National Oceanic and Atmospheric Administration (NOAA) Administrative Order (NAO) 212-15B:

Management of NOAA Data and Information

Data Management Handbook

Document Owner: NOAA Data Governance Committee (DGC)

This Handbook is the implementation of NAO 212-15B that establishes the requirements and procedures for full data life cycle management in all domains of NOAA data and information.

Version History

October 1, 2024: This Data Management Handbook supersedes prior Data Management Procedural Directives, and is effective October 1, 2024. The DGC will revise or reaffirm the Handbook at least every two years or as needed. The authoritative location for the Handbook is: https://sites.google.com/noaa.gov/noaa-data/handbook (internal NOAA site).

Table of Contents

Table of Figures

Chapter 1 - Purpose and Scope of the NAO for Management of NOAA's Data and Information

A. Purpose

NAO 212-15B establishes the Department of Commerce (DOC) 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**. This Data Management Handbook assumes the full authority of the NAO to provide further guidance and details on how this policy is to be implemented. The intent of this Handbook is to define requirements, track performance objectives and metrics for executives to understand the current state and progress, as well as outline procedures that Assistant Chief Data Officers (ACDOs), data managers, program managers, project leads, data stewards, archivists, and data analysts, or any other roles involved in data life cycle management must take to meet the purpose of the NAO.

B. Policy Scope and Background

This Handbook is approved, issued, and maintained by the NOAA DGC pursuant to the authorities granted in NAO 212-15B, Management of Data and Information. The Handbook will be reviewed annually and updated every two years, or more frequently as needed. This Handbook supersedes the Environmental Data Management Framework and all previous Environmental Procedural Directives as issued by the Environmental Data Management Committee (EDMC). Any references to the previous EDMC Procedural Directives should be interpreted to reference this Handbook.

Background

Data management ensures maximum use, value, and trust of data by present and future generations by establishing processes and policies that enable broad discovery, access, interoperability, and reusability. It is intrinsic to data integrity, and data integrity is a vital component to scientific integrity, a core pillar of NOAA's scientific efforts. Scientific integrity is important to everyone at NOAA, and it is our duty to present and future generations that depend on NOAA's scientific legacy. While there are laws and regulations that guide the implementation of data management, the overarching goal is to enable innovation to address environmental and societal challenges.

This Handbook is based on current applicable laws and regulations (Appendix A) related to data management as well as our Agency's and the federal data community standards and best practices. In particular, the Handbook builds off of and addresses elements of the Foundations for Evidence-Based Policymaking Act (the Evidence Act), Department of Commerce 2022-2026 Strategic Plan, NOAA's Data Strategy, and the NOAA Public Access to Research Results

(PARR) Plan. Recognizing that these strategies and plans will change over time, this Handbook is intended to be a living document that can be updated to reflect changes in rules, regulations, strategies, and plans.

Below is a brief history and description of a series of guiding policy documents that form the foundation of this document.

In 2002, the Office of Management and Budget (OMB) issued Circular A-16 to provide a coordinated approach to managing geospatial data across the federal government. OMB issued supplemental guidance in 2010 to "facilitate the adoption and implementation of a coordinated and effective federal geospatial asset management capability."

In 2013, OMB issued Open Data Policy (OMB Memorandum M-13-13) directing all agencies to "manage information as an asset throughout its life cycle to promote openness and interoperability, and properly safeguard systems and information." Similarly, the White House Office of Science and Technology Policy (OSTP) issued a Memorandum on February 22, 2013 entitled "Increasing Access to the Results of Federally Funded Research" directing each federal agency that conducts over $100 million annually in research and development expenditures to develop a plan to support increased public access to the results of that research. In response to the OSTP Memorandum, the NOAA Research Council issued the NOAA Plan for Increasing Public Access to Research Results (PARR) in February 2015.

In addition, the Foundations for Evidence-Based Policymaking Act of 2018 (also referred to as the Evidence Act) requires agencies to modernize data management practices. In particular, Title II of the law, the OPEN Government Act, requires agencies to create Open Data Plans to make federal data publicly available by default and provide comprehensive searchable data inventories of all agency data assets.

The NOAA Data Strategy, issued in July 2020, is consistent with the OMB Memorandum, M-19-18 Federal Data Strategy - A Framework for Consistency, and the Federal Data Strategy's 2020 Action Plan. It builds upon statutes and OMB information policy and guidance, with supplementary guidance on how agencies should manage and use federal data. As with the Federal Data Strategy, the NOAA Data Strategy will constantly evolve to be regularly updated for consistency with new statutory or OMB information policy guidance. The NOAA Data Strategy is consistent with the Department of Commerce Strategic Plan for 2022-2026, the NOAA Information Resources Management Strategic Plan 2021-2025, and the NOAA 2020 Business Brief.

Finally, in August of 2022 OSTP released a Memorandum, "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research", that updates the PARR requirements requiring data to be available at the time of publication.

In addition to the laws and policies, in the environmental data community there are the principles of Findable, Accessible, Interoperable, and Reusable (FAIR), Collective Benefit,

Authority to Control, Responsibility, and Ethics (CARE), and Open Data/Open Science. FAIR was defined in a 2016 article “FAIR Guiding Principles for scientific data management and stewardship,” with the intention to support the reuse of data in the scientific community (Wilkinson et al., 2016). CARE principles were created by the Global Indigenous Alliance to build on FAIR principles to ensure the rights and interests of Indigenous People are respected in data management. Open Data/Open Science is a community effort, emphasized by the White House’s 2023 Year of Open Science, to ensure data, metadata, and science are publicly available and unrestricted to advance open and equitable research.

NOAA uses NAOs to determine policy and Handbooks to provide detailed information to implement those policies. The definition of terms used to describe the intent and scope:

- Policy: the course of action determined by NOAA and codified in a NAO
- Requirements: what is needed to be in compliance with applicable laws, policies, and strategic plans
- Objectives and Metrics: how to determine and track if a requirement is being met
- Procedures: details on how to meet the requirements
- Best Practices: optional or additional actions that can be taken to improve data management and the value to the public.

Scope

This Handbook is designed to provide NOAA data practitioners the requirements, metrics, and procedures and/or best practices to follow in alignment with the above laws, regulations, memoranda, and strategic plans. The scope of this document includes procedures to manage data throughout the data life cycle from data planning, obtaining data through long-term preservation and broad data access, but does not include data security, data collection planning or execution, or data quality control. The procedures apply to all NOAA Line Offices and Staff Offices, including affiliates and grantees funded with NOAA resources. Per the NAO, all offices are mandated to follow these procedures except where it is noted as a best practice. The DGC through the ACDOs will have oversight to evaluate compliance and recognize that resources may not be initially available to meet all requirements.

Chapter 2 - Key Terms and Understanding Data Management

A. Purpose

The purpose of this Chapter is to provide an overview of data life cycle management and define key terms that will be used throughout the Handbook. If there are cases where the definitions of these terms need to vary in a particular chapter or section it will be explicitly stated. In the absence of that clarification the terms in the chapter apply to all other sections of the Handbook.

B. Data Life cycle Management

The Data Life Cycle Model

A data life cycle model provides a common conceptual framework to define roles as well as actions, operations, or processes required to manage data as a strategic asset, meet the law and policy drivers described in the background, and achieve the principles of FAIR, CARE, and Open Data/Open Science. The data model in Figure 1 provides the high-level reference of how each step fits into the big picture. This model was developed for NOAA's needs and was based on a number of existing federal models.

There are six primary components (Plan, Obtain, Process, Preserve, Access, and Disposal) and four cross-cutting components (Document, Track and Monitor, Quality, and Security). These components apply to all data as defined in the scope of this Handbook, but the implementation and application differs for different types of data. The general description of each component is provided below with specific information on each component that is in scope for this Handbook described in Chapters 3 and 4 for environmental data and administrative data respectively.

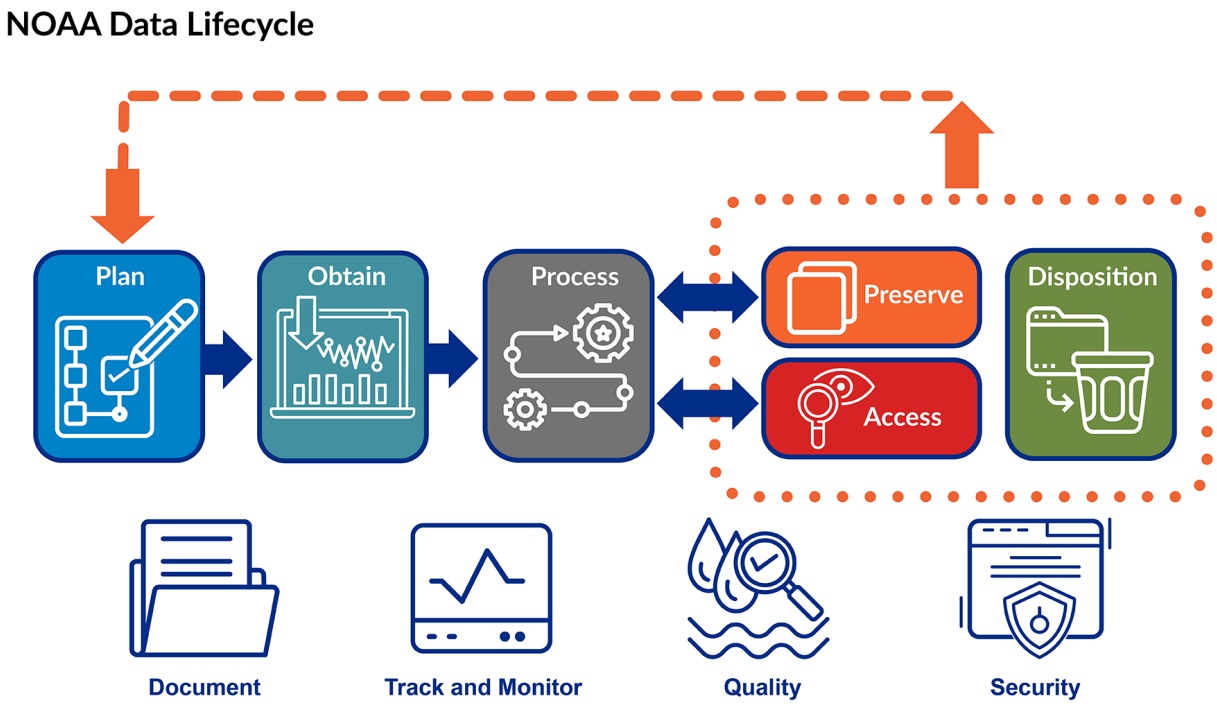


Figure 1: The NOAA Data Life Cycle

Primary Components

Plan - Plan and design the project, program, or research from onboarding to project closure, including methods and resources for data management.

Obtain - Obtain the necessary data through observation using NOAA-owned or leased systems, purchasing commercial data, or producing through models, data integration, or data analysis.

Process - Work with the data to make it useful to users and create products, including but not limited to transformations, calibrations, flagging, binning, subsetting, and scientific algorithms.

Preserve - Identify essential records, determine data required to be preserved, determine the method of preservation, align preservation with records retention schedules, curate the records, and coordinate discovery and access to archived records.

Access - Ensure the discoverability and availability and use of data by providing internal and public access to the data and metadata as appropriate in a timely manner.

Disposition - Re-appraisal of data based on community needs and records retention schedule to evaluate ongoing data management needs, including the option of destruction and deletion.

Cross-Cutting Components

Describe - During data collection and creation, document data and collection processes using standard metadata practices to ensure provenance and independent understandability.

Track and Monitor - Track data and metadata throughout the life cycle and monitor application of data management principles.

Data Quality - Ensure the quality, objectivity, utility, and integrity of the data. Further definition of and instructions on this component is the responsibility of the Science Council and outlined in NAO 202-735D-2: Scientific Integrity. It will not be expanded on in this Handbook.

Data Security - Protect data from unauthorized access, corruption, and theft. Further definition of and instructions on this component is the responsibility of the Office of the Chief Information Officer and will not be expanded on in the Handbook.

C. Key Terms

The following select terms are critical for this Handbook.

