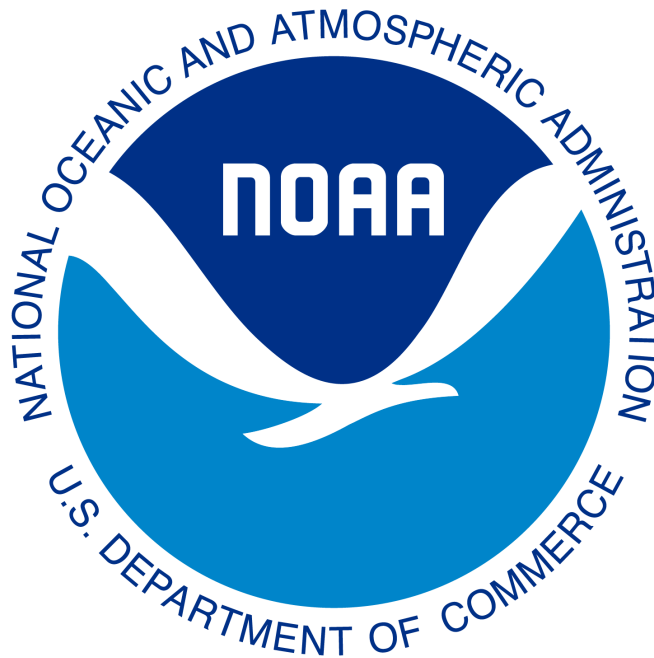


# **NOAA Response**

to the Science Advisory Board Report on

# **Open Data and Open Science**



November 2023

# Executive Summary

In its December 2022 Report on Open Data/Open Science (Report),<sup>1</sup> the Science Advisory Board (SAB, or the Board) 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. These approaches are increasingly being mandated by Federal law and expected by NOAA user communities. Adherence to open data principles maximizes return on the data collection and stewardship investments made by the American public. Open science enhances NOAA's commitment to high quality science and transparency. In this response, NOAA seeks to highlight areas of success and commit to strengthen the practices required of open data and the culture of open science across the agency.

NOAA is proud of its achievements making its data open. The NOAA Data Strategy,<sup>2</sup> along with the NOAA Data Strategic Action Plan,<sup>3</sup> are public commitments to make NOAA data open to the fullest extent possible. These build on decades of commitment to full and open data sharing, including NOAA's engagement with the World Meteorological Organization and the World Data System and role as a founder of the Group on Earth Observations. The NOAA Cloud Strategy<sup>4</sup> and NOAA Cloud Strategic Action Plan<sup>5</sup> are public commitments to providing the infrastructure to store, access, and interact with NOAA data in an open manner. The Board challenged NOAA to build out some of the underpinnings of FAIR data, including Persistent identifiers (PIDs), metadata, open-access application programming interfaces (APIs), and open licensing. We are pleased that the Board has recognized our progress, and will see our continued commitment in this response. NOAA largely concurs with SAB Recommendation 1, concerning open data, the only issue being with the Board's recommendation to prioritize FAIR principles over other mission requirements.

As realized by the Board, NOAA's adoption of open science is proceeding more slowly, and largely internally. NOAA is moving forward in the areas of open-source software and licensing. This response highlights localized progress toward meeting the Board's recommendations for training, engagement, and workforce development across the agency, and acknowledges they must be built upon. This report covers progress and challenges of open science adoption in NOAA. NOAA believes that open science drives innovation. However, based on real and perceived challenges of open science, there is hesitation from science teams and the NOAA Science Council surrounding full-scale adoption. NOAA must learn from its early adopters, and develop an internal communications strategy to explain to our scientists what open science is and why it benefits them, the agency, and the public. It will take time, but this buy-in will help NOAA meet Recommendations 2-4.

---

<sup>1</sup> [https://sab.noaa.gov/wp-content/uploads/SAB\\_Report\\_Dec2022\\_OSOD.pdf](https://sab.noaa.gov/wp-content/uploads/SAB_Report_Dec2022_OSOD.pdf)

<sup>2</sup> <https://sciencecouncil.noaa.gov/wp-content/uploads/2022/08/2020-Data-Strategy.pdf>

<sup>3</sup> <https://www.noaa.gov/sites/default/files/2022-11/NOAA-Data-Strategic-Action-Plan.pdf>

<sup>4</sup> <https://sciencecouncil.noaa.gov/wp-content/uploads/2022/08/2020-Cloud-Strategy.pdf>

<sup>5</sup> [https://sab.noaa.gov/wp-content/uploads/2022/04/SAB\\_MtgPres\\_April2022\\_NOAARespCloud\\_4-22-22\\_Final.pdf](https://sab.noaa.gov/wp-content/uploads/2022/04/SAB_MtgPres_April2022_NOAARespCloud_4-22-22_Final.pdf)

