



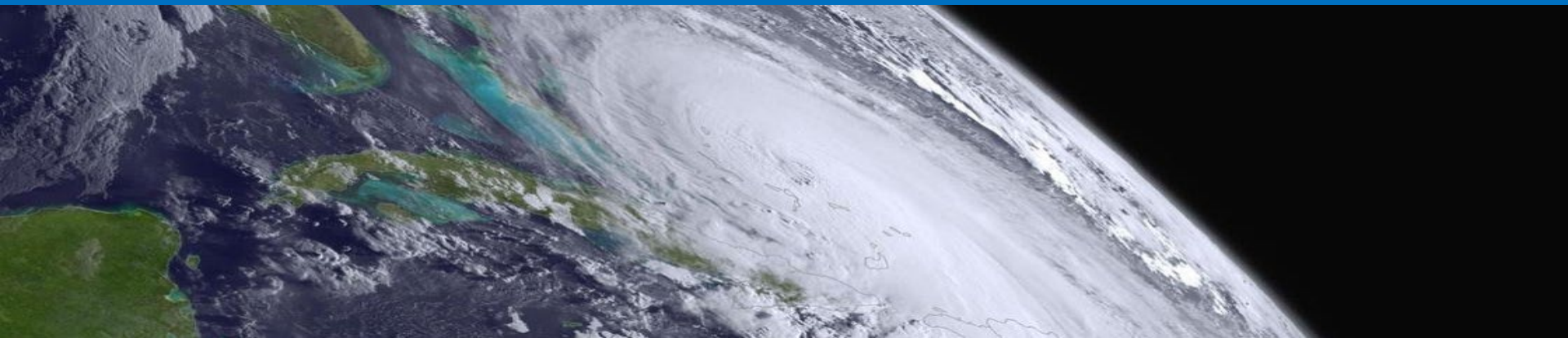
**NOAA**

Jason Cooper  
CDO Senior Policy Advisor  
(LANTERN Detailee)

November 17, 2023

# Open Data and Open Science at NOAA

Response to the December 2022 Science Advisory Board report



# Open Data / Open Science Report

## Findings:

- NOAA is succeeding in making data FAIR through the NOAA Data Strategy and the accompanying Action Plan
- NOAA is succeeding in making publications open and available through the Institutional Repository
- NOAA should embrace Open Science agency-wide

**NOAA Data Strategy**  
Maximizing the Value of NOAA Data

**NOAA**  
Institutional  
Repository

# SAB Recommendations

**Recommendation 1:** NOAA should follow the principles of FAIR open data and, whenever possible, these principles should be prioritized over other mission requirements.

**Recommendation 2:** Encourage and support the use of open source software as a key component of open science.

**Recommendation 3:** NOAA should engage with the open science community around open reproducible research and support workforce training on how to do open, collaborative, and reproducible science in support of the NOAA mission.

**Recommendation 4:** Consider providing consistent guidance across the agency for best practices, checklists, and dashboards to track adherence to open science principles, policies, and mandates across the enterprise, while still supporting NOAA's distributed culture of data and science.

# Background and Approach

- The SAB challenged NOAA to:
  - implement Findable, Accessible, Interoperable, and Reusable (FAIR) data principles,
  - promote the use of open source software,
  - engage with the open source community, and
  - apply open science principles that are consistent across NOAA.
- Responsibility assigned to the NOAA CDO to author the NOAA response, who supported a LANTERN detailee to lead and coordinate a NOAA-wide response
- Several NOAA initiatives are underway that pertain to the SAB recommendations



## NOAA SCIENCE ADVISORY BOARD REPORT ON OPEN DATA/OPEN SCIENCE

PRESENTED TO THE NOAA SCIENCE ADVISORY BOARD  
DECEMBER 1, 2022



# NOAA Data Governance Committee

NOAA has expanded data roles, responsibilities and relationships within line offices and at the enterprise level

## Relationship to this Report

- Open Data Implementation
- Cross-NOAA Coordination
- DOI Working Group
- Open Source Software
- Open Publications



# NAO 212-15B and the Data Management Handbook

Now in the final stages of review and approval, **Administrative Order 212-15B** "*Management of NOAA Data and Information*" will establish the NOAA Data Management Policy. The purpose of the policy is to ensure data are treated as a strategic asset and managed to realize the maximum value from NOAA's investment in observations, modeling, and research, per the NOAA Data Strategy.