Assistant Chief Data Officer (ACDO) - Positions in each Line Office (LO) that lead data governance within the LO, and work in close collaboration with LO leadership and data program staff. The ACDO coordinates the implementation of the NOAA Data Strategy within the LO, including oversight of policies that ensure NOAA data are strategically and efficiently managed on behalf of the NOAA enterprise, maximizing the value of NOAA's data assets through sound and coordinated data governance and management practices. The ACDO is responsible for monitoring, tracking, and reporting on the metrics described within each data management section listed in this handbook ACDOs are responsible for performing outreach within their LO to ensure that data management roles are described and communicated to NOAA managers and rating officials to ensure consistent expectations in performance plans where possible.

Administrative Data - Derived from the operation or management of an organization or institution (Elias, 2014), and is collected for the purposes of registration, transaction, and record keeping and often associated with the delivery of a service (Woollard, 2014). Administrative data differs from experimental, scientific, and observational data in that it is found (rather than systematically made) data and is not primarily collected for research purposes. As such, it can be large, complex, and not necessarily collected in an organized manner that allows for linkages to other information or data (Connelly et al, 2016). Through the curation, enhancement, documentation, and accessibility of administrative data, NOAA can better understand how it achieves its mission and serves society (McGrath-Lone et al, 2022). This is especially important for the research and observational components of NOAA whose value in decision making and societal impacts is found elsewhere in the agency. Applications of administrative data allow for research in service delivery, user engagement, program management, social science, and

economic valuation. Administrative data can include personal identifiable information (PII) or other sensitive, controlled unclassified information. Public data collection is also subject to the Paperwork Reduction Act. Data collection and warehousing will follow federal and NOAA policies where applicable. (NAO-212-15B)

Data - Recorded information, regardless of form or the media on which the data are recorded. (Evidence Act).

Dataset - An identifiable grouping of data ***generally*** represented by a single metadata record and a Data Management Plan. Datasets can be characterized by an interrelated family of more specific units: grouping, content, relatedness, and purpose (Renear et al, 2010). A dataset may refer to a digital rendition of factual materials, or a product of a given version of an algorithm/model. A dataset may contain one or many physical samples or data files in an identical format, having the same geophysical variable(s) and product specification(s), such as the geospatial location or spatial grid.

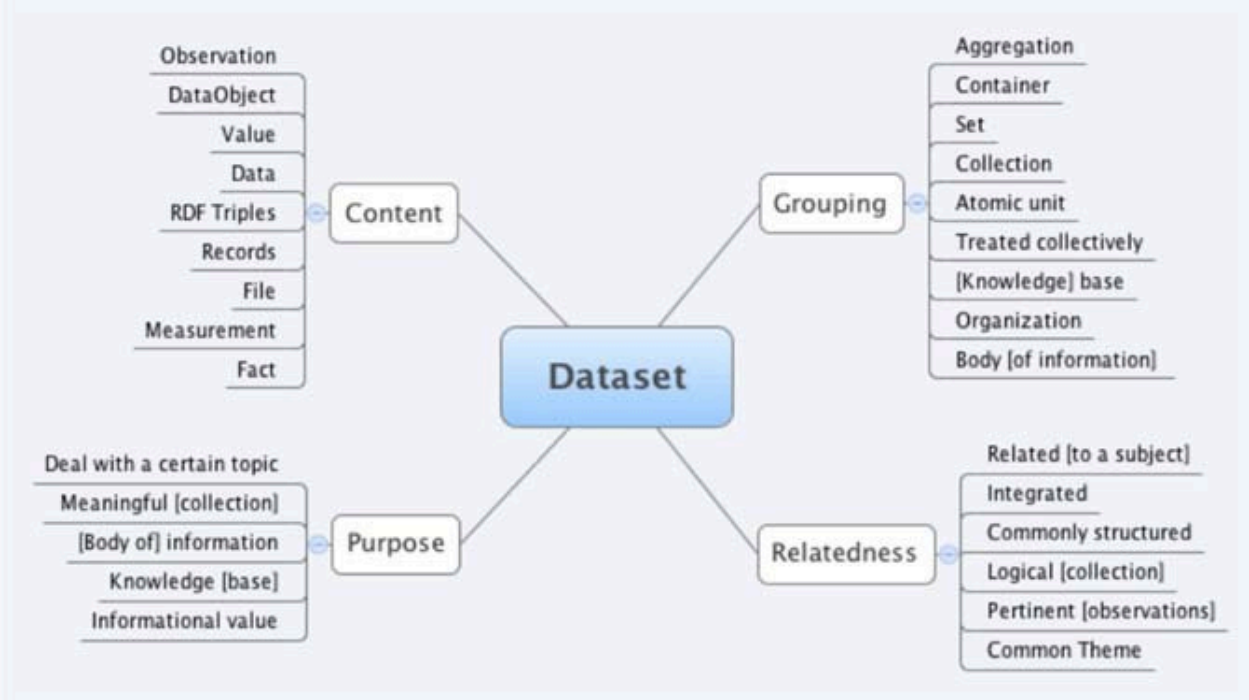


Figure 2: Dataset Diagram (Renear et al 2010)

Data Asset - A collection of data elements or datasets that may be grouped together. (NAO-212-15B)

Data Access - Services and functions which make the stored information holdings accessible to users by providing data search, discovery, retrieval, and dissemination functions. Access can refer to either the functionality, the services providing the functionality, or the entity providing the corresponding services. This includes data access from Cloud services, as referenced in

Reference Model for an Open Archival Information System (OAIS), Recommended Practice CCSDS 650.0-M-2, Magenta Book 2012.

Data Dissemination - 'Dissemination' means the government-initiated distribution of information to a non-government entity, including the public. The term 'dissemination,' as used within OMB Circular A-130, does not include distribution limited to federal government employees, intra- or interagency use or sharing of federal information, and responses to requests for agency records under the Freedom of Information Act (5 U.S.C. § 552) or the Privacy Act (5 U.S.C. § 552a). (OMB Circular A-130 Section 10 #18*)*

Data Steward - A person (or group of people) responsible for managing data assets in accordance with guidelines in this Handbook.

Data Life Cycle - The stages through which information passes, typically characterized as a creation or collection, processing, dissemination, use, storage, and disposition, to include destruction and deletion, per the OMB Circular A-130. (NAO-212-15B)

Data Management - A combination of two major activities conducted in coordination, data management services and data stewardship, which together constitute a comprehensive end-to-end process including movement of data and information from the observing system sensors to the data user. This process includes obtaining data, quality control, metadata development, cataloging of metadata, validation, reprocessing, storage, retrieval, dissemination, and archiving of data.

Data Producer - Individuals or groups who are responsible for creating, managing, or submitting environmental data to NOAA National Centers for Environmental Information (NCEI) and/or the NOAA Central Library.

Data Sharing - Making data publicly visible and accessible in a timely manner at no cost (or no more than the cost of reproduction), in a format which is machine-readable and based on open standards, along with metadata necessary to find and properly use the data.

Environmental Data - Data that includes, but is not limited to 1) recorded and derived observations; 2) measurements of the physical, chemical, biological, geological, and geophysical properties and conditions of the oceans, atmosphere, space environment, Sun, and solid earth, 3) correlative data, such as socio-economic data; 4) model outputs using or predicting data; and 5) related documentation and metadata. (NAO 212-15B)

Extramural Data - New environmental data created by extramural funding recipients (i.e., NOAA-funded grants, contracts, and cooperative agreements).

Information - Any communication or representation of knowledge such as facts, data, or opinions in any medium or form, including textual, numerical, graphic, cartographic, narrative, electronic, or audiovisual forms per the OMB Circular A-130.

Internal NOAA Source Data - Data that are generated by NOAA-owned sensors or systems or NOAA federal or affiliate employees. (See Appendix E, Data Licensing)

Metadata - Structural or descriptive information about data such as content, format, source, rights, accuracy, quality, provenance, frequency, periodicity, granularity, publisher or responsible party, contact information, method of collection, and other descriptions per the Evidence Act.

Machine Readable - A format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost, per the Evidence Act.

National Geospatial Data Asset - A core set of geospatial datasets as defined and managed by the National Geospatial Data Asset Management Plan in accordance with OMB Circular A-16 Supplemental Guidance.

NOAA-Level Metadata Catalog - An inventory of all NOAA datasets, documented with formal standards-compliant metadata and searchable by the public.

NOAA Observing Systems of Record - A list of observing systems as defined by the Observing Systems Committee (OSC), a subcommittee of the NOAA Observing Systems Council (NOSC).

Open Data - Publicly available data structured in a way that enables the data to be fully discoverable and usable by end users. In general, open data will be public, accessible, described, reusable, complete, timely, and managed post-release. See *OMB M-13-13 Section I* for additional details.

Persistent Identifier (PID) - A long-lasting reference to a digital resource, contributor, or organization such as Digital Object Identifiers (DOIs), Open Researcher and Contributor Identification, Research Organization Registry, etc.

Program - The organizational units the LO chooses to manage, track, and report data with the ACDOs as the responsible individual for defining, curating, and providing the list of Programs for use and tracking across NOAA.

Program Manager - Any individual or organization who is the lead for the life cycle of the data, typically the individual or organization funding and managing the acquisition of data although each Program Office may define it as it works best for their mission.

Timely – Defined as no later than publication of a peer-reviewed article based on the data, or two years after the data are collected and verified, or two years after the original end date of the grant (not including any extensions or follow-on funding), whichever is soonest, unless a delay has been authorized by the NOAA funding program.

Chapter 3 - Data Management for Environmental Data

A. Purpose

The purpose of Chapter 3 is to establish the requirements, objectives and measurements (if applicable), and procedures for implementation of the DOC NOAA Environmental Data Management policy (NAO 212-15B) for environmental data. Following the scope section that describes in more depth the relevant data for this Chapter, each section provides details for one component of data management.

B. Scope

This Chapter applies to all NOAA environmental data and to personnel and organizations managing these data, including federal and affiliate employees. The requirements described in each of the following sections apply to any NOAA-funded or NOAA-executed environmental data programs, including contracts, grants, and externally-funded programs or projects that collect or produce data from observing systems, field data collections, laboratory analyses, models, product generation, or research.

Specifically, the following datasets are in scope of the requirements in this Chapter:

- Derived data and/or data products
- Digital audio or video recordings of environmental phenomena
- Data collected in a laboratory or other controlled environment
- Model output
- Digital forecast data
- Socio-economic data
- Graphic/image displays of forecast data not otherwise included in digital form
- Third-party data purchased or acquired from external sources that NOAA disseminates or uses in externally facing products or services
- Environmental data in accordance with the Evidence Act and PARR
- All legacy data archived in digital form at the NOAA NCEI

The following are out of scope for the requirements in this Chapter:

- Third-party data purchased or acquired from external sources for internal NOAA access and use only.
- Legacy data that were produced by NOAA Programs which no longer exist and were never archived, unless those results are deemed worthy of public access by an existing Program which identifies the necessary resources for accessibility. Records retention schedules still apply to legacy data even if they are not subject to the requirements in this Handbook.
- Archival information disseminated by NOAA before June 30, 2008, and still maintained

by NOAA as archival material.
- Information relating solely to correspondence with individuals or persons.
- Data appropriately labeled as classified, Controlled Unclassified Information (CUI), or restricted and cannot be made public based on law, regulation, security classification, or contract.

In addition to datasets, scientific publications are included in the scope of the Data Citation Section. Further guidance and procedures to meet the full PARR requirements will be included in a subsequent version.

C. Dataset Tracking and Reporting

Purpose

It is critical for data management to have a comprehensive list of all in-scope environmental datasets in the organization. The datasets on the list will be the basis for reporting under the Evidence Act and future metrics. The comprehensive list will be used to evaluate impact, determine compliance, highlight gaps to support resource requests, and demonstrate progress. Unless otherwise specified, all performance metrics listed in this Chapter will be reported by each LO ***ACDO*** to the NOAA ***Chief Data Officer (CDO)*** via the NOAA ***DGC*** at least annually. Questions regarding LO reporting procedures and mechanisms should be directed to the LO ACDO.

Data Management Requirements

NOAA Programs, as determined by the LO ACDOs, are required to maintain a current list of all in-scope datasets that will be used to evaluate and report data management metrics stated in the following sections. Refer to Chapter 3, Section B Scope for the definition of in-scope and Chapter 2, Section C. Key Terms for the definition of a dataset.