## Background

This Report comes at an opportune moment. As NOAA migrates more and more systems to the cloud, the agency has the opportunity to fundamentally change the way it conducts science and manages the data and products it creates. Data collection, product development and software procurement and development occur in vastly different environments, from NESDIS enterprise algorithms producing consistent satellite data products spanning multiple satellite deployments, to NWS working in an operational forecasting environment that processes vast amounts of weather data in real time, to Fisheries and OAR labs on the cutting edge of a very specific problem space. A more data-centric or data-first management enterprise and interoperable research environments can open the door to expanded collaboration and data sharing opportunities. NOAA is responding to these opportunities with initiatives brought into focus by the FAIR principles, the Evidence Act,<sup>6</sup> and the Office of Science and Technology Policy (OSTP) Year of Open Science.<sup>7</sup>

As the Report states, NOAA's distributed nature makes consistent NOAA-wide implementation of any policies or guidance a challenging task. However, the recent hiring of a NOAA Chief Data Officer (CDO) and Line Office Assistant Chief Data Officers (ACDOs), and expanding enterprise data governance through the new newly-chartered NOAA Data Governance Committee (DGC), makes it more achievable than ever before. CDO staff and DGC working groups can develop consistent policy and guidance, and implement dashboards to track compliance across line offices. Shining a light on parts of NOAA with strong open data and open science implementation provides models for those who are slower to embrace these important movements in the scientific community. Continued progress will advance NOAA's commitment to increasing the value of NOAA data, realizing a greater return on the investment of data collection, research and product development.

NOAA notes a correlation between this Report and a previous SAB report. In December 2019, the Data Archive and Access Requirements Working Group (DAARWG) delivered a report<sup>8</sup> containing several recommendations that will advance open science and open science in a cloud environment. NOAA has made strides implementing licensing guidance and promoting open-source software, which are described in this response.

This response is divided into four sections, mirroring the recommendations in the SAB Report on Open Data and Open Science: 1) open data, 2) open source software, 3) open science and 4) guidance. Each section contains a narrative intended to frame the response, followed by the recommendation and NOAA response.

---

<sup>6</sup> <https://www.congress.gov/115/plaws/publ435/PLAW-115publ435.pdf>

<sup>7</sup> <https://www.whitehouse.gov/ostp/news-updates/2023/01/11/fact-sheet-biden-harris-administration-announces-new-actions-to-advance-open-and-equitable-research/>

<sup>8</sup> [https://sab.noaa.gov/wp-content/uploads/2021/11/SAB\\_Mtg\\_Dec19\\_Report\\_DAARWGReport\\_final\\_121319.pdf](https://sab.noaa.gov/wp-content/uploads/2021/11/SAB_Mtg_Dec19_Report_DAARWGReport_final_121319.pdf)

# 1. Open Data

In 2019, Congress established the Evidence Act. Title II, the OPEN Government Data Act, requires agencies to create Open Data Plans to make Federal data publicly available by default, provide comprehensive searchable data inventories of all agency data assets, and designate a Chief Data Officer for the coordination of Title II activities and policies. This law is foundational to directing NOAA's Open Data implementation in the past several years. NOAA recognizes the complementary nature of the Board's recommendations to the Evidence Act.

In its response to OSTP's 2013 Public Access to Research Results (PARR),<sup>9</sup> NOAA issued a series of Procedural Directives to guide NOAA scientists and data managers, including the Data Access Procedural Directive.<sup>10</sup> It requires data to be discoverable, accessible, and include appropriate metadata. While this and the other Procedural Directives issued in the wake of the original PARR memo pre-date the FAIR principles, the spirit of those principles is present. In February 2023, the NOAA Chief Scientist and Chief Data Officer submitted an implementation plan to OSTP to address updated guidance provided in the 2022 Nelson Memo.<sup>11</sup> As part of that implementation, NOAA has established a cross-Line Office PARR working group, which meets monthly and is responsible for developing an updated NOAA PARR Plan and any associated policy or policies. Working Group deliverables are tied to the timeline provided in the Memo.

In addition, a DGC working group has updated all existing procedural directives, seen as a very positive outcome of the response to the initial PARR guidance, and included the content in an all-encompassing **Data Management Directives Handbook**. The Handbook provides consistent data management guidance, with references to open science and the FAIR data principles:

Findable	Accessible
<p>F1. (Meta)data are assigned a globally unique and persistent identifier</p> <p>F2. Data are described with rich metadata (defined by R1 below)</p> <p>F3. Metadata clearly and explicitly include the identifier of the data they describe</p> <p>F4. (Meta)data are registered or indexed in a searchable resource</p>	<p>A1. (Meta)data are retrievable by their identifier using a standardized communications protocol</p> <p>A1.1 The protocol is open, free, and universally implementable</p> <p>A1.2 The protocol allows for an authentication and authorization procedure, where necessary</p> <p>A2. Metadata are accessible, even when the</p>

<sup>9</sup> <https://obamawhitehouse.archives.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research>

