.	The NOAA Data Catalog is the harvesting site for NOAA data to data.gov. However, due to several issues and complications with the harvesting, a task team is being stood up by the CDO and DGC to address the outstanding issues.

DAARWG Recommendations & NOAA Response

DAARWG/SAB Comments applicable to Recommendations 1-2:
- *Action 1.1 Establish the NOAA CDO organizational structure.*
- *1.1.1 Define the roles and responsibilities of the NOAA CDO.*
- *1.1.2 Establish a budget for the NOAA CDO to ensure the resources are available to have a meaningful impact.*

- **DAARWG/SAB Recommendation 1**: The Plan should state who ultimately owns and evolves the Data Strategy for NOAA - is it the CDO, or another individual or body?

  - **NOAA Response 1**: The NDSAP states the following [Ref: Page 6, last paragraph in the "Timeline, Roles, and Priorities" Section]:

    "The NOAA Data Strategic Action Plan lays out how the NOAA CDO and existing bodies such as the NOAA Environmental Data Management Committee (EDMC), will make progress toward the goals and objectives of the Data Strategy while leveraging expertise and established groups in NOAA Line Offices and Staff Offices. The NOAA CDO will ultimately oversee the execution of the Data Strategic Action Plan. Goals 1 and 2 of the Data Strategy focus on establishing the NOAA CDO organizational structure, governance, and supportive processes within the line/staff offices to enable the execution of the actions in this plan. Therefore, the priority actions for the remainder of FY22–23 will advance these goals."

DAARWG Recommendations & NOAA Response

DAARWG/SAB **Comments applicable to Recommendations 1-2:**
- *Action 1.1 Establish the NOAA CDO organizational structure.*
- *1.1.1 Define the roles and responsibilities of the NOAA CDO.*
- *1.1.2 Establish a budget for the NOAA CDO to ensure the resources are available to have a meaningful impact.*

- **DAARWG/SAB Recommendation 2**: The Plan should clarify the source of the CDOs budget and how it can be protected from cuts in future years.

  - **NOAA Response 2**: The NDSAP states the following [Ref: Page 6, Goal 1, Actions 1 & 3]:

    "*Goal 1 aligns NOAA enterprise data governance roles establishing collaborative leadership, empowered by policy and budget, with representation on NOAA's strategic and decision-making councils and advisory boards.*"

    - Action 1) Establish the NOAA CDO organizational structure.
      - Milestones and Deliverables:
        - A. Define the roles and responsibilities of the NOAA CDO.
        - B. Establish a budget for the NOAA CDO to ensure the resources are available to perform the functions of the office.
    - Action 3) Formalize NOAA Line Office data governance and roles.
      - Milestones and Deliverables:
        - D. Assist Line Office leadership in defining budget and staffing requirements.

DAARWG Recommendations & NOAA Response

__DAARWG/SAB Comment applicable to Recommendation 3:__ Having the CDO represented on every NOAA council and committee that has a "data connection" might stretch the office too thin, and would not necessarily reduce stovepiping unless the CDO is empowered to approve or correct work by those other bodies.

- **DAARWG/SAB Recommendation 3**: Carefully consider which groups should actually have CDO representation, and ensure all groups follow existing and future data-related policies.

  - **NOAA Response 3**: Recommendation is noted. The CDO is an official member of the CIO and Observing System Councils, as well as a participant in other NOAA councils and committees. The NOAA CDO proposed a plan to establish Line Office Assistant Chief Data Officer (ACDO) positions to lead Line Office data governance requirements and report up to the NOAA CDO. This has been briefed and approved by the NOAA CIO Council (and with each of the NOAA LO ACIOs), and the NOSC. [Reference: Draft NOAA LO ACDO Recommendations document here for more details].

DAARWG Recommendations & NOAA Response

DAARWG/SAB Comments applicable to Recommendations 4, 5 & 6: *We are most concerned about the scope and function of the proposed Data Governance Body (DGB) and its relation to the existing Environmental Data Management Committee (EDMC). Having both DGB and EDMC may be confusing, counterproductive, and a dilution of resources. Also, we note that NOAA already has Councils and Committees, but does not seem to have any other "governance bodies." The Federal Data Strategy cited in Action 2.1 says data governance bodies "are commonly called by such names as Data Governance Boards, Data Councils, or Data Strategy Teams."*

- **DAARWG/SAB Recommendation 4:** Either (a) very clearly delineate the scope and relationship of DGB & EDMC, or (b) rename EDMC as DGB, or (c) recharter EDMC to include the functions and authority intended for DGB.