DGC is responsible for:
- Establishing and maintaining a NOAA-level document repository
- Examining Agency dataset trends and reporting to DOC

ACDOs, or their designee, are responsible for:
- Coordinating across Offices and Programs within their LO to ensure comprehensive and consistent tracking and reporting of datasets
- Ensuring that Programs regularly update the information
- Reporting on the metrics to the DGC at least annually

Program Managers,or their designee, are responsible for:

- Identifying all datasets produced or hosted by their program
- Tracking conformance with the data management requirements as defined by the sections in this Handbook

- Reviewing and updating the information at least annually to meet reporting requirements

Performance Objectives and Measurements

Objective C.1: Dataset tracking is up to date with all LO Offices and Programs reporting.

Metric C.1.1: Number of datasets that are tracked.

Procedures

1. Each Line Office's ACDO will administer dataset tracking for their LO.
    a. Establish roles and responsibilities for tracking (for example: ACDO, Program Managers, Data Management Working Group Members, Program Data Officers, project data managers and data specialists, extramural funding federal Program Officers and Contracting Officers, *etc.*) and designate point(s) of contact with responsibility for understanding and providing updated information for tracking at least annually to the ACDO
    b. Ensure all points of contact understand the data stewardship requirements as defined in this *Handbook*
    c. Collaborate with partner organizations that provide data services and report on services to ensure accurate reporting to the DGC, including NOAA Geoplatform, NOAA Institutional Repository, and NCEI.
    d. Establish and maintain a dataset tracking system with a comprehensive list of all in-scope datasets as defined in Chapter 3, Section B. Scope with at least the following information:
        i. Logistical (start date, last update, anticipated end date, responsible party, active or inactive/archived status)
        ii. Type of data effort(s) (Observing System of Record, program, project, model, grant, contract, *etc.*)
        iii. Existence and link to Data Management Plan with data of last review
        iv. Existence and link to metadata in compliance with the Data documentation Section
        v. Status of Data archiving, including whether archived or not, appraisal status and/or archive recommendation status, data repository location
        vi. Link to the Data DOI(s)
        vii. Link to public Data Access and any associated license(s)
    e. Prepare and deliver a tracking report for the DGC at least annually

2. The DGC will establish a standardized and centralized review process to manage and track submissions of Data Management Plans to a NOAA data document repository; tracking will include:
    a. Approval information (date submitted by responsible party, date approved by approving individual, *etc.*)
    b. If applicable, information from the LO Data Management Plans (DMPs) repository (date submitted by responsible party, date accepted by repository, *etc.*)

D. Data Management Planning

Purpose

The goal of data management planning is to ensure that all parties involved have common expectations and coordinate data plan execution to ensure proper documentation, accessibility, and preservation for future use. The purpose of this section is to define the requirements for DMPs, who should write the plan, what metrics are collected, and the procedures for writing and managing DMPs. This section directs managers of all data production or data collection programs and systems to ensure DMPs are developed for their data. A Data Management Plan Template with questions to be addressed regarding NOAA environmental data are provided (See Appendix C).

Data Management Requirements

NOAA Programs are required to consider, in advance of funding or executing data collection, production, or acquisition, how to meet data requirements in this Handbook to ensure public accessibility to the data, and the long-term preservation according to the appropriate NOAA records retention schedule.

DGC is responsible for:

- Maintaining a NOAA DMP Repository.
- Reviewing and determining whether to approve LO/SO-specific modifications to the DMP Template.

ACDOs, or their designee, are responsible for:

- Coordinating and providing guidance on DMP creation and submission for their LO;
- Ensuring a centralized review process is established and maintained to manage and track submissions to the NOAA DMP repository managed by the DGC, ensuring a standardized approach is in place;
- Seeking DGC approval for Line and Staff Office-specific modifications of the Data Management Plan Template (Appendix C) as listed in the Procedures;
- Tracking and reporting on the availability of funding for data management using the information provided in DMPs.

Program Managers, or their designee, are responsible for:

- Identifying resources within their own budget to manage data they produce;
- Communicating new or existing data generating programs to their ACDO;
- Writing, reviewing, and following the plan throughout the data life cycle, and revising as circumstances warrant as well as submitting the plan to ACDO;
- Coordinate with the respective supervisors to ensure that employees responsible for any aspect of data management have that role clearly stated in their performance plan and have the authority, training, and means to carry out their responsibilities;

- Coordinating with NOAA NCEI or, if out of scope for NCEI, another applicable data repository to ensure that the plan is appropriate, resources are identified as needed, and the schedule is feasible for long-term data preservation.
- NOAA Programs are required to consider, in advance, how to ensure public accessibility and long-term preservation of externally-funded data, to provide guidance for proposers to use in developing a plan for data access, and to track and enforce conditions imposed on awardees

Performance Objectives and Measurements

Objective D.1: All NOAA environmental data are covered by a DMP.

Metric D.1.1: Percentage of datasets that are covered by a DMP.

Objective D.2: NOAA environmental DMPs are effectively and completely written to meet data management requirements throughout the data life cycle, inform decisions, and set common expectations.

Procedures

1. All NOAA environmental data shall have an associated DMP. All NOAA Programs or Systems that produce or collect environmental data shall have DMP(s) for the data they produce internally or for data commissioned via contracts or grants.
2. A single DMP may cover multiple data types that are managed similarly by the program.
3. DMPs may be hierarchical. Specifically, a master plan applicable to a group of observing systems or data types may be developed, supplemented by more specific plans that inherit the provisions of the master plan and provide data-specific details. In this case, the specific plans shall include a reference to the applicable master plan.
4. DMPs shall be followed and maintained throughout the full Data Management Life cycle for all of the data within the scope of the Plan.
5. DMPs shall be revised or superseded as needed if circumstances change.
6. The scale and complexity of DMP(s) may vary with the scale, complexity, and significance of the data being produced; however, all DMPs shall:
    a. Provide an appropriate level of detail to address all elements of the data management life cycle.
    b. Be followed and maintained throughout the full life cycle of the data.
7. DMPs shall be based on the Template (Appendix C). This generic Template may be modified to better meet the needs of NOAA LO/SO as follows:
    a. NOAA LO/SO may not reword existing questions in the Template without approval by the DGC and updating of this Handbook's Procedures. The intent of standardized wording is to ensure consistency across the resulting plans and to enable automated assessment.
    b. NOAA LO/SO may extend the Template by adding questions, in which case the LO ACDO will be notified of such extensions and coordinate as appropriate.
    c. NOAA LO/SO may adopt a simplified version of the Template (by removing or making optional some questions), subject to approval by the DGC. Approval is

required to ensure this directive is not weakened by making key questions optional.

8. DMPs shall be submitted to the ACDO, or designee for review, approval, and depositing into the designated NOAA Google Drive DMP Repository.
    a. ACDOs or designee shall have the authority to create and manage their designated section of the NOAA DMP Repository.
    b. DMP authors shall coordinate submission with their ACDO. DGC Membership, including the list of ACDOs can be found on the internal NOAA Data Community site.
9. NOAA Programs that fund Grants, Cooperative Agreements or Contracts shall ensure, *prior to issuing a Federal Funding Opportunity (FFO) Announcement, Grant or Contract Solicitation,* that the program develops Data Management Guidance describing in general terms how the Program intends for the resulting data to be managed by the recipient and what, if any, NOAA support (such as archiving at NCEI) may be provided.
    a. The Data Management Guidance shall be based on the Template in Appendix D, but may be tailored to fit the needs of the Program.
    b. The Data Management Guidance shall be included or referenced in the requirements documents for financial assistance and contract actions, in order to assist proposal submitters preparing Data Management Plans.
    c. During development of the Data Sharing Guidance, programs that plan to fund the production of data shall consult with NOAA NCEI regarding the feasibility and potential costs for archiving the data at NCEI and to enable NCEI to reach a preliminary understanding of potential capacity, resource, or technical issues.
    d. NOAA Programs shall ensure offerors and recipients of NOAA funding are aware of their role in satisfying the requirements of this Handbook by including text from Appendix D in all requirements documents for financial assistance awards and contracts that include data collection or production.
        i. The text in Appendix D shall be included in all FFO Announcements and Contract Solicitations anticipated to produce data in order to ask for Data Management Plans in proposals and to advise proposers of these requirements if funded.
        ii. The text in Appendix D shall be included in all Announcements and Solicitations not anticipated to produce data.
        iii. The text in Appendix D shall be included in all Notices of Award and Contracts anticipated to produce data in order to require recipients of Grants, Cooperative Agreements, or Contracts to share data, submit manuscripts, cite data used in publications, and report funding sources using Open Funder Registry (formally known as FundRef).
    e. Data Management Plans shall be evaluated as part of the proposal review process; proposal submitters and reviewers shall be informed of this evaluation criterion.

E. Data Documentation

Purpose

The goal of data documentation is to ensure that NOAA environmental data are documented with human-presentable and machine-readable metadata to enable discovery, access, and use of these resources. The purpose of this NOAA Data Documentation Section is to define the requirements for metadata, performance metrics, and procedures to meet this goal and supports NAO 212-15B, Management of Environmental Data and Information (2023), which states that "environmental data will be visible, accessible and independently understandable to users, except where limited by law, regulation, policy, … or security requirements." In addition, per Congress, NOAA is required by the Evidence Act to maintain a comprehensive data catalog that feeds to Data.gov, and by the Geospatial Data Act to document geospatial data with standard metadata in the national GeoPlatform.

Data Management Requirements

NOAA Programs that produce or maintain on-line or off-line environmental data are required to ensure their data are properly documented with metadata, as defined through the metrics below, and that the metadata are included in the NOAA-level metadata catalog.

All NOAA environmental data, as defined in Chapter 3, Section B. Scope, shall be described with a metadata record that conforms to the current NOAA environmental metadata standard. Metadata records shall be created for online digital data and offline data, including data on paper, ROM disks, magnetic tape, and other physical media. Offline data may be documented at the collection or facility level if necessary (e.g., one record summarizing a room-full of analog data holdings). If an online copy of offline data is available, the offline version is not required to have a separate metadata record. Metadata records shall be made publicly accessible online, such as via a Web Accessible Folder or an Application Programming Interface (API). Metadata records shall be maintained in such a manner that the NOAA-level metadata catalog can extract, reuse and make publicly accessible metadata in accordance with current NOAA-level metadata catalog processes.

Program Managers, or their designee, are responsible for:

- Enforcing the provisions of this Section's procedures for the data they produce or maintain including metadata requirements, metadata content, metadata granularity, and metadata format.
- Ensuring that metadata includes clear data license information. (See Appendix E)
- Maintaining the metadata, including reviewing and updating in a timely manner consistent with the needs of the dataset.

NOAA Catalog Working Group is responsible for:

- Coordinating across the enterprise to govern the NOAA-level metadata catalog and the

public search interface.
- Ensuring that all datasets in the NOAA-level metadata catalog are findable.
- Coordinating metadata from the NOAA-level metadata catalog with any external catalogs that the NOAA metadata will be incorporated into.
- Ensuring metadata from the NOAA-level catalog are available for harvest into data.gov.

NOAA Enterprise Metadata Working Group is responsible for:
- Developing and promoting best practices for data documentation for internal and external NOAA stakeholders, including guidance on the creation and maintenance of metadata content and templates used across NOAA programs to promote cross-NOAA data discovery and interoperability.
- Maintaining the NOAA Metadata Completeness Rubric.

NOAA Metadata Tool developers are responsible for:
- Ensuring that systems supporting the authoring and discoverability of standard International Organization for Standardization (ISO) Geographic Metadata Standard 19115-2 metadata follow NOAA content best practices.

DGC is responsible for:
- Ensuring support for the implementation, tracking, and maintenance of this section.

Performance Objectives and Measurements

Performance objectives have been established in order to ensure the completeness of NOAA's public data inventory, sufficient metadata for data discovery and use, and traceability to NOAA Programs responsible for data production. Metrics will be computed automatically to the extent possible.

Objective E.1: All NOAA datasets, whether on-line or off-line, have a publicly-accessible ISO Geographic Metadata record.

Objective E.2: Metadata records in the NOAA Data Catalog include all required information and all relevant optional information

> **Metric E.2.1**: Percentage of datasets documented with metadata in the NOAA Data Catalog.

Objective E.3: Metadata records in the NOAA Data Catalog identify the Program responsible for the record.

Procedures

Note: Data archived at the NOAA NCEI are documented in accordance with these requirements and procedures at or before inclusion as NCEI data holdings and are therefore considered to have met all requirements of this Section. Data not submitted and accepted by NCEI must independently satisfy the requirements below.

Metadata content: Metadata records shall include, at a minimum, the information necessary to enable data consumers to discover, access, evaluate, and use the dataset, and may include references to supplementary documentation.