<sup>10</sup> <https://nosc.noaa.gov/EDMC/PD.DA.php>

<sup>11</sup> <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>

	data are no longer available
<p style="text-align: center;"><b>Interoperable</b></p> <p>I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.</p> <p>I2. (Meta)data use vocabularies that follow FAIR principles</p> <p>I3. (Meta)data include qualified references to other (meta)data</p>	<p style="text-align: center;"><b>Reusable</b></p> <p>R1. (Meta)data are richly described with a plurality of accurate and relevant attributes</p> <p>R1.1. (Meta)data are released with a clear and accessible data usage license</p> <p>R1.2. (Meta)data are associated with detailed provenance</p> <p>R1.3. (Meta)data meet domain-relevant community standards</p>

**Figure 1. FAIR Guiding Principles for scientific data management and stewardship, Wilkinson et al. (2016).<sup>12</sup>**

The NOAA Executive Panel tasked the NOAA Chief Data Officer to convene a cross-agency team to develop a current and future state high-level analysis of NOAA’s data dissemination activities. In September, NOAA finalized a Data Dissemination Report, with a stated objective to “Identify dissemination platforms that are findable, accessible, interoperable, and allow for the reuse of digital assets that offer interoperability (FAIR principles) for data delivery, including compliance with the Evidence Act and open science policies.” It also developed a NOAA-internal dissemination catalog and assessment tool, which accommodates all NOAA systems, with an initial emphasis on flagship systems that disseminate NOAA’s most valuable data. The end goal of these integrated activities is to ensure NOAA’s dissemination systems are in alignment with all Federal and NOAA policy requirements, support open access to all of NOAA’s data as applicable, and are responsive to user requirements of high-quality NOAA data and services.

Observational data from commercial sources occupies a unique space in the era of open data and open science. Restrictions on data can impact release of derived products, and can inhibit our ability to enable reproducibility of scientific results. NOAA must negotiate with

<sup>12</sup> <https://www.nature.com/articles/sdata201618>

the commercial sector for the right to share commercial data as openly as possible. There are sometimes creative ways to address that interest without unduly burdening the commercial provider's broader business opportunities, such as providing data openly after a time delay or at a coarser resolution. This is especially important as NOAA increasingly seeks to augment its data holdings using commercial data. NOAA serves a vital conduit of environmental data. There is a vibrant downstream commercial sector that is built off of free access to government-funded data that is also important to the US economy and NOAA's mission.

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

**NOAA Statement:** NOAA partially concurs with this recommendation. NOAA appreciates the recommendation to prioritize openness of data, and as the Board acknowledged, has made great strides in that direction. However, it is difficult to consider risking other mission requirements (for example, generating fisheries stock assessments, model output data for forecasting, or producing the National Climate Assessment) in favor of fast-tracking FAIR data principles. 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.

### **What we are doing now:**

- Currently 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 Data Management Directives Handbook assumes the full authority of NAO 212-15B to provide guidance and details on how the policy is to be implemented superseding the NOAA Data Management Procedural Directives. The Handbook defines requirements, supports the tracking of performance objectives to understand the current state of data management and progress towards our data management goals. The intent of the handbook is to outline procedures that Assistant Chief Data Officers (ACDOs), data managers, program managers, project leads, data stewards, archivists, data analysts, or any other roles involved in data life cycle management must take to meet the purpose of the NAO 212-15B. The Handbook was released internally in July 2023, and will be made public once the NAO is approved.

## What we plan to do in the future:

- The NOAA CDO will promulgate the principles through the new ACDOs. The CDO will continue to track execution of the NOAA Data Strategy through the Action Plan. NOAA will develop metrics and report on compliance with guidance in the Handbook designed to further adherence to open data principles.
- NOAA appreciates the subcommittee finding that “FAIR principles can be challenging to implement for legacy datasets and products due to the lack of resources and expertise.” Legacy data would be more usable, and thus more used, if FAIR principles were applied. The CDO will encourage the National Centers for Environmental Information (NCEI) to evaluate the implementation of the principles to archived datasets, and submit a plan for implementation. NCEI would not be starting from scratch: in 2019, NOAA data management experts published the Data Stewardship Maturity Matrix,<sup>13</sup> and could apply or adapt it as needed to focus more squarely on FAIR data principles.
- NOAA also notes the difficulty “linking” related datasets. For example, if a researcher discovers a new species, it is difficult to find the associated conductivity, temperature, and depth (CTD) data needed to identify areas conducive for species habitat. NOAA will investigate parent, child, and sibling PID structures to link related datasets.