  - **NOAA Response 4**: The EDMC is undergoing a transition to the NOAA Data Governance Committee (DGC). A new Terms of Reference was developed to clarify the scope, roles and responsibilities of the new NOAA DGC. The ToR was approved by the EDMC, CIO Council, and currently pending final approval by the NOSC. In addition to the new ToR, we have also been working on other EDMC to DGC transition related materials including:

    - Summary Transition from EDMC to NDGC document;

    - EDMC to DGC Transition Plan gantt chart

DAARWG Recommendations & NOAA Response

DAARWG/SAB Comments applicable to Recommendations 4, 5 & 6: *We are most concerned about the scope and function of the proposed Data Governance Body (DGB) and its relation to the existing Environmental Data Management Committee (EDMC). Having both DGB and EDMC may be confusing, counterproductive, and a dilution of resources. Also, we note that NOAA already has Councils and Committees, but does not seem to have any other "governance bodies." The Federal Data Strategy cited in Action 2.1 says data governance bodies "are commonly called by such names as Data Governance Boards, Data Councils, or Data Strategy Teams."*

- **DAARWG/SAB Recommendation 5**: If a separate (non-EDMC) body is to be created, consider following established NOAA nomenclature and having it be called the Data Council; EDMC could then report to the Data Council instead of to both the NOAA Observing Systems Council and CIO Council as currently done.
    - **NOAA Response 5**: A new Terms of Reference was developed to clarify the scope, roles and responsibilities of the new NOAA Data Governance Committee (DGC). The ToR was approved by the EDMC, CIO Council, and currently pending final approval by the NOSC.

DAARWG Recommendations & NOAA Response

__DAARWG/SAB Comments applicable to Recommendations 4, 5 & 6:__ We are most concerned about the scope and function of the proposed Data Governance Body (DGB) and its relation to the existing Environmental Data Management Committee (EDMC). Having both DGB and EDMC may be confusing, counterproductive, and a dilution of resources. Also, we note that NOAA already has Councils and Committees, but does not seem to have any other "governance bodies." The Federal Data Strategy cited in Action 2.1 says data governance bodies "are commonly called by such names as Data Governance Boards, Data Councils, or Data Strategy Teams."

- **DAARWG/SAB Recommendation 6**: Consider whether the CDO should be Chair ex officio of the Data Council, DGB, or EDMC, as appropriate.

  - **NOAA Response 6**: The CDO will be the ex officio Chair of the NOAA Data Governance Committee (DGC). This is stated in the Terms of Reference approved by the EDMC, CIO Council, and currently pending final approval by the NOSC.

DAARWG Recommendations & NOAA Response

_DAARWG/SAB Comments applicable to Recommendations 7__:_
- _Action 3.2: Publish and promote a NOAA Open Data Plan._
- _How would the proposed "NOAA Open Data Plan" differ from the NOAA Plan for Public Access to Research Results (PARR Plan) issued in 2015, in response to the OMB Holdren Memo of 2013? NOAA seems to have already satisfied this Federal Data Strategy requirement, and has long been ahead of many other agencies._

- **DAARWG/SAB Recommendation 7**: The Plan should avoid duplicating or rewriting past policies unless there is an identified shortcoming with the prior documents

    - **NOAA Response 7**: The NOAA EDMC Procedural Directives are dated back to 2008 and some have not been reviewed in 14 years. Given the new federal policies and legislation such as the Evidence Act, and the Federal Data Strategy's annual Action Plan, these PDs need to be updated.

    - Due to the fact that OMB guidance on agency Open Data Plans is still pending, as well as the subsequent Commerce guidance that will be issued after OMB, NOAA's Action 3.2 to publish an Open Data Plan is on hold.

DAARWG Recommendations & NOAA Response

_DAARWG/SAB Comments applicable to Recommendations 8__:_

- _Action 3.3: Maintain a comprehensive NOAA data inventory, including standard metadata, that is published to NOAA and Federal data catalogs (data.gov) to ensure that NOAA's data are discoverable and accessible to external stakeholders._
- _This "comprehensive NOAA data inventory, including standard metadata" seems to already exist, in the form of the NOAA Data Catalog established in ~2016 at (https://data.noaa.gov/dataset) and designed to be harvested by data.gov._

- **DAARWG/SAB Recommendation 8**: The Plan should avoid duplicating past actions or reinventing past methods unless there is an identified shortcoming with the prior approach.
  - **NOAA Response 8**: The NOAA Data Catalog is the harvesting site for NOAA data to data.gov. However, due to several issues and complications with the harvesting, a task team is being stood up by the CDO and DGC to address what the next steps should be.