1. *Completeness*: Completeness is a measure of how many of the metadata elements in a current standard metadata record contain information about the data being documented.
    a. Completeness includes detailed information about necessary characteristics of each variable to ensure that those variables can be used or compared with other similar measurements.
    b. Completeness includes contextual information that is sufficient for deciding if data are appropriate for use in specific situations.
    c. Appendix F Part 1 summarizes the mandatory and optional content for dataset or collection level metadata records.
    d. Appendix F Part 2 describes the current metadata completeness tool ("the Rubric tool") which is used to evaluate metadata completeness.
    e. Appendix F Part 3 describes the high-level sections of metadata records.
2. *Metadata Quality*: Metadata quality is a measure of how well a complete descriptive metadata record supports the discovery, access, and reuse of the data being documented. Quality is not necessarily the same as completeness: The 'Rubric tool' does not assess quality, it assesses only completeness (i.e., are mandatory elements populated with a value, not is the value in the mandatory element accurate or correct).
    a. Human-review of content prior to publishing metadata is highly recommended to ensure a minimal level of information quality. Metadata tools mentioned in Appendix F Part 2 support a 'second set of eyes' quality review step.
3. *Accessibility*: Accessibility is a measure of how well descriptive metadata can be found using online tools that can search standards-based metadata records and how the data that is described can be acquired.
    a. Metadata records for data that are accessible on-line shall include links to machine-readable data-access methods as required by the NOAA Data Access Section, and should include a description and location of the dissemination system(s) used to provide access to each dataset, including the methods and interfaces supported.
    b. Metadata records for data not accessible on-line shall include the physical location (address) of the data, contact information for the data custodian or data position, planned data dissemination date (if any), instructions for off-line access (if available), and any constraints on access.
4. *Data Quality & Integrity*: Metadata must include a description of quality, authenticity and lineage (to establish provenance) for data and systems to ensure users can be confident in their use of data and data access points.
    a. Document data quality and integrity by providing as much information about the completeness, accuracy, consistency, context and uncertainty of the data as possible in metadata or in documentation openly available to data users (i.e., websites or system documentation pages).

Metadata granularity: Wherever possible, metadata records shall describe aggregations of closely-related observations, often referred to as a “collection”. The intent of aggregation is to

simplify discovery and use of the data; the appropriate level of aggregation depends on the nature of the data.

1. Data from an ongoing time series, and collections of related data from multiple locations, shall be represented by an aggregated metadata record.
2. Other levels of aggregation shall be at the discretion of the producers and stewards of the data.
3. If necessary, individual observations that are part of an aggregation may have individual metadata records in addition to the aggregation record; these individual records shall include a reference to the aggregation record.

Metadata format: Metadata records shall be formatted as machine-readable, syntactically-valid Extensible Markup Language (XML) files based on international standards.

1. The current default encoding for NOAA metadata shall be ISO. The ISO version currently in effect at NOAA is ISO 19115-2 (2009) conceptual model with ISO 19139-2 (2012) XML schema.
2. Legacy metadata records in U.S. Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata format shall no longer be accepted. NOAA data managers are required to use existing FGDC Current Standard for Digital Geospatial Metadata transforms to convert metadata to the ISO Geographic Metadata Standard so that ISO-compliant metadata can be included in the NOAA metadata catalog and other discovery and access tools.
3. No metadata standards or formats other than those described in this Handbook are currently accepted for registration in the NOAA Data Catalog.

Additional Recommendations: Metadata records that conform to the ISO 19115-2 standard, as documented in this Section, may only be part of the descriptive information necessary to make a data collection or dataset understandable or usable.

1. *Ancillary metadata*: Ancillary metadata, such as data dictionaries that define the individual characteristics of specific variables and references to other works that provide contextual information about the data are invaluable resources that should be identified in a standard metadata and data package.
2. *Human-readable documentation:* To provide access to visually-impaired users, it is a recommended practice that archival data packages include a Section 508-compliant human-readable PDF of the metadata content.
3. *File-level documentation:* The ISO 19115-2 metadata standard supports detailed, structured representation of descriptive information about individual variables. If the file-level and field-level metadata to document the structure and content of the geospatial data are not included in the ISO record, it is important to provide an alternate, human-readable representation of these details to provide the necessary context for other data users. Additional documentation in the form of data dictionaries, such as that used by the Environmental Research Division Data Access Program (ERDDAP), and other contextual documentation or links to related documentation is encouraged.

F. Data Preservation

Purpose

The goal of data preservation is to ensure NOAA data are broadly available, independently understandable, and reusable across sectors, communities, and generations. These goals are supported by and depend on the requirements and procedures in the other Sections. The purpose of this Section is to define requirements, metrics, and procedures for appraising and preparing data for long-term preservation and access as well as governing and managing data, including if those data are archived within NOAA or using an external repository. This Section specifically ensures compliance with mandates and directives for records management from the National Archives and Records Administration (NARA), the NOAA Records Disposition Handbook, the Evidence Act, and PARR. The OAIS Reference Model (RM), ISO 14721, (Appendix G Part 1) defines the requirements for preservation of digital data in a repository. Guidance in this Section is based on the concepts described in the OAIS RM.

Data Management Requirements

NOAA Programs are required to manage data they create, collect, or receive in accordance with NOAA records management schedules. Data for which there is no documented disposition authority must be preserved until a disposition authority (i.e., a 'records schedule') is approved by NARA. NOAA Programs shall collaborate with NCEI to appraise all environmental data for archiving and shall archive data deemed appropriate for long-term preservation at NCEI or an alternate, recommended repository. NCEI will lead an appraisal process that determines which scientific records are preserved in a NOAA archive based on uniqueness, replicability, long-term value, legal mandates, quality, documentation, and records retention requirements. The designated repository is required to retain data for at least the legally-mandated retention period, as determined by the appropriate records retention schedule, and to provide discovery and access to data in accordance with applicable federal laws, regulations, and mandates.

The DGC is responsible for:

- Defining, managing and adjudicating an appeal process for appraisal decisions made by NCEI.
- Advocating for resources at the NOAA level.
- Working with ACDOs and NCEI to maintain a list of appropriate archival facilities referenced in Appendix G for data that are out of scope for NOAA archival facilities.

ACDOs, or their designee, are responsible for:

- Comprehensive tracking of data streams and data assets that need to be appraised for archiving.
- Providing a short rubric to help Program Managers identify which data are high priority for long-term retention, based on applicable records management schedules and requirements.
- Identifying datasets deemed high-priority for archiving but which are not yet archived.

- Tracking which data have been archived and links to where they are accessible.
- Advocating for resources at the LO level.
- Coordinating with NCEI to recommend an appropriate archival facility.
- Coordinating with Program Managers to identify or define appropriate retention periods for data, as necessary.
- Working with NCEI through the appraisal process to produce a list of data that should be archived.

Data producers are responsible for:

- Working with NCEI in the development of their DMP to plan for appropriate long-term preservation.
- Archiving data (when and where) as outlined in their DMP and in accordance with NOAA records schedule requirements.
- Preparing and maintaining data and metadata for archive appraisal and long-term preservation and use per the Data Documentation and Data Access Sections.
- Identifying and communicating access needs and limitations.
- Ensuring that they have the resources to archive data per their DMP.
- Resourcing and communicating gaps in resources.

NCEI is responsible for:

- Appraising data for archival retention in coordination with ACDOs, Program Managers and/or Data Stewards, documenting the decision, notifying the data submitter of the decision, and allowing for an appeal to the DGC with a recommendation for external reviewers knowledgeable in data management and/or scientific records as appropriate.
- Providing archiving guidance, archival documentation tools, preservation and archival access services to Data Stewards for data deemed appropriate to archive with NCEI.
- Advising Data Stewards about potential alternate trusted repositories based on specific repository characteristics, such as those defined in NCEI Policy for archiving of NCEI-produced data and products (2019) and outlined in Appendix G Part 2.

The archival facility (NCEI or alternate authorized repository, both have the same responsibilities) is responsible for:

- Providing a statement of collection scope that defines what types of data/records are accepted (within scope) and not accepted (out of scope).
- Providing detailed guidelines to users on how data should be prepared and documented for archiving.
- Appraising data for archival retention in coordination with ACDOs, Program Managers and/or Data Stewards and documenting the decision
- Providing a process to request archiving in the repository.
- Providing archival documentation tools for the repository.
- Providing preservation services that meet NOAA, NARA, and federal requirements, regulations and laws related to preservation and access to federal data and records.
- Maintaining an inventory of archived data.
- Providing discovery and access to archived data.

- Providing input information to NOAA ACDOs to ensure that metrics defined in this Section are reported accurately in a timely manner.

Performance Objectives and Measurements

Objective F.1: Archive all of NOAA's data determined to have long-term archival value, with a priority on data that have the highest value for NOAA and the public.

Objective F.2: Archived data are discoverable and accessible to the public.

Procedures

Preservation Planning: Beginning with the DMP, the Program Manager, Data Steward and/or LO ACDO shall coordinate with NCEI staff to review data that have been or will be collected or created, so that appropriate resources are identified and available for ensuring that data collected in accordance with the DMP are properly managed for the entire data life cycle. NCEI staff have extensive experience with planning and appraisal processes and are familiar with current NOAA, DOC, and other federal requirements for managing scientific data.

Preservation Management:

1. *Appraisal*
    a. The archival process begins with an appraisal step. Appraisal is used to confirm that data are correctly identified as long-term or short-term records, what specific characteristics are known about the data being appraised, what repository is most appropriate for retaining the data, and providing evidence supporting a decision to accept or decline the data for a specific repository. Appendix G Part 3 provides the detailed appraisal process procedures. Appendix G Part 4 and 5 provide a summary of the appraisal questions and the full questionnaire respectively.
    b. The designated repository for NOAA data shall provide a mechanism to create a request to archive data in the repository which is used to appraise the data. This request process may vary depending on the repository that is selected by the Data Manager. The request process will likely include requirements for identifying the data source, parties responsible for funding, collecting, processing or taking other actions on the data and other characteristics and related metadata for the data. Data that are identified in the appraisal process to be within scope of the designated repository (NCEI or other repository), shall be acquired by the repository and added to the repository holdings in accordance with the repository's published processes.
    c. The NCEI Archive Collecting Policy defines the scope of NCEI data holdings. NCEI maintains several well-defined tools to support the request to archive process, including but not limited to Send2NCEI (for relatively small, non-repeating data), Advanced Tracking and Resource tool for Archive Collections - ATRAC (for relatively large, repeating, or otherwise complex data),

and domain-specific tools for paleoclimate data, bathymetric, water column acoustics and other geophysical data. NOAA data deposited at NCEI adheres to all federal laws, orders, and regulations related to archival management, duration of retention, and public accessibility.

d. Non-NCEI repositories have different collection scope and management policies, request to archive tools or practices, retention policies, and other requirements. Regardless of a non-NCEI repository's policies, NOAA data are required by Federal law to be managed in accordance with all Federal laws, orders, and regulations related to duration of retention and public accessibility.

2. *Ingest*
    a. Repository ingest processing is the mechanism for transferring data from the data provider/data source to the repository. This transfer may be done as a one-time transaction or as a repetitive, automated or semi-automated or manual process. The best practice for a particular dataset is to consult with the data repository in advance, i.e., the data management planning stage, to identify which mechanism is most appropriate for the data that will be transferred.
    b. A key aspect of the ingest process is to ensure that there is no data loss or data corruption during the transfer process. Each repository shall identify, document, and maintain the IT infrastructure necessary to support secure transmission of data from the data source to the repository. This may include, but is not limited to, providing a file manifest describing the content of the data transfer, valid filename constructions that can be verified using repository software, or other validation schemes to ensure that there is no loss or modification of data during the transfer process.
    c. The NCEI request to archive process is typically initiated using Send2NCEI or ATRAC. Other requests to archive processes may be used, but NCEI must appraise and decide if data are within the NCEI collection scope prior to accepting data for archiving. At present, NCEI uses multiple ingest processes and procedures that are based on the type of data, i.e., a specific ingest process for bathymetric, water column, acoustic, or paleoclimate data that is optimized for those data types or an automated acquisition and ingest process for a satellite data stream that is optimized for that type of data.
    d. Non-NCEI repositories will have variable ingest processes and requirements. When using a non-NCEI, the NOAA program or office using a non-NCEI repository must understand the repository's scope, appraisal processes, and ingest processes. Appendix G Part 2 identifies requirements that define alternate repositories used to manage NOAA data records for the long-term. Any repository used by NOAA must meet and maintain accurate information about their processes and capabilities.
3. *Data management*
    a. Data management in the context of the preservation function (as defined by the OAIS Reference Model) focuses on ensuring adequate documentation to make the archival data independently understandable to a designated community of data consumers.

b. NCEI accepts a wide diversity of data types as outlined in the Collection Scope policy. By necessity, this diversity of scope anticipates a similar diversity of file formats that support these different data types. Data represented in preferred file formats are more easily acquired, ingested and documented by NCEI. Other data formats may be accepted, at the discretion of NCEI. NCEI archival metadata shall include complete documentation of all variables, in accordance with the guidelines defined in Section E. Data Documentation.
c. NOAA data archived at NCEI are managed in accordance with all applicable NOAA, NARA, and federal regulations, requirements, and laws including but not limited to the Evidence Act, Public Access to Research Results, Open Data Policy, and other mandates.
d. Non-NCEI repositories will have variable metadata processes and requirements. When using a non-NCEI repository, the NOAA program or LO using a non-NCEI repository must understand the repository's requirements for data formats and documentation. However, NOAA data are required to be managed in accordance with all applicable NOAA, NARA, and federal regulations, requirements, and laws including but not limited to the Evidence Act, NOAA records schedules, Public Access to Research Results, U.S. Open Data Policy, and other mandates.