**SAB Recommendation 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 (e.g., cc-0 or cc-by).

**NOAA Statement:** NOAA concurs with this recommendation.

## What we are doing now:

- **PIDs.** NOAA issues Digital Object Identifiers (DOIs) through Datacite. The issuing of DOIs is tightly controlled by NCEI and a few other NOAA offices for two reasons. First is the number of staff that would be required to mint, manage, and track the additional licenses. Second is to ensure that DOIs are issued for well-documented, stable, accessible data collections. To address this issue, the DGC in 2022 convened a DOI task team, which delivered a report, “*A Proposal for a Comprehensive Data DOI solution for NOAA Hosted Data*,” to address gaps in DOI issuance needs. Identified gaps include:
  - i. DOIs are not available on all NOAA’s data holdings.
  - ii. Data DOIs are not available in a timely manner to meet journal and scientific workflow schedules.
  - iii. Data DOIs are generally not available on NOAA data hosted outside of NCEI.
  - iv. DOIs are not automatically issued for all datasets hosted at NCEI.
  - v. Lack of easily accessible information on the data DOI request and mint process.

---

<sup>13</sup> <https://datascience.codata.org/articles/10.5334/dsj-2019-041>



vi. Lack of a consistent DOI usage reporting infrastructure.

- **Metadata.** The NOAA Enterprise Metadata Working Group and several line office working groups exist to develop templates, guidance, etc. which are used to expand metadata availability, consistency, and completeness across the agency.
- **Open-access APIs.** NOAA is developing an inventory of data dissemination points across the agency, as well as a catalog tool, described above, which will allow for assessment of dissemination system alignment with Federal and NOAA policy requirements, and support open access to all of NOAA's data as applicable.
- **Standard open data license.** NOAA has provided data licensing guidance to its line and staff offices: "All Internal NOAA Source Data that are appropriate for public release should be formally dedicated to the public domain via the Creative Commons 1.0 Universal Public Domain Dedication (CC0-1.0), which removes all copyright that may exist from the data so that it may be used by anyone, for any purpose." Implementation guidance for metadata fields and landing pages is also included. The NOAA Office of Coast Survey has been a trailblazer on this topic, releasing internal, commercial, and voluntarily provided bathymetry data under CC licenses.<sup>14</sup>

NCEI has issued guidance for its staff to follow when working with data providers:

- a. Environmental data and information **produced by NOAA or any Federal agency** are available fully and openly to data users unless explicitly exempt from release by Federal law, regulation or policy. These data are in the public domain in the United States. To serve our international partners and customers, and to remove uncertainty in the United States, NCEI will work to ensure a Creative Commons Zero (CC0-1.0) license is applied.
- b. Environmental data and information **collected or created under NOAA grants or cooperative agreements** must be made visible, accessible, and independently understandable to general users in a timely manner (typically no later than two (2) years after the data are collected or created [*NOTE: this "embargo" period will be eliminated as NOAA responds to new PARR guidance*]). Except where limited by law, regulation, policy or security requirements, NCEI will strongly encourage providers to apply a CC0 1.0 or CC BY 4.0 license in grant or agreement language, or upon submission.
- c. NCEI may archive data **produced by non-Federal entities** for which the creator either surrenders or maintains intellectual rights. NCEI will seek to minimize access restrictions for environmental data and information acquired from non-Federal sources, and will encourage providers to apply a CC0 1.0 or CC BY 4.0 license. If it is determined that accepting limited-access data is not in the best interest of the public, NCEI may decline to accept data.

---

<sup>14</sup> <https://www.nauticalcharts.noaa.gov/data/data-licensing.html>



### **What we plan to do in the future:**

- **PIDs.** The expansion of DOI issuance will result in higher cost and development of some automated controls to ensure DOIs are not issued for poorly-documented, inaccessible, incomplete, or transient data. The updated PARR guidance requires this expansion. 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 DOI task team report.
- **Metadata.** A DOI link resolves to a landing page, which includes metadata about the collection. NOAA believes that the expansion of DOI issuance will necessarily improve collection-level metadata to include discovery, usage, format, lineage, etc.
- **Open-access APIs.** NOAA will assess the openness of these access points to ensure that none require a user authentication or account unless the data are protected by legal mandate. NOAA will assess the feasibility of requiring all data to be accessible via API, and develop guidance on providing access to data not currently accessible via API.
- **Standard open data license.** NOAA will extend licensing guidance to NOAA-funded and NOAA-acquired data, including commercial data (See 1.3).

### **SAB Recommendation 1.2: Issuing PIDs for NOAA datasets should be made the highest priority and bottlenecks removed.**

**NOAA Statement:** NOAA concurs with this recommendation.

**What we are doing now:** NOAA issues DOIs for all NOAA publications, and for datasets upon request.

**What we plan to do in the future:** As stated in 1.1, the expansion of DOI issuance will result in nominally higher licensing costs. More significantly, it will necessitate development of automated controls to ensure DOIs are not issued for poorly-documented, inaccessible, incomplete, or transient data. The updated PARR guidance requires this expansion. 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. In cases where a DOI is not issued, NOAA will require data collections to be assigned an alternative PID. NOAA will investigate whether collection metadata IDs and universally unique identifiers (UUIDs) could be part of the solution. These IDs are already in use in parts of NOAA. Leveraging them could lead to accelerated adoption of this recommendation.