The accompanying **Data Management Directives Handbook**

- Released internally in June 2023; will be made public once the NAO is approved.
- Provides guidance and details on how the policy is to be implemented; supercedes the NOAA Data Management Procedural Directives.
- Defines requirements, supports the tracking of performance objectives to understand the current state of data management and progress towards our data management goals.
- Audience: ACDOs, data managers, program managers, project leads, data stewards, archivists, data analysts, etc.

# Commercial Data and Licensing

NOAA **Commercial Data Buys Task Team** has been chartered to provide guidance to NOAA programs on the acquisition and use of commercial data.

**Commitment:** In its negotiations with commercial data providers, NOAA will:

- Seek to minimize access restrictions for environmental data and information acquired from non-Federal sources
- Consider contractually requiring providers transfer data ownership to NOAA or apply a Creative Commons CC0 or CC BY 4.0 license.

NOAA very much appreciates the approach suggested in the Report, “to negotiate licenses so that after a suitable embargo period, data is made FAIR.”

- A recent NESDIS data buy included a delayed release option of only 24 hours after data collection.
- NOAA has seen that companies are willing to sacrifice external commercial interests if they are compensated for those losses.
- In instances in which commercial interests, laws, policies, or international commitments prohibit openness to this data, NOAA will carefully weigh the sacrificing of open data principles against the value of the data to the NOAA mission.

# The Digital Object Identifier (DOI) Conundrum

The DGC 2022 **Digital Object Identifier DOI Task Team** Report: *"A Proposal for a Comprehensive Data DOI solution for NOAA Hosted Data,"* identified gaps in DOI issuance needs:

- DOIs are not available on all of NOAA's data holdings.
- Data DOIs are not available in a timely manner to meet journal and scientific workflow schedules.
- Data DOIs are generally not available on NOAA data hosted outside of NCEI.
- DOIs are not automatically issued for all datasets hosted at NCEI.
- Lack of easily accessible information on the data DOI request and minting process.
- Lack of a consistent DOI usage reporting infrastructure.

NOAA must maintain controls to ensure DOIs are not issued for poorly-documented, inaccessible, incomplete, or transient data.

**Commitment:** NOAA will invest in a license to issue more DOIs and in more automated assessment and approval tools, and begin to close the gaps identified in the task team report.





# Open Source Software: NAO 201-118

**NOAA Administrative Order 201-118** *"Software Governance and Public Release Policy"* will outline NOAA's policy regarding open source software and public release. Goals of the NAO are:

- All appropriate mission-oriented software developed by NOAA Federal employees or with Federal funds and products produced therefrom can be released to the public, following consistent protocols across NOAA.
- Community development of software where appropriate, in the interests of transparency, quality, and efficiency.
- Quality of software, both regarding finalized products and as a part of the ongoing development process.
- Compliance with law and best practices for software development and public release, in particular with respect to Intellectual Property issues, while retroactively addressing all such issues for existing software as needed, and addressing such issues up-front for newly developed software.
- Educate NOAA personnel with respect to requirements, policies, and best practices for software governance and public release.

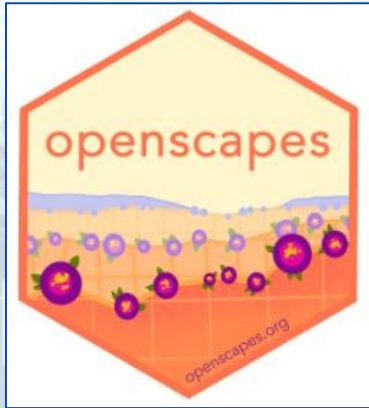
# Open Source Software: NAO 201-118

**NOAA Administrative Order 201-118** "Software Governance and Public Release Policy" addresses two recommendations directly:

- Any new projects should agree at the project start to use a widely accepted permissive open source software license and terms requiring this should be included in NOAA contracts and partnership agreements.
  - Establishes open software requirements for software developed by NOAA, software developed with NOAA funding, and jointly developed software
- NOAA-developed open source software will require maintenance over time. NOAA should develop agency wide guidance on how this will be supported if the project that develops the software is ended while other NOAA projects that rely on the software continue.
  - Requires software reporting and establishes a retirement process

# Open Science: Early Adopters

NOAA Fisheries:  
Openscapes



NWS: Earth Prediction  
Innovation Center



# Open Science: Leveraging Early Adopters

- The OSTP Year of Open Science initiative has raised awareness of open science principles.
- Communicating and demonstrating benefits is key to accelerating the shift to open science in NOAA research communities.
- NOAA can use lessons from early adopters and the open science community at large to develop guidance for implementing open science principles agency-wide.