4. *Archival storage*
    a. Archival storage refers to the IT infrastructure that supports the long-term inclusion, retention, migration, and physical storage of digital data in the repository. Standards for archival storage technologies are defined by NARA and federal IT regulations and laws. These standards include, but are not limited to, requirements for IT systems management like locations of backup systems and frequency of system backups.
    b. NCEI adheres to and enforces relevant IT infrastructure requirements for archival IT systems.
    c. Non-NCEI repositories will have variable archival storage, backup, and other IT system processes and requirements. However, NOAA data are required to be managed in accordance with all applicable NOAA, NARA, and federal regulations, requirements, and laws including but not limited to the Evidence Act, Public Access to Research Results, Open Data Policy, and other mandates.
5. *Archival access*
    a. Data preservation requires the repository to provide access capabilities for the preserved data. The designated repository for NOAA environmental data shall provide at least one access capability for a data consumer to obtain a copy of archived data. Data must be available to the public, in accordance with the Evidence Act and other federal mandates unless expressly exempt from public release by Federal law or regulation. Data that are not expressly protected from public access under a Freedom of Information Act (FOIA) exemption must be released to the public upon request.

G. Data and Publication Citation

Purpose

The goal of data and publication citation is to:

- support traceability and scientific reproducibility through direct unambiguous connection to data used in derived products and results
- enable acknowledgement or credit for Data Producers, data distributors, and other contributors in the data production and stewardship process
- encourage submission of data and satisfactory metadata to NCEI or another long-term repository for long term preservation
- encourage submission of publications and satisfactory metadata to the NOAA Central Library for long-term access, as required by NOAA policy
- track data impact through references in scientific literature
- enable gathering additional bibliometric information regarding how data are being used
- improve NOAA's ability to catalog its data holdings.

The purpose of this Section is to establish the requirements, metrics, and procedures to assign and obtain unique PIDs to data and publications retained, produced, or otherwise managed by NOAA. This includes data archived at NCEI and publications deposited in the holdings of the NOAA Institutional Repository. This Section also describes requirements for related landing pages providing dataset information and access instructions, and guidance for citing NOAA data and publications by internal and external users. The purpose and syntax of DOIs, appropriate levels of granularity for a data collection, how to obtain a DOI, and the recommended data citation format are included in Appendix H. This Section applies to NOAA and non-NOAA data providers, to NOAA users of NOAA data, the NOAA Central Library, and to NCEI.

Data Management Requirements

To meet the goals of data citation there are requirements for both NOAA data-producing programs and NOAA users of data. NOAA Programs that produce environmental datasets for publication or long-term data archives are required to obtain PIDs for those data. In keeping with best scientific practice NOAA users of data are required to cite data used in projects and results.

Program Managers are responsible for:

- Understanding requirements in this Section and receive assigned identifier(s), and provide data and comprehensive metadata as outlined in this Section.

Data Users are responsible for:

- Citing data used in projects and results, using the assigned DOI if available.
- Informing data stewards of the requirements for data without identifiers citing external data in a manner similar to NOAA data, including the identifier if available.
- Having their data listed as the first reference in the citation section of the paper or list the data in the availability section of the journal.

NOAA Data Citation Working Group supports NOAA by:

- Acting as an advisory panel for the overall process of managing NOAA DOIs.
- Acting as a central point of contact to NCEI and the NOAA Central Library.

NOAA NCEI is responsible for:

- Issuing NOAA data DOIs.
- Ensuring long-term maintenance and availability of landing pages, metadata, and data.
- Collecting metrics on the count of NOAA data DOIs created.
- Coordinating and linking related data and publication DOIs.

NOAA Central Library is responsible for:

- Responding to questions.
- Receiving requests for NOAA publication DOI assignment.
- Assigning identifiers to publications.
- Collecting metrics on citation of NOAA DOIs in publications.
- Coordinating and linking related data and publication DOIs.

Publications published (or submitted for publication) prior to the effective date of this Handbook shall not be required to retroactively add citations for data. Future publications based on data that have not been issued identifiers shall not be required to amend the data citations therein if identifiers are issued to data after publication or submission. NOAA data and publications can request a reserved DOI that can be included in the data or publication to facilitate linkage. See Appendix H Part 1 on reserving DOIs.

Performance Objectives and Measurements

Objective G.1: NOAA Published or Archived datasets have a DOI that is used to cite the data.

Procedures

Citing data and publications

1. These are the NOAA procedures for requesting, issuing, or reserving a DOI for data and publications.
2. A DOI shall be considered a "NOAA DOI" only if it is created using the officially-recognized NOAA DOI registrar service and the data or publication meets the requirements outlined in this handbook. The prefix for a NOAA DOI is defined by the registrar that has the current contract for providing DOIs to NOAA. Consequently, the prefix for NOAA DOIs may periodically change if a new registrar is selected, but in no case will an existing DOI prefix change as a result of a registrar vendor change. If using an outside NOAA source for DOIs (such as Zenodo or GenBank) the rules and procedures of that source should be followed. DOIs from outside sources are not NOAA DOIs.
3. NOAA staff, or any non-NOAA individual, who would like to use a DOI for a publication or data may request that a DOI be assigned to a NOAA archival dataset or NOAA publication.
4. NOAA individuals and NOAA projects that use NOAA data and publications, such as in a

research paper, a derived product, a model assimilation, or a policy decision, shall be required to cite the data or publications used.

5. If a DOI has been assigned to data or a publication, the citation shall include the DOI and shall follow the citation guidance provided in the landing page for the data, or include the equivalent information in a specific citation format that has been mandated by the journal or publication.
    a. If a DOI has not been assigned, the citation shall include at least the data or publication title/name, creator/author, publisher, date of access, and subset used (if applicable). Data users are encouraged to request assignment of a DOI to data of interest by contacting NCEI at NCEI.DOI@noaa.gov.
    b. Non-NOAA users of NOAA data and publications are strongly encouraged to cite NOAA resources in the manner described above.
6. NOAA shall maintain an agency-wide license for a DOI registrar to mint DOIs and shall renew the license annually until this provision is rescinded.
7. NOAA DOI Syntax
    a. The syntax, eligibility, granularity, and other characteristics of NOAA dataset and publication identifiers and their citation from any source are defined as follows:
        i. A NOAA DOI shall be structured as a numeric prefix (assigned to NOAA by the DOI registrar) followed by a suffix comprising an arbitrary sequence of letters and digits.

        ii. NOAA DOIs shall not be structured to contain the names of organizations, programs, observing systems, or data.
        iii. The DOI standard (ISO 26324) shall be used for NOAA data and publication identifiers. DOIs are currently being issued by over 5000 naming authorities (including publishers and science data centers), with over 55 million DOIs assigned as of 2012 to datasets, books, and other digital objects.

Data DOI Procedures

1. Requesting a NOAA Data DOI
    a. NOAA data DOIs are issued when data are accepted by NCEI for archival and long-term preservation. Not all DOI requests will necessarily be granted by NCEI. NCEI does not issue DOIs to dataset that already have a DOI.
        i. When there is uncertainty about whether a DOI will be assigned, send questions to NCEI.DOI@noaa.gov.
        ii. If a DOI is requested from NCEI, DOIs may be assigned at a different level of granularity than requested based on previously determined collection level data.
        iii. Requests to assign a NOAA DOI for data that were archived prior to this policy going into effect may be sent to NCEI at NCEI.DOI@noaa.gov.
        iv. The NCEI DOI Granularity Decision Tree tool shall be used by NCEI to assist with making decisions about appropriate data DOI granularity.

2. NCEI Data DOI Issuance
    a. NCEI shall issue DOIs for all new accessions added to the repository unless there is a verified reason not to.
    b. Long term preservation and documentation shall be required prior to receiving a NOAA data DOI from NCEI.
        i. Data shall have a comprehensive metadata record that complies with the NOAA Data Documentation Section and NCEI best practices.
        ii. Metadata for data shall include an image that is related to the content of the data, in accordance with recommended practices.
    c. NCEI shall designate personnel who have the authority to issue and manage DOIs and to decline requests to NCEI for DOIs if eligibility criteria are not met. This group is referred to as the NCEI Data Citation Working Group.
    d. When a data DOI is assigned by NCEI, NCEI shall update the metadata record to include the DOI and the recommended citation text.
        i. NOAA data DOIs shall be assigned the object type "dataset" during the DOI registration process.
        ii. When a NOAA data DOI has been assigned by NCEI, NCEI personnel shall notify relevant individual(s) or organization(s) listed as creators or providers of the data.
        iii. Any DOIs for data issued by a NOAA project under a separate license, even if that license was funded by NOAA, shall be considered project-specific DOI rather than a NOAA DOI. Data producers are strongly encouraged to use the NOAA DOI license in collaboration with NCEI.
        iv. Datasets submitted for preservation and archiving at NOAA that have previously been assigned a DOI by another organization shall be handled as described or referenced in Appendix H Part 2.
    e. DOIs shall generally be assigned at as coarse a level of granularity as possible. Data comprising many small components shall generally be assigned a single DOI rather than separate DOIs for each component.
        i. Data components that differ only in time or in space (latitude, longitude, elevation, and depth) shall be aggregated under a single DOI (e.g., a single DOI would be assigned to measurements that are archived as individual monthly files).
        ii. Datasets that are continuously updated (e.g., from automated sensors) shall be assigned a single DOI rather than broken into multiple intervals with separate DOIs.
    f. NOAA data DOIs will resolve to a human-readable landing page at an internet-accessible address which provides basic information about the data and links to additional metadata and the actual data. This landing page will reference the numeric data.
        i. DOIs shall be resolvable using the mechanism established by the International DOI Foundation.
        ii. Landing page addresses shall be updated with the DOI registrar when

data are moved.
   iii. Dataset landing pages shall be automatically produced from the metadata record.
   iv. Landing pages shall include a link to the actual metadata record.
   v. Landing pages shall include links or instructions for accessing data (unless data have been dispositioned and deleted from the archive).
   vi. Landing pages shall include guidance on how to cite the data.
g. NOAA data DOIs shall persist for as long as NCEI or successor organizations, exist.
   i. NCEI shall maintain the corresponding landing page in perpetuity.
   ii. Data that have been retired from NOAA archival repository holdings shall continue to have a landing page which shall include an explanation of the data disposition and may refer to other similar and/or related data.

3. NCEI Reserving Data DOIs
   a. Upon request, NCEI may create a DOI for data accepted for archive, but not release the DOI or data, which shall be referred to as a 'reserved DOI', as allowed by the related Data Access Section of this handbook. Additional information on reserved DOIs can be found in Appendix H Part 1.
   b. A reserved DOI and access to those data from a Federal source may be limited to no more than 12 months, as defined under Data Access.
   c. A reserved DOI and access to those data from a grantee/contract/CI source may be limited to no more than 24 months, as defined under Data Access.