**SAB Recommendation 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.

**NOAA Statement:** NOAA concurs with this recommendation.

**What we are doing now:** The **NOAA Commercial Data Buys Task Team** is a limited-duration team of topical experts chartered to provide guidance to NOAA programs on the acquisition and use of commercial data, as defined by Federal Acquisition Regulation and the US Group on Earth Observations. This guidance will not only enable NOAA operational and research mission needs, but also meet statutory requirements and policies in a cost-effective manner. This will lead to the “establishment and maintenance of a sustainable and mutually beneficial partnership with the commercial sector with particular attention to applicable data rights in contractual arrangements.” NOAA has significant experience with space weather and radio occultation data purchases, commercially procured bathymetry data, and privately-owned uncrewed systems. The task team will collect and share insights gained from recent data buys and make recommendations for maximizing the value of commercial data to research communities and citizens.

**What we plan to do in the future:** 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 that 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 radio occultation 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.

## 2. Open Source Software

The move to the cloud and the emergence of open science tools affords the opportunity to rethink software development. To make software development more open and sustainable, NOAA can standardize and containerize code so that packages can be reused across applications, programs, and even line offices. NOAA can encourage this software to be open source and carry an open license, so that other NOAA staff, as well as our academic and other partners, can improve the software and ideally use it to reproduce research results.

Commercial data and software has a place in NOAA, particularly for mission-support functions. NOAA will use commercial off-the-shelf software for cases which meet NOAA's needs while remaining aware of reproducibility issues. Similarly to data, NOAA can negotiate the rights to share software. Opportunities exist to broadly share data/software through open licensing. To meet the recommendations in this Report, NOAA must train developers to identify and use open source software development tools.

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

**NOAA Statement:** NOAA concurs with this recommendation. Ensuring that NOAA software (with some exceptions due to security concerns, etc.) is available through open source licenses builds transparency and trust in NOAA science and services, promotes private sector innovation, and through open science, enables broader contributions to NOAA systems that serve NOAA's mission.

**What we are doing now:** For software developed by NOAA or with NOAA funding, including software used for scientific research and operational products, NOAA has drafted **NOAA Administrative Order (NAO) 201-118**, "*Software Governance and Public Release Policy* to outline NOAA's policy regarding open source software and public release.

**What we plan to do in the future:** Coinciding with the approval of the Software Governance NAO, NOAA will draft a Handbook to implement the policy. The Handbook will provide NOAA with requirements, guidelines, and processes to reach compliance with the NAO. The Handbook will outline what software is required to be released to the public, standards for release including documentation and quality assurance, guidance on open source licenses including public domain, release and disclosure procedures, language for grants and contracts to ensure open source software when development starts, and pathways for current software to become compliant.

## **SAB Recommendation 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.**

**NOAA Statement:** NOAA concurs with the recommendation, and has stated compelling reasons in 1.1.

**What we are doing now:** NAO 201-118 is in the final stages of review, with the following purpose:

1. 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. To enable community development of software where appropriate, in the interests of transparency, quality, and efficiency.

3. To ensure quality of software, both regarding finalized products and as a part of the ongoing development process.
4. To ensure compliance with law and best practices for software development and public release, in particular with respect to Intellectual Property issues, while:
  - a. retroactively addressing all such issues for existing software as needed; and
  - b. addressing such issues up-front for newly developed software.
5. To educate NOAA personnel with respect to requirements, policies and best practices for software governance and public release.

**What we plan to do in the future:** Once NAO 201-118 is approved, the accompanying Handbook will include implementation guidance to assist staff, as outlined in response to Recommendation 2. Note that the NOAA Office of General Counsel is already providing consistent guidance to NOAA Line Offices.

**SAB Recommendation 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.

**NOAA Statement:** NOAA concurs with this recommendation, with the caveat that the decision on the best software solution must be based on a broad range of factors (requirements, security, partnerships, standards, hosting, etc.), and that in some cases, exceptions to open source and public release may apply, including software related to security and commercial-off-the-shelf (COTS) software. Draft NAO 201-118 takes this same position, providing 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.

The Earth Prediction Innovation Center (EPIC) and the Unified Forecast System (UFS) are examples in which NOAA is already using open source licenses for some of our primary weather models<sup>15</sup> (For more on EPIC, see Recommendation 3). NOAA appreciates the subcommittee finding that retroactively making software open source is extremely difficult.

**What we are doing now:** Draft NAO 201-118 addresses open software requirements for software developed by NOAA or with NOAA funding, as well as jointly developed software:

- “Software developed by NOAA or with NOAA funding specifically for its mission will be developed and publicly released as open source software unless legally prohibited or superseded by formal, written agreements between NOAA and an external entity.”
- “For software jointly developed, modified or added upon in collaboration between Federal employees and any entity external to NOAA (including contractors, grantees, cooperative institutes, private entities, interagency partners, international partners, Cooperative Research and Development Agreement partners, individual contributions via Github, etc.), steps must be taken to ensure that the development results in

---

<sup>15</sup> <https://github.com/ufs-community/ufs-weather-model/blob/develop/LICENSE.md>

intellectual property rights that enable the software to be publicly released as open source software.”