# Summary

- NOAA is succeeding in making data FAIR through the NOAA Open Data Strategy and the accompanying Action Plan. NAO 212-15B and the Data Management Directives Handbook will further NOAA's commitment to open data.
- NOAA is implementing open licensing for its data, and negotiating open licenses for commercial data.
- NOAA must move forward to provide PIDs or DOIs for all NOAA data.
- NOAA is moving forward into an era in which open source software is an expectation. NAO 201-118 "*Software Governance and Public Release Policy*" will outline NOAA's policy regarding open source software and public release.
- NOAA must build upon open science successes and extend open science into all corners of the agency.



# BACKUP SLIDES



# SAB Recommendation 1

#	Recommendation	NOAA Response
1	NOAA should follow the principles of FAIR open data and, whenever possible, these principles should be prioritized over other mission requirements.	NOAA is continuously working to improve adherence to open data and the FAIR principles, and has made significant progress embedding them into our mission. However, directing resources toward meeting open data principles cannot occur at the expense of other mission requirements.
1.1	Specifically, since FAIR is open to interpretation, NOAA data should all be required to have PIDs, metadata, open-access APIs, and a standard open license (eg. cc-0 or cc-by).	The NOAA Data Governance Committee is advancing best practices and standards with respect to metadata, APIs, and licensing. NOAA will invest in a license to issue more DOIs and in more automated assessment and approval tools, and begin to close the gaps identified in a recent DOI task team report.
1.2	Issuing PIDs for NOAA datasets should be made the highest priority and bottlenecks removed.	To comply with new PARR requirements and this recommendation, NOAA will need to invest in a license to issue more DOIs, as well as human resources to manage the process, and in more automated assessment and approval tools.
1.3	NOAA should consider the impact of any use restrictions on purchased commercial data on reproducibility and scientific impact and strive to minimize the use of non-open data whenever practical, as well as negotiating contracts that transition this data to open data after an appropriate time period.	In its negotiations with commercial data providers, NOAA will seek to minimize access restrictions for environmental data and information acquired from non-Federal sources, and will consider contractually requiring providers transfer data ownership to NOAA or apply a Creative Commons CC0 or CC BY 4.0 license.

# SAB Recommendation 2

#	Recommendation	NOAA Response
2	Encourage and support the use of open source software as a key component of open science.	NOAA has drafted NOAA Administrative Order (NAO) 201-118, “ <i>Software Governance and Public Release Policy</i> ” to outline NOAA’s policy regarding open source software and public release.
2.1	NOAA should develop agency wide guidance recommending the use of permissive open source software licenses for most projects, unless there are compelling reasons otherwise.	NAO 201-118 seeks to ensure all appropriate mission-oriented software developed by NOAA Federal employees or with Federal funds and products produced therefrom can be released to the public, following consistent protocols across NOAA.
2.2	Any new projects should agree at the project start to use a widely accepted permissive open source software license and terms requiring this should be included in NOAA contracts and partnership agreements.	NAO 201-118 states that mission-oriented software developed by NOAA Federal employees or with Federal funds and products produced therefrom should be released to the public as open source software.
2.3	NOAA-developed open source software will require maintenance over time and NOAA should develop agency wide guidance on how this will be supported if the project that develops the software is ended, while other NOAA projects that rely on the software continue.	Software retirement is an important topic covered by NAO 201-118. The processes for software retirement will also be documented in the Handbook, including procedures for announcing, and developing a mechanism for publicizing the end of software development and maintenance. NOAA staff will be required to report their software to the Technology Partnerships Office.