Publication DOI Procedures

1. Requesting a NOAA Publication DOI
   a. The NOAA Central Library assigns DOIs to all NOAA Publications submitted to the NOAA Institutional Repository (IR) that meet the following criteria:
      i. Must be a NOAA Publication;
      ii. Must be submitted to the NOAA Institutional Repository or in the Library's electronic holdings (please see the NOAA Institutional Repository Document Policy for a full list of publications that are accepted for inclusion);
      iii. Must be a final version (drafts are not eligible for DOI assignments).
   b. Long-term preservation and documentation shall be required prior to receiving a NOAA publication DOI. A document must have sufficient bibliographic metadata, as defined by the NOAA Central Library, prior to receiving a NOAA publication DOI.
   c. Submitters may request a publication DOI prior to submission to the NOAA IR via the Library's Institutional Repository Submission form.
2. NOAA Library Publication DOI Issuance
   a. NOAA Library shall designate personnel who have the authority to issue and manage publications DOIs and to decline requests for DOIs if eligibility criteria are not met.
   b. When a NOAA publication DOI is assigned, NOAA Library shall update the

metadata record to include the DOI and the recommended citation text.

c. Publication DOIs shall be assigned the object type "text" during the DOI registration process.
d. Documents which have already received a DOI from another publisher (journal articles in particular) shall not be assigned a NOAA publication DOI.
e. Publication DOIs shall resolve to the actual document or to a landing page describing and linking to the document.
f. NOAA Central Library shall ensure the resolvability and persistence of DOIs they assign.
g. The level of granularity at which a NOAA publication DOI is assigned shall be at the discretion of NOAA Central Library. A publication DOI may be assigned at a different level of granularity than requested and shall be established in collaboration with publication authors to the extent possible.
h. Publication DOIs shall generally be assigned at as coarse a level as possible.
i. NOAA Central Library shall ensure the resolvability of publication DOIs they assign through landing pages
j. NOAA Central Library shall ensure the persistence of DOIs they assign.

3. Reserving Publication Identifiers
    a. The NOAA Central Library has begun minting DOIs to NOAA publications upon submission to the NOAA IR automatically.
    Authors/offices do not have to request a DOI prior to submission, however they are able to if they intend to embed the DOI on their publication.
    b. The NOAA Central Library may agree to create a reserved DOI for a publication that has not yet been deposited in the Library's IR.
    c. Interested parties may request a reserved DOI via the NOAA IR Submission Form.
    d. Individuals request identifiers with the intent of embedding the link within the document.
    e. Documents are submitted through normal procedures IR Submission Form, Research Publication Tracking System or NOAA Repository email at noaa.repository@noaa.gov)
    f. The NOAA Central Library reserves the right to cancel a reserved publication DOI if 120 days have elapsed since the request was initiated and the publication has not been submitted to the NOAA IR.

General questions about this Section may be sent to the NOAA Data Citation Working Group at NOAA.Data.IDs@noaa.gov.

- Specific questions about the process for obtaining a NOAA data DOI or to request a data DOI may be sent to the NCEI Data Citation Working Group at ncei.doi@noaa.gov.
- Specific questions about the process for obtaining a NOAA publication DOI or to request a publication DOI may be sent to the NOAA's Institutional Repository at noaa.repository@noaa.gov.

H. Data Access

Purpose

The goal of data access is to ensure that all NOAA environmental data and associated documentation are freely available, publicly accessible in a timely fashion, and reusable with an open license, as required by the Evidence Act and the *Office of Science and Technology Policy (OSTP) Memorandum on "Increasing Access to the Results of Federally Funded Scientific Research"* (PARR 2013, and see below regarding PARR 2022) and other federal laws, orders, and guidance relevant to public data.

Data Management Requirements

Many of the requirements for data access are described in the above Sections of this document. Approved submission of data to the NOAA NCEI, and provision of public access by NCEI, shall be considered to meet all the provisions of this requirement for discoverability. Data not submitted to NCEI, or not accepted by NCEI, shall independently satisfy the requirements below.

Existing LO directives and policies for provision of data and products must be adhered to prior to making the data accessible. For example, the National Weather Service (NWS) has a policy that public comment/review is required prior to making new data/products available.

Recognizing that some data may not be made accessible for legal or technical reasons, any exceptions should be rare. Decisions to withhold access must be explicit and documented in the DMP.

General Requirements for Data Access

Unique requirements for data access not explicitly described in above Sections of this document are summarized here; redundant requirements are omitted. These are based on the Open Data Policy 2013, Evidence Act 2018, and Geospatial Data Act 2018.

- General data access requirements:
    - Machine-readable
    - Publicly available
    - In open, non-proprietary, formats
    - Free of charge, or at no more than the cost of reproduction when physical media or special delivery mechanisms are required
    - Under an open license
    - Without restrictions that would impede use or reuse
        - Where a log in would improve service delivery or is required for security purposes, a common Government system such as Login.gov is used.

    - Using the open data and interoperability standards established and maintained in the DGC Standards, Services, Platforms, Tools, and Software (SSPTS) Catalog.
    - Data dissemination will be done in accordance with the NOAA Dissemination Task Team Report.
- At least one of the following machine-readable data access methods shall be provided:
    - Data shall preferably be made available via a web service or API that supports machine-to-machine data access and enables users to request the desired subset of the data.
    - Data may be made available via bulk download via File Transfer Protocol (FTP) or Hypertext Transport Protocol (HTTP) in addition to Web service(s). However, bulk download as the only method of access is discouraged. Websites, web pages, or portals requiring human intervention to operate do not qualify as meeting the machine-readable access requirement, but may be provided in addition to machine-readable methods.
    - Archived data stored on robotic tape drive may be made available via asynchronous ordering service.
- Geospatial-specific data requirements
    - Geospatial datasets that are or contribute to a NGDA data themes should be identified as such, meet NGDA standards, and be made publicly discoverable via GeoPlatform.gov. Geospatial datasets that do not directly contribute to NGDA themes should also meet those standards to the extent practicable.
- Restricted data
    - Datasets restricted from public access due to law, privacy, confidentiality pledge, security, trade secret, contractual, or other valid restrictions (such as being under development) need to have the reason access is restricted and how access is restricted specified in the DMP and reviewed and approved by the ACDO.
- Data shall be accessible until one of the following conditions has been met
    - Data have passed their retention schedule, have been dispositioned according to NARA rules, and are deemed of no future value to NOAA or other potential users.
    - A DMP has been filed and accepted by the ACDO explaining the reasons for eliminating or reducing accessibility. Data not archived at NCEI is also governed by a retention schedule that needs to be followed for accessibility.
    - A new version of the data has been made available which supersedes the older version.

Timeliness Requirements for Data Access

Generally, NOAA data should be made publicly available with minimal delay, but timely access will not be the same in all cases and will vary between funding mechanisms and acquisition approaches. More recently, PARR 2022 explicitly defines timeliness as: **data underlying peer-reviewed scholarly publications are made publicly accessible at the time of publication.** While this will not be a requirement until (no later than) December 31, 2027, it is

recommended that programs start working towards this goal now and meet it whenever possible.

The DMP will document the data access schedule following the current timeliness requirements:

- NOAA-produced data should be made publicly available either:
    - at the time of publication of a peer-reviewed article based on the data
    - with no delay beyond latency imposed by funding dependencies, data processing, transmission, and possibly archive ingest
    - not to exceed one year after initial manual data collection, processing, or quality control (unless approved in a DMP).
- Extramural (Data obtained via NOAA-funded grants, cooperative agreements, or contracts) should be made publicly available either:
    - at the time of publication of a peer-reviewed article based on the data
    - two years after the data are collected and verified
    - two years after the original end date of the grant (not including any extensions or follow-on funding)
- Holding data by the Data Producer for any reason other than listed above (i.e., withholding data access solely for the purpose of being the first to publish) is not permitted for any NOAA funded data.

Extramural Requirements for Data Access

Many of the requirements for data access apply to all NOAA data, including datasets generated by both NOAA Programs and by NOAA-funded grants, contracts, and cooperative agreements (extramural data). NOAA Programs should plan to adjust *requirements documents* per a revised *NOAA PARR Plan*, currently to be published by December 31, 2024, with an effective date TBD but no later than December 31, 2027. The PARR 2022 requirement with respect to timeliness only addresses data associated with publications, and that timeliness of other data is not dependent on the PARR effective dates

Until that time:
NOAA Programs shall strive to ensure that environmental data produced as a result of NOAA-funded Grants, Cooperative Agreements, or Contracts are made publicly accessible, in a timely fashion (typically within 2 years), free of charge or at no more than the cost of reproduction, and that any exceptions or extensions are explicitly justified on a case-by-case basis. Data sharing could be limited by law, regulation, policy, security requirements, commercial or international agreements, or valid technical considerations.

- Approved submission of data to NCEI, or to a publicly-accessible data repository approved by the funding program, shall be sufficient to satisfy the data accessibility requirement as of the date the submission was performed.
- Data accessibility shall be confirmed, during or before the close-out phase if possible, by (a) verifying that an online link for data access was reported by the funding recipient and (b) checking such links to confirm the presence of expected data.

- NOAA Programs shall have flexibility to authorize exceptions or delays to data accessibility on a case-by-case basis with valid justification; these instances shall be reported to DGC.
- Submission of manuscripts to the NOAA Institutional Repository shall be confirmed, either as part of normal reporting by the Repository or by Programs querying the Repository to verify the existence of manuscripts.
- NOAA programs may use existing legal methods (e.g., as defined in the Department of Commerce Grants and Cooperative Agreements Manual) in the case of non-compliance by grantees.
- Follow guidelines as outlined in Appendix D - Extramural Data Sharing Guidance
- Data shall be publicly accessible no later than publication of a peer-reviewed article based on the data, or two years after the data are collected and verified, or two years after the original end date of the grant (not including any extensions or follow-on funding), whichever is soonest, unless a delay has been authorized by the NOAA funding program.

Performance Objectives and Measurements

Objective H.1: Ensure that all NOAA environmental data and associated documentation are freely available, publicly accessible in a timely fashion, and reusable.

Metric H.1.1: Percentage of datasets that are accessible by the public.

Metric H.1.2 : Percentage of datasets that are restricted.

Objective H.2: Ensure NOAA-generated (both NOAA and extramural) datasets meet federal requirements for scientific data.

Objective H.3: Ensure recipients of NOAA extramural funding meet federal requirements for scientific data if they intend to generate environmental data.

Procedures

Note: Provision of public access by the NOAA NCEI shall be considered to meet all the provisions of requirements for data access. Data not made accessible by NCEI shall independently satisfy the data access requirements.

1. Planning for data access
    a. Create a DMP *(required)*
    b. Ensure the DMP includes **data access approach** (target access provider(s)) and **data access schedule(s)** *(required)*
        i. If the *DMP* specifies access limitations and/or exceptions ensure this is approved by the ACDO; *(required if applicable)*
    c. Ensure identified data points of contact and authors have Open Researcher and Contributor Identification (ORCID) *(encouraged)*
2. Documentation for data access
    a. Understand the documentation standard(s) and format(s) required by the target access provider (*i.e.*, data repository(ies) and/or data access service(s))

(recommended)

b. Select a documentation tool appropriate for the target access provider, see Appendix F Part 2 for available NOAA metadata tools. *(recommended)*
c. If possible, create and maintain native documentation (metadata, ancillary files, and/or data dictionaries, *etc*.) to meet the target access provider documentation requirements.*(recommended)*
d. If necessary, transform native documentation to the format required by target access provider *(recommended)*
e. Include individual ORCIDs in documentation *(encouraged)*

3. Prepare data package (including documentation)
    a. In advance of data package completion (ideally at least one year), confirm requirements and procedures for target access provider *(recommended)*
    b. Select a data submission tool appropriate for the target access provider, see Appendix I Data Access Details, Part 1 for suggestions *(recommended)*
    c. Complete quality control/quality assurance (QA/QC) protocols, if any, and any other analyses to meet standards for scientific integrity should applied and documented appropriately in the metadata *(required)*
    d. Confirm that the data package does not contain any information that is restricted from public access due to law, privacy, confidentiality pledge, security, trade secret, contractual, or other valid restrictions. If it does, updated the DMP and submit for ACDO approval *(required)*
    e. Coordinate with authors of associated peer-reviewed scholarly publications to ensure data access date(s) and publication date(s) are aligned *(encouraged)*
        i. If known, include publication DOI(s) in data documentation; for NOAA Technical Memoranda/Reports, contact the NOAA Institutional Repository to reserve a publication DOI *(recommended)*
    f. Ensure the data package includes, at a minimum, the data underlying the publication(s) that would be needed by an independent researcher to reproduce the study *(recommended)*
    g. Prepare a test package of a subset of data to ensure all requirements are met; share test package with target access provider if appropriate *(encouraged)*
4. Obtain approval to submit to target access provider
    a. At a minimum, first line supervisors should provide review and approval of data packages before submission *(encouraged)*
5. Submission to target data access provider, see Appendix I Data Access Details, Part 1 for suggestions. *(required)*
6. Next steps:
    a. Update metadata (i.e., InPort) with data access URL (i.e., ERDDAP) and data DOI (if any) *(required)*
    b. Update associated publication with dataset citation (including data DOI, if applicable) and re-submit publication to NOAA Institutional Repository *(recommended)*

Chapter 4 - Data Management for Administrative Data

This is a placeholder for further development of the requirements, metrics, and procedures for Administrative Data.