- NOAA is developing step-by-step guidance on NAO 201-118 compliance for inclusion in a Handbook for staff reference. Guidance has already developed and cleared legal review of draft contract and grant terms, for inclusion in the Handbook.

**What we plan to do in the future:** Once NAO 201-118 is approved, NOAA will communicate the new policy and guidance to project managers and those responsible for contract or grant development and acquisition, in order to begin use of language requiring open source software development in the contract and grant terms.

**SAB Recommendation 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.

**NOAA Statement:** NOAA concurs with this recommendation. Retirement of legacy software is critical due to resource needs and overhead for maintenance, and also to ensure projects are using or developing the most relevant software systems to advance science or technical capabilities.

**What we are doing now:** NOAA leverages Git (Github/Gitlab) for most large software development projects, for configuration management, documentation, and support. While there is a NOAA GitHub instance,<sup>16</sup> repositories are managed at the project level. Documentation and user forums are key to communicating across projects which versions of the software are maintained and supported for use and development. When older versions of the software are no longer supported, or the software system is retired, projects are advised to migrate to new systems if they exist, or to take on maintenance responsibilities themselves if they do not exist. One example is the Hurricane Weather and Research Forecasting (HWRF) model, which was replaced by the UFS-based Hurricane Analysis and Forecast System (HAFS). The HWRF system will no longer be maintained beyond the 2023 Atlantic Hurricane season, and users and developers for hurricane modeling have been advised to transition to HAFS.

**What we plan to do in the future:** 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.

As part of development of the Handbook to accompany NAO 201-118, NOAA staff will be required to report their software to the Technology Partnerships Office. When resourcing allows, NOAA will make a publicly available list of open source software available.

---

<sup>16</sup> <https://github.com/NOAAGov>

### 3. Open Science

This highly-collaborative approach to scientific research has the potential to move scientific discovery forward more quickly, shortening the time from research project inception to actionable results. By increasing access to scientific data and the scientific process, open science is also increasing diversity in research communities. As the Report states, NOAA is not as far along the path toward full adoption of open science as with open data. In the Federal Government and the scientific community at large, the concept of open science is not as mature. NOAA recognizes benefits and challenges of open science:

<b>Benefits</b>	<b>Challenges</b>
<ul style="list-style-type: none"><li>● Greater trust in scientific results</li><li>● Reproducible results</li><li>● Greater inclusion of early career and more diverse scientists</li><li>● Collaboration with colleagues inside and outside of NOAA</li><li>● More scientists have access to the tools needed to do open science</li><li>● Common tools: Jupyter Notebooks, processing capability, open source software</li><li>● Onboarding new people into science teams</li><li>● Litigation protection</li><li>● Easier to continue research beyond one scientist's career</li></ul>	<ul style="list-style-type: none"><li>● Potential for decisions to be made with pre-published data</li><li>● Slowing of the publication process</li><li>● Favoring easily reproducible research over...<ul style="list-style-type: none"><li>○ Research that includes modeled outputs that can vary each time the model is run</li><li>○ Exploratory research which no clear hypothesis going in</li></ul></li><li>● Research involving many large or disparate datasets (esp. interdisciplinary)</li><li>● Professional embarrassment</li><li>● Scientists dependent on commercial software will have their careers stunted<ul style="list-style-type: none"><li>○ Critical research results will be devalued</li></ul></li><li>● Getting "scooped!"</li></ul>

**Figure 2. Benefits and challenges of open science, from a literature review and an ad hoc OAR open science panel event.**

NOAA fully understands that open science drives innovation. However, the challenges above are causing hesitation surrounding full-scale adoption. Pockets of NOAA have taken the initiative to move toward open science. The agency must learn from its early adopters, and develop an internal communications strategy to explain to our scientists what open science is and why it benefits them, the agency, and the public. NOAA's move toward open science must be deployed in a way that continues to promote innovation, and staff, particularly more tenured scientists, are provided the tools and guidance to embrace open science confidently. A sound communications strategy can mitigate anticipated resistance.

### **SAB 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.

**NOAA Statement:** NOAA concurs with this recommendation.

**What we are doing now:** NOAA has cited early engagement and training above, and wishes to highlight two current initiatives. NOAA Fisheries and NWS Unified Forecast System are the open science leaders within NOAA.

**Openscapes**<sup>17</sup> is a non-profit organization promoting a scientific culture that is more efficient and collaborative, and can uncover environmental solutions faster. NOAA Fisheries is working with Openscapes to advance open science. The Openscapes Mentors program supporting open science from 2023 through 2026. The overarching vision of NOAA Fisheries open science is to support scientists, developers, and policy analysts in fulfilling open science principles. Openscapes is concerned with open science training in workflow and technical skills needed at the individual and team level focused on helping all staff engaged in data-driven science and decision-making through an active and engaged mentor group across Fisheries. Early adopters are quickly seeing the benefits of open science in their research, and will very likely promote open science regardless of any policy or mandate to do so. Fisheries is beginning to communicate the benefits of Openscapes, and open science in general, to other line offices.