# SAB Recommendation 3

#	Recommendation	NOAA Response
3	NOAA should engage with the open science community around open reproducible research and support workforce training on how to do open, collaborative, and reproducible science in support of the NOAA mission.	Openscapes and EPIC are shining stars of NOAA with respect to open science. NOAA must build upon these early successes and extend open science into all corners of the agency.
3.1	When NOAA scientists publish scientific papers, the software and configurations used for figures, tables, and core results should be made available at time of publication.	To meet this recommendation, NOAA needs to develop comprehensive guidance for scientists seeking to preserve and provide access to their software.
3.2	NOAA should invest in workforce development in broad support of open source software, make research results reproducible at the time of publication, and more generally open science.	NOAA Fisheries is leading the way in workforce development. Scientists at Fisheries have established open science and open data communities of practice that are focusing on workforce development.
3.3	NOAA should sponsor or leverage an annual conference or other annual event, such as a session at a larger scientific conference, with an accompanying report to engage consistently with the external scientific community around open data, reproducible research, and more generally open science.	The 2022 NOAA Environmental Data Management Workshop featured a session titled “Open Science and Open Science Infrastructure.” In 2023, NOAA has participated in several conferences to engage with the external scientific community. NOAA will include open data and open science principles at the annual NOAA Enterprise Data Management Workshop. The agency will consider partnering with USGS and encourage that presentations given there also be presented at extramural conferences.



# SAB Recommendation 4

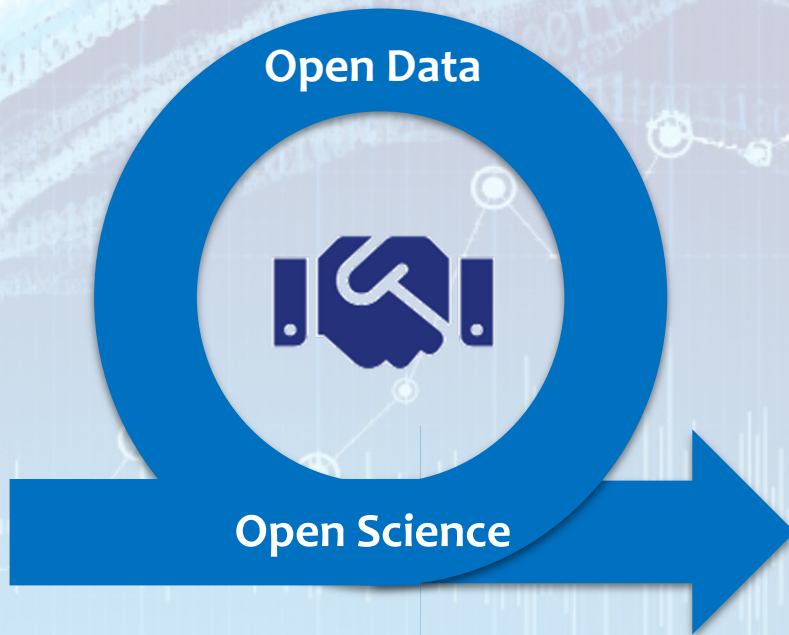
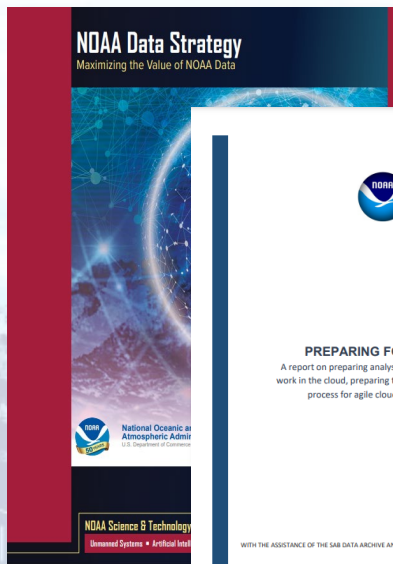
#	Recommendation	NOAA Response
4	Consider providing consistent guidance across the agency for best practices, checklists, and dashboards to track adherence to open science principles, policies and mandates across the enterprise, while still supporting NOAA's distributed culture of data and science.	Best-practices are emerging from our early adopters. The DGC is assisting with several aspects of open science. A substantial coordinated effort and resources will be required going forward to fully realize open science in NOAA. There are options we are considering as an organization to promote the importance of open science to our mission and partners. In the short term, NOAA will explore collecting best practices from early adopters and the wider research community.

# Sources for Response

- Draft NAO 212-15B and the Data Management Directives Handbook
- Draft NAO 201-118 and upcoming Handbook
- DOI Task Team Report
- NOAA Enterprise Metadata Working Group
- Data Dissemination Report
- PARR Working Group
- Commercial Data Buys Task Team
- OpenScapes, EPIC, workshops

# DAARWG Engagement

# & New Opportunities



# 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**