Examples of data that may be included in the Chapter are:

- Access statistics, forecast performance metrics, and other business process information that are about data but do not themselves constitute environmental data.
- Responses to requests for agency records under the FOIA, the Privacy Act, the Federal Advisory Committee Act or other similar law.
- Press releases, fact sheets, press conferences or similar communications in any medium that announce, support the announcement or give public notice of information NOAA has made publicly available elsewhere. Reference records, including library holdings and World Data Center holdings.
- Public filings.
- Responses to subpoenas or compulsory document productions.
- Requirements documents (*scopes of work, performance work statements, statements of objectives, Federal Funding Opportunities, etc.*).
- Policy manuals and management information produced for the internal management and operations of NOAA, and not primarily intended for public dissemination.
- Information presented to Congress as part of legislative or oversight processes, such as testimony of NOAA officials, and information or drafting assistance provided to Congress in connection with proposed or pending legislation, that is not simultaneously disseminated to the public. (However, information, which would otherwise be covered by applicable guidelines, is not exempted from compliance merely because it is also presented to Congress.)
- Information limited to adjudicative processes, such as pleadings, including information developed during the conduct of any criminal or civil action or administrative enforcement action, investigation or audit against specific parties, or information distributed in documents limited to administrative action determining the rights and liabilities of specific parties under applicable statutes and regulations.

Appendix A - Reference Documents

Connelly, R., Playford, C. J., Gayle, V., & Dibben, C. (2016). The role of administrative data in the big data revolution in social science research. Social science research, 59, 1-12.

Consultative Committee for Space Data Systems, “Producer-Archive Interface Methodology Abstract Standard,” CCDS 651.01-B-1, Blue Book, Feb 2014.
https://public.ccsds.org/pubs/651x1b1.pdf

Consultative Committee for Space Data Systems, “Producer-Archive Interface Methodology Abstract Standard,” CCDS 651.0-B-1, Blue Book, May 2004.
http://public.ccsds.org/publications/archive/651x0b1.pdf

Elias, P. (2014). Administrative data. Facing the Future: European Research Infrastructures for the Humanities and Social Sciences. Berlin: Scivero Verlag, 47.

Federal Advisory Committee Act
https://www.gsa.gov/policy-regulations/policy/federal-advisory-committee-management/legislation-and-regulations/the-federal-advisory-committee-act

Federal Data Strategy
https://strategy.data.gov/

Federal Data Strategy Curated Data Skills Catalog
https://resources.data.gov/assets/documents/fds-data-skills-catalog.pdf

Foundations for Evidence-Based Policymaking Act of 2018
https://www.congress.gov/bill/115th-congress/house-bill/4174

International Standard (ISO 15489-1) effective September 15, 2001 entitled “Information and documentation -- Records management.”
https://www.iso.org/standard/62542.html

Mc Grath-Lone, L., Jay, M. A., Blackburn, R., Gordon, E., Zylbersztejn, A., Wijlaars, L., & Gilbert, R. (2022). What makes administrative data “research-ready”? A systematic review and thematic analysis of published literature. International journal of population data science, 7(1).

National Archives and Records Administration, Subchapter B - Records Management (36 CFR 1220-1238).
https://www.ecfr.gov/current/title-36/chapter-XII/subchapter-B

National Archives and Records Administration. (2007, September 20). *Appraisal policy of the National Archives*. National Archives and Records Administration.

http://www.archives.gov/records-mgmt/initiatives/appraisal.html

NOAA 2020 Data Strategy
https://sciencecouncil.noaa.gov/wp-content/uploads/2022/08/2020-Data-Strategy.pdf

NOAA Administrative Order (NAO) 212-15A effective December 2, 2008: “Management of Environmental and Geospatial Data and Information.”
http://www.corporateservices.noaa.gov/~ames/NAOs/Chap_212/naos_212_15.html

NOAA Administrative Order (NAO) 202-735D-2: Scientific Integrity
https://www.noaa.gov/organization/administration/nao-202-735d-2-scientific-integrity#:~:text=The%20purpose%20of%20this%20National,inform%20management%20and%20policy%20decisions.

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

NOAA Fiscal Year 2022-2026 Strategic Plan
https://www.noaa.gov/organization/budget-finance-performance/value-to-society/noaa-fy22-26-strategic-plan

NOAA Information Quality Guidelines
https://www.noaa.gov/organization/information-technology/policy-oversight/information-quality/information-quality-guidelines

NOAA NESDIS Data Lifecycle
https://www.nesdis.noaa.gov/news/nesdis-data-lifecycle

NOAA Policy on Partnership in the Provision of Environmental Information (19 January 2006).
http://www.noaa.gov/partnershippolicy

NOAA Records Disposition Handbook
http://www.ofa.noaa.gov/~ames/Records_Management/disposition_handbook.html

OMB Circular A-130, Management Information as a Strategic Resource
https://www.cio.gov/policies-and-priorities/circular-a-130/

Renear, Allen H., Simone Sachi, Karen Wickett. Definition of dataset in the scientific and technical literature. Proceedings of the American Society for Information Science and Technology. Volume 47, Issue 1. Feb 2011.
https://doi.org/10.1002/meet.14504701240

The Paperwork Reduction Act of 1995 (44 U.S.C. 3501 et seq.).
https://digital.gov/resources/paperwork-reduction-act-44-u-s-c-3501-et-seq/

The United States Geological Survey Science Data Lifecycle Model
https://pubs.usgs.gov/of/2013/1265/pdf/of2013-1265.pdf

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 16–18 (2016).
https://doi.org/10.1038/sdata.2016.18

Woollard, M. (2014). 3.1 administrative data: Problems and benefits. a perspective from the united kingdom1. Facing the future: European research infrastructures for the humanities and social sciences, 49.

Appendix B – Acronym List

Acronym	Definition
ACDO	Assistant Chief Data Officers
API	Application Programming Interface
ATRAC	Advanced Tracking and Resource tool for Archive Collections
CARE	Collective Benefit, Authority to Control, Responsibility, and Ethics
CDO	Chief Data Officer
CUI	Controlled Unclassified Data
DGC	Data Governance Committee
DM	Data Management
DMP	Data Management Plan
DOC	Department of Commerce
DOI	Digital Object Identifier
EDMC	Environmental Data Management Committee
FAIR	Findable, Accessible, Interoperable, and Reusable
FFO	Federal Funding Opportunity
FGDC	Federal Geographic Data Committee
FOIA	Freedom of Information Act
FTP	File Transfer Protocol
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protocol
IR	Institutional Repository
ISO	International Organization for Standardization
LO	Line Office
NAO	NOAA Administrative Order

Acronym	Definition
NARA	National Archives and Records Administration
NCEI	National Centers for Environmental Information
NESDIS	National Environmental Satellite, Data, and Information Service
NGDA	National Geospatial Data Asset
NMFS	National Marine Fisheries Service
NOAA	National Oceanic and Atmospheric Administration
NODD	NOAA Open Data Dissemination
NOS	National Ocean Service
NOSC	NOAA Observing Systems Council
NRC	National Research Council
NWS	National Weather Service
OAIS	Open Archival Information System
OAP	Ocean Acidification Program
OAR	Ocean and Atmospheric Research
ORCID	Open Researcher and Contributor Identification
OMB	Office of Management and Budget
OSC	Observing Systems Committee
OSTP	Office of Science and Technology Policy
PARR	Public Access to Research Results
PID	Persistent Unique Identifier
PII	Personal Identifiable Information
QA	Quality Assurance
QC	Quality Control
RM	Reference Model

Acronym	Definition
S2N	Send2NCEI
SDIS	Scientific Data Information System
TBD	To Be Determined
URL	Uniform Resource Locator
USGS	United States Geological Survey
XML	Extensible Markup Language

Appendix C - Data Management Plan Template

Please provide the following information to the LO ACDO for submission to the NOAA DMP Repository.[12]

Reference to Master Data Management Plan (if applicable)

As stated in Chapter 3 Section E, DMPs may be hierarchical. If this DMP inherits provisions from a higher-level DMP already submitted to the Repository, then this more-specific Plan only needs to provide information that differs from what was provided in the Master DMP.

DMPs covering multiple timeframes, geographies, data types, collection methods, etc., are encouraged to use lists, tables, figures, etc., to organize requested information in a manner most appropriate from the complexity of the data.

URL of higher-level DMP (if any) as submitted to DMP Repository:

1. **General Description of Data to be Acquired and Managed**
   - 1.1. Name of the Data, data collection Project, or data-producing Program:
   - 1.2. Project Purpose and Abstract including summary description of the data:
   - 1.3. Is this a one-time data collection, or an ongoing series of measurements?
   - 1.4. Actual or planned temporal coverage of the data:
   - 1.5. Actual or planned geographic coverage of the data:
   - 1.6. Is there a data acquisition plan? Does acquisition include restricted designations?
   - 1.7. Type(s) of data and approximate data volume:
     (e.g., digital numeric data, imagery, photographs, video, audio, database, tabular data, etc.)
   - 1.8. Data collection method(s):
     (e.g., satellite, airplane, unmanned aerial system, radar, weather station, moored buoy, research vessel, autonomous underwater vehicle, animal tagging, manual surveys, enforcement activities, numerical model, etc.)
     - 1.8.1. Collection platform, if known:
   - 1.9. If data are from a NOAA Observing System of Record, indicate name of system:
     - 1.9.1. If data are from another observing system, please specify:
2. **Point of Contact for this Data Management Plan (author or maintainer)**
   - 2.1. Name:
   - 2.2. Title:
   - 2.3. Affiliation or facility:
   - 2.4. E-mail address:

2.5. Phone number:

3. **Responsible Party for Data Management**
Program Managers, or their designee, shall be responsible for ensuring the proper management of the data produced by their Program. Please indicate the responsible party below.

3.1. Position Title:

3.2. Name of current position holder:

3.3. Email of current position holder

4. **Resources**
Programs must identify resources within their own budget for managing the data they produce.

4.1. Are there the appropriate level of resources available to manage these datasets throughout the data life cycle?

4.2. If the data are to be publicly accessible, are there the appropriate level of resources available to disseminate (make publicly available) these datasets?

4.3. Are there the appropriate level of resources available to archive (submit to an approved data repository for long term preservation) these datasets? If so, where do you plan to archive these data?

4.4. What is the approximate percentage of the overall project budget for these data devoted to data management to ensure compliance with data management requirements (specify percentage or "unknown"):

5. **Data Lineage and Quality**
NOAA has issued Information Quality Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information which it disseminates.

5.1. Processing workflow of the data from collection or acquisition to making it publicly accessible *(describe or provide URL of description)*:

5.1.1. If data at different stages of the workflow, or products derived from these data, are subject to a separate data management plan, provide reference to other plan and provide data source details:

5.2. Quality control procedures employed *(describe or provide URL of description)*:

6. **Data Documentation**
The Data Documentation Procedure requires that NOAA data be well documented, specifies the use of ISO 19115-2: 2009 and related standards for documentation of new data, and provides links to resources and tools for metadata creation and validation.

6.1. Does metadata comply with the Data Documentation requirements?

6.1.1. If metadata are non-existent or non-compliant, please explain:

6.2. Name of organization or facility providing metadata hosting:

6.2.1. If service is needed for metadata hosting, please indicate:

6.3. URL of metadata folder or data catalog, if known:

6.4. Process for producing and maintaining metadata *(describe or provide URL of description)*:

7. **Data Access**

NAO 212-15 states that access to environmental data may only be restricted when distribution is explicitly limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements. This restriction would include PII and other sensitive data (export controlled data) and data restricted by contract or other written, binding agreement (permitted to be withheld under the Evidence Act) including commercial data licensed via contract, data obtained from another third party subject to a restrictive license (international partner, CRADA, etc.).The Data Access section of this Handbook contains specific guidance, recommends the use of open-standard, interoperable, non-proprietary web services, defines timeliness, provides information about resources and tools to enable data access.