NOAA's **Earth Prediction Innovation Center (EPIC)**.<sup>18</sup> is the catalyst to drive a collaborative weather research community to enhance the Unified Forecast System (UFS). Designed to support the Weather Enterprise, the UFS is the source system to simplify NOAA's current operational suite of 21 stand-alone forecast systems, utilizing common codes and infrastructure. EPIC is moving the UFS into a public-facing cloud computing environment. The EPIC Community Portal provides community access to the UFS code base, tutorials, advanced user support, application training, codefests, workshops and funding opportunities. EPIC creates an open science environment for UFS that enables collaborations and contributions within the broad weather community to advance operational forecast skill. Key success metrics for EPIC include the number of community contributors to UFS, the speed at which new code innovations move from research to operations or product applications, and the improvement of the skill of the forecast system.

**What we plan to do in the future:** NOAA must build upon these early successes and extend open science into all corners of the agency. NOAA can develop metrics, communicate those

---

<sup>17</sup> <https://www.openscapes.org/>

<sup>18</sup> <https://epic.noaa.gov/>



metrics, and capture best practices. NOAA will leverage NASA TOPS training modules to train scientists how to do open, collaborative, and reproducible science in support of the mission.

### **SAB Recommendation 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.

**NOAA Statement:** NOAA concurs with this recommendation. NOAA is committed to meeting the new PARR guidance requiring data collected and/or used in research activities to be made public at the time of publication.

**What we are doing now:** As mentioned in Recommendation 2.3, NOAA projects use GitHub to store their software. NCEI is developing guidance on submitting software to the archive alongside the data. NOAA has some experience making software and configurations for figures and tables available. The 2018 National Climate Assessment may have included the most comprehensive effort to document data sources in figures.<sup>19</sup> The effort was extremely labor-intensive, but deemed necessary given the visibility and scrutiny of the assessment.

**What we plan to do in the future:** To meet this recommendation, NOAA needs to develop comprehensive guidance for scientists seeking to preserve and provide access to their software. The Climate Assessment effort provides a model for others to follow, which would be less onerous for less complex, comprehensive papers.

### **SAB Recommendation 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 Statement:** NOAA concurs with this recommendation. As stated in the Report, when compared with progress on open data, there has been less progress on supporting open science at NOAA through open source software. However, there are pockets of NOAA in which the use of open source software has taken root.

#### **What we are doing now:**

- In May, NOAA Fisheries presented their work with Openscapes to the NOAA Data Governance Committee. Fisheries has found success implementing open science principles by demonstrating the benefits rather than by simple mandate. Socializing this success across the agency will encourage adoption of open science principles.
- NCEI has been exploring archiving inputs and instructions for producing a dataset or product in lieu of the actual data since at least 2015. In 2017, NESDIS created a policy to “determine the cost effectiveness of on-demand regeneration of products derived from

---

<sup>19</sup> <https://data.globalchange.gov/report/nca4/figure>

Environmental Data v. long-term preservation of the derived products.”<sup>20</sup> If successful, this policy would both reduce archival storage costs and ensure reproducibility. However, issues have arisen with software maintenance, IT security issues stemming from archival of executable code, and especially in the realm of model data, in which a model simulation run thousands of times can produce thousands of slightly different outputs. NOAA will investigate the reproducibility of results for data used in published research to see if the issues are more manageable.

- Throughout FY 2023, the NOAA Library has offered outreach campaign weeks and ad-hoc webinars focusing on different aspects of open science. Examples include [Misconceptions Surrounding Open Access](#), [Supporting Open Science Through Open Access & Data Initiatives](#), and [Using ORCID for Disambiguation](#).
- Regarding workforce development, NOAA Fisheries is leading the way. Scientists at NOAA Fisheries have established open science and open data communities of practice that are focusing on workforce development. The Fisheries Information System (FIS) Program was an early adopter of Open Science practices and organizes the FIS Professional Specialty Groups to lead peer-learning and training. The Fisheries Integrated Modeling System (FIMS) initiative, another early adopter, is helping develop next generation stock assessment models by embracing Open Science, and their staff help lead Open Science trainings across NMFS. The NMFS Open Science group helps coordinate team-trainings on reproducible workflows and tools via the Openscapes program. NMFS Open Science also hosts the NMFS R User Group which runs twice-monthly training and maintains a calendar of open R trainings (2-3 per week) from across the wider Open Science community. The NOAA Fisheries Integrated Toolbox (FIT) provides infrastructure and training for open science software development. Socializing these successes across the agency will encourage adoption of open science principles.

**What we plan to do in the future:** NOAA will continue to leverage early adopters to advance open science across the agency. The agency can prioritize familiarity with and enthusiasm for open science in its vacancy announcements. NOAA can recognize open science adoption in its annual Science Report.<sup>21</sup>

Starting FY24, NMFS Office of Science and Technology designated a full-time lead for its Open Science initiative. 2-3 open science representatives have been identified at each science center to support open science outreach and training with supervisor support for 8 hrs/month. Team-based training in reproducible workflows started in 2020, will continue FY24-26, with a goal of 100 staff trained per year.

---

<sup>20</sup> [http://prod.nesdis.acsitefactory.com/sites/g/files/anmtlf151/files/2021-08/npd\\_6010\\_01a.pdf](http://prod.nesdis.acsitefactory.com/sites/g/files/anmtlf151/files/2021-08/npd_6010_01a.pdf)

<sup>21</sup> <https://sciencecouncil.noaa.gov/council-products/noaa-science-report/>

**SAB Recommendation 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.

**NOAA Statement:** NOAA concurs with this recommendation.

**What we are doing now:** The 2022 NOAA Environmental Data Management Workshop featured a session titled “Open Science and Open Science Infrastructure.”

In 2023, NOAA has participated in conferences to engage with the external scientific community:

- In April, the NOAA Weather Program Office and the Natural Hazards Center partnered to convene a Workshop on Open Data and Reuse in Social Science Weather Research.<sup>22</sup> The workshop sought to gather insights and recommendations on ethical publication and reuse of social, behavioral, and economic science data in the context of weather-related research.
- In May, several NOAA staff attended the USGS Community for Data Integration Workshop.<sup>23</sup> Attendees quickly realized the value of this mostly-internal collaboration and sharing of open science tools and methodologies.
- In July, NOAA hosted the Unifying Innovations in Forecasting Capabilities workshop<sup>24</sup> in Boulder, CO. The workshop invited participants to share challenges and successes related to Unified Forecast System (UFS) contributions, and voice their thoughts on where the future will take us. The workshop’s goal was to engage the greater Weather Enterprise in the ongoing effort to accelerate contributions to the UFS. The theme for this year’s workshop was Innovation and Community.
- Also in July, more than 40 NOAA scientists, data managers, and program managers across six line offices participated in the annual meeting of the Earth Science Information Partners (ESIP), a collaborative non-profit community working to meet environmental data challenges across Earth science disciplines. Eleven NOAA-led breakout sessions covered diverse topics including: AI-ready data, open-source analytics, cloud-native technologies, and more. The meeting theme was Opening Doors to Open Science. In partnership with NASA and USGS, NOAA is a long-standing sponsor of ESIP's mission.

---

<sup>22</sup> <https://hazards.colorado.edu/uploads/basicpage/odrw-agenda.pdf>

<sup>23</sup> <https://www.usgs.gov/centers/community-for-data-integration-cdi/2023-cdi-workshop>

<sup>24</sup> <https://epic.noaa.gov/eventsposts/uifcw-2023/>

- NOAA is an active participant on the American Meteorological Society's Science and Technology Activities Commission "Board on Open Science, Data, and Software"<sup>25</sup>

**What we plan to do in the future:** NOAA can do much more to showcase our progress and promote adoption in the areas of open data, reproducible research, and open science. 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 NOAA will encourage that presentations given there also be presented at extramural conferences and workshops such as AGU, AMS, Ocean Sciences, and the ESRI User Conference.

## 4. Open Science Guidance

NOAA agrees with the subcommittee recommendation to provide consistent guidance across the agency for best practices, checklists, and dashboards to track adherence to open science principles, policies and mandates across the enterprise. 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.

**SAB 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.

**NOAA Statement:** NOAA concurs with this recommendation.

**What we are doing now:** Best-practices are emerging from our early adopters. The DGC is assisting with several aspects of open science: open data, DOIs, and even open source software and publishing, which are on the fringes of DGC scope.

**What we plan to do in the future:** While NOAA appreciates the Board's suggestion that the Data Governance Committee lead these activities, the committee does not have the expertise to develop agency-wide guidance on open science. We acknowledge that 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. As the adoption of open science continues, NOAA will use these and general open science principles to develop a checklist for scientists to follow to make their data and scientific practices open. This can even

---

<sup>25</sup> <https://www.ametsoc.org/index.cfm/stac/boards/board-on-open-science-data-and-software/terms-of-reference/>

be included in the automated assessment of data and products submitted for archive from our distributed communities.

## Conclusion and Acknowledgments

NOAA provides essential environmental data and scientific output to researchers, decision-makers and the American public. NOAA acknowledges and appreciates the Science Advisory Board for their thoughtful recommendations for improving the openness of NOAA data and science. This response is reflective of our progress as of November 2023. NOAA is committed to advancing open data and open science as outlined in this response, and would appreciate the opportunity to provide periodic updates to the SAB on our progress.