7.1. Do these data comply with the general data access requirements?

7.2. Intended data access method(s):

(Specify Web Service; API; FTP Bulk Download; HTTP Bulk Download; Website, web page, or portal; Asynchronous Ordering Service; To Be Determined; Unable to Provide Access; Limited Access Only; or No Access Intended)

7.2.1. Is the data, in part or in whole, restricted to the public: (yes or no)?

7.2.2. If Restricted,To Be Determined, or Unable to Provide Access, please explain:

7.3. Name of organization of facility providing data access:

7.3.1. If data hosting service is needed, please indicate URL of data access service or other methods, if known.

7.4. Tentative dissemination date by which data will be made publicly available per timeliness requirement:

(Specify [Exact Date]; No Delay; one year from Collection; one year from QA/QC; or

2 years from QA/QC; Exceeds 2 years from QA/QC; To Be Determined; Unable to Provide Access; Limited Access Only; or No Access Intended)

8. **Data Preservation and Protection**

The NOAA Procedure for Scientific Records Appraisal and Archive Approval[14] describes how to identify, appraise and decide what scientific records are to be preserved in a NOAA archive.

8.1. What is the actual or planned long term data archive location:
(Specify NCEI facility, or other specific archive, To Be Determined, Unable to Archive, or No Archiving Intended)

8.1.1. If data preservation is to be provided by a data repository other than NOAA NCEI, please explain (e.g., data out-of-scope for NCEI, appraisal failed, preferred community of practice, cost considerations, technical considerations, etc.)

8.1.2. If To Be Determined, Unable to Archive or No Archiving Intended, please clarify that selection:

8.2. Data storage facility prior to being sent to an archive facility (if any):

8.3. Approximate delay between data collection and submission to an archive facility:

8.4. How will the data be protected from accidental or malicious modification or deletion prior to receipt by the archive? Discuss data back-up, disaster recovery/contingency planning, and off-site data storage relevant to the data collection:

9. **Additional Line Office or Staff Office Questions**
Line and Staff Offices may extend this template by inserting additional questions in this section.

Approval:

_____	_____
Name	Title
_____	_____
Signature	Date

Appendix D - Extramural Data Sharing Guidance

This Handbook specifies requirements for NOAA Programs that issue grants, cooperative agreements, or contracts. In particular, NOAA Programs are required to consider, in advance, how to ensure public accessibility and long-term preservation of externally-funded data, to provide guidance for proposers to use in developing a plan for data access, and to track and enforce conditions imposed on awardees. These requirements do not apply to FFO Announcements and Contract Solicitations published before June 1, 2016, or to Grants, Cooperative Agreements, and Contracts funded as a result of Announcements published before that date. The previous version of this directive applies to earlier grants.

Part 1 - Data Management Guidance Template for NOAA Funding Programs

High-level Data Management Guidance (DM Guidance) based on the template below shall be produced by NOAA Programs seeking to award Grants, Cooperative Agreements, or Contracts anticipated to result in the production of environmental data. Programs may tailor the guidance to suit their needs. The DM Guidance shall be included or referenced in the NOAA FFO Announcement or Contract Solicitation. The same Guidance may be used for multiple Announcements or Solicitations as long as the Guidance is applicable to each one.

NOTE: If the scope of the Announcement or Solicitation does not permit the creation of data, then this section does not apply and this guidance is not required. Instead, include the text referenced in PART 2 of this Appendix for projects not expected to generate environmental data in the Announcement. This directive does apply to Broad Agency Announcements, which may generate proposals that would create, collect, or acquire data or peer-reviewed publications and the text reference in PART 2 should be included in the Announcement.

Data Management Guidance to Proposal Writers

1. **Responsible NOAA Official** for questions regarding this guidance and for verifying accessibility of data produced by funding recipients:
    - Name:
    - Title:
    - Affiliation or facility:
    - E-mail address:
    - Phone number:
2. **Data Accessibility:** The NOAA Program recommends (or requires) that public access to grant/contract-produced data be enabled as follows *(one or more of the following, or alternate text as appropriate, can be provided as guidance; Data Management Plans submitted with Proposals should reflect the option(s) provided by the Program)*
    - Funding recipients are expected to submit data to NOAA NCEI, which will

provide public access and permanent archiving.
    - The NOAA Program has held preliminary consultation with NCEI regarding these pending data.
- Data are to be submitted to the following International Council for Science World Data System facility: _____.
  (see list at https://worlddatasystem.org/members/member_directory)
- The following NOAA facility *(other than NCEI)* will operate a publicly accessible online data server for these data: _____.
- An existing publicly accessible online data server at the funded institution is to be used to host these data (describe in proposal).
- Data are to be submitted to a public data repository appropriate to this scientific domain (describe in proposal). *(Options could include Dryad - http://datadryad.org, Figshare - http://figshare.com, Dataverse http://dataverse.org, Pangaea - http://www.pangaea.de, Acadis (for Arctic data) - https://www.aoncadis.org, Merritt (for University of California) - https://merritt.cdlib.org, or others.)*
- Funding recipients will establish their own data hosting capability (describe in proposal).
- Proposal may request permission not to make data publicly accessible (proposal to explain rationale for lack of public access, and if funded to obtain approval from Responsible NOAA Official listed above).

3. **Technical recommendations** *(one or more of the following, or alternate text as appropriate, can be provided as additional guidance; Data Management Plans submitted with Proposals should reflect the option(s) provided by the Program)*
    - The NOAA Program recommends (or requires) the following data format(s), data access method(s), or other technical guidance:_____
    - The NOAA Program is not offering specific technical guidance. Proposals are to describe their proposed approach. Use of open-standard formats and methods is encouraged.
4. **Resources** *(NOTE: NCEI may charge a fee for archiving, particularly for large or unusual datasets; NOAA Programs are advised to contact NCEI in advance)*
    - Proposals are permitted to include the costs of data sharing or archiving in their budgets.
    - NOAA Program resources for data sharing or archiving have already been identified; proposals should not include such costs.
        i. Proposals may include such costs if data volume is expected to exceed _____ .

Programs may also consider sharing the NOAA Data Management Plan in Appendix C as an example.

Part 2 - Text to be included in Announcements and Awards

The following text is for inclusion in FFO Announcements and Contract Solicitations and Notices of Award and Contracts.

Text to be included in Announcements and Solicitations for projects NOT expected to generate environmental data

This announcement is not seeking proposals that generate environmental data. Therefore, a Data Management Plan is not required as part of the Proposal.

Text to be included in FFO Announcements and Contract Solicitations for projects that may generate environmental data (including Broad Agency Announcements)

1. Environmental data and information collected or created under NOAA grants or cooperative agreements must be made discoverable by and accessible to the general public, in a timely fashion (typically within two years), free of charge or at no more than the cost of reproduction, unless an exemption is granted by the NOAA Program. Data should be available in at least one machine-readable format, preferably a widely-used or open-standard format, and should also be accompanied by machine-readable documentation (metadata), preferably based on widely- used or international standards.
2. Proposals submitted in response to this Announcement must include a DMP of up to two pages describing how these requirements will be satisfied. The DMP should be aligned with the DM Guidance provided by NOAA in the Announcement. The contents of the Data Management Plan (or absence thereof), and past performance regarding such plans, will be considered as part of proposal review. A typical plan should include descriptions of the types of environmental data and information expected to be created during the course of the project; the tentative date by which data will be shared; the standards to be used for data/metadata format and content; methods for providing data access; approximate total volume of data to be collected; and prior experience in making such data accessible. The costs of data preparation, accessibility, or archiving may be included in the proposal budget unless otherwise stated in the Guidance. Accepted submission of data to the NOAA NCEI is one way to satisfy data sharing requirements; however, NCEI is not obligated to accept all submissions and may charge a fee, particularly for large or unusual datasets.
3. NOAA may, at its own discretion, make publicly visible the DMP from funded proposals, or use information from the DMP to produce a formal metadata record and include that metadata in a Catalog to indicate the pending availability of new data.
4. Proposal submitters are hereby advised that the final pre-publication manuscripts of scholarly articles produced entirely or primarily with NOAA funding will be required to be submitted to NOAA Institutional Repository after acceptance, and no later than upon publication. Such manuscripts shall be made publicly available by NOAA one year after publication by the journal.

Text to be included in Notices of Award and Contracts of projects anticipated to generate environmental data or peer-reviewed publications

Special Award Conditions:

Data Sharing: Environmental data collected or created under this Grant, Cooperative Agreement, or Contract must be made publicly visible and accessible in a timely manner, free of charge or at minimal cost that is no more than the cost of distribution to the user, except where limited by law, regulation, policy, or national security requirements. Data are to be made available in a form that would permit further analysis or reuse: data should be encoded in a machine-readable format, preferably using existing open format standards; data must be sufficiently documented, preferably using open metadata standards, to enable users to independently read and understand the data. The location (internet address) of the data should be included in the final report. Pursuant to NOAA Information Quality Guidelines, data should undergo quality control (QC) and a description of the QC process and results should be referenced in the metadata. Failure to perform quality control does not constitute an excuse not to share data. Data without QC are considered "experimental products" and their dissemination must be accompanied by explicit limitations on their quality or by an indicated degree of uncertainty.

1. **Timeliness:** Data accessibility must occur no later than publication of a peer-reviewed article based on the data, or two years after the data are collected and verified, or two years after the original end date of the grant (not including any extensions or follow-on funding), whichever is soonest, unless a delay has been authorized by the NOAA funding program.
2. **Disclaimer:** Data produced under this award and made available to the public must be accompanied by the following statement: "These data and related items of information have not been formally disseminated by NOAA, and do not represent any agency determination, view, or policy."
3. **Failure to Share Data:** Failing or delaying to make environmental data accessible in accordance with the submitted DMP, unless authorized by the NOAA Program, may lead to enforcement actions, and will be considered by NOAA when making future award decisions. Funding recipients are responsible for ensuring these conditions are also met by sub-recipients and subcontractors.
4. **Funding acknowledgement:** Federal funding sources shall be identified in all scholarly publications. An Acknowledgements section shall be included in the body of the publication stating the relevant Grant Programs and Award Numbers. In addition, funding sources shall be reported during the publication submission process using the FundRef mechanism (http://www.crossref.org/fundref) if supported by the Publisher.
5. **Manuscript submission:** The final pre-publication manuscripts of scholarly publications produced with NOAA funding shall be submitted to the NOAA Institutional Repository at https://repository.library.noaa.gov after acceptance, and no later than upon publication, of the paper by a journal. NOAA will produce a publicly-visible catalog entry directing users to the published version of the article. After an embargo period of one year after publication, NOAA shall make the manuscript itself publicly visible, free of charge, while continuing to direct users to the published version of

record.

6. **Data Citation:** Publications based on data, and new products derived from source data, must cite the data used according to the conventions of the Publisher, using unambiguous labels such as Digital Object Identifiers. All data and derived products that are used to support the conclusions of a peer-reviewed publication must be made available in a form that permits verification and reproducibility of the results.

Appendix E - Data Licensing Requirements

CASE 1: Internal NOAA Source Data
NOAA adopted the well-known and internationally recognized licenses from Creative Commons, consistent with OPEN Government Data Act and the Federal Data Strategy guidance. All Internal NOAA Source Data that are appropriate for public release (not classified or controlled unclassified information) 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. This license ensures maximum use of our data, to spur and encourage exploration and innovation throughout the industry.

Data refers to recorded information, regardless of form or the media on which the data are recorded. This includes data and products regardless of the level of processing.

Internal NOAA Source Data refers to data that are generated by NOAA-owned sensors or systems or NOAA federal employees, including:

- NESDIS satellite data and data products, climatologies, and atlases
- NWS radar data and model output
- NMFS habitat surveys, coral ecosystems analyses
- OAR carbon mooring arrays, atmospheric chemistry monitoring stations
- NOS bathymetry data collected by NOAA ships
- Derived products developed by NOAA systems or NOAA federal employees using internal and/or external data, when the agreement with the external data provider confirms that they do not maintain ownership of or restrict the use of derived products created with their data (e.g., NWS model output)

License field in Metadata: CC0 1.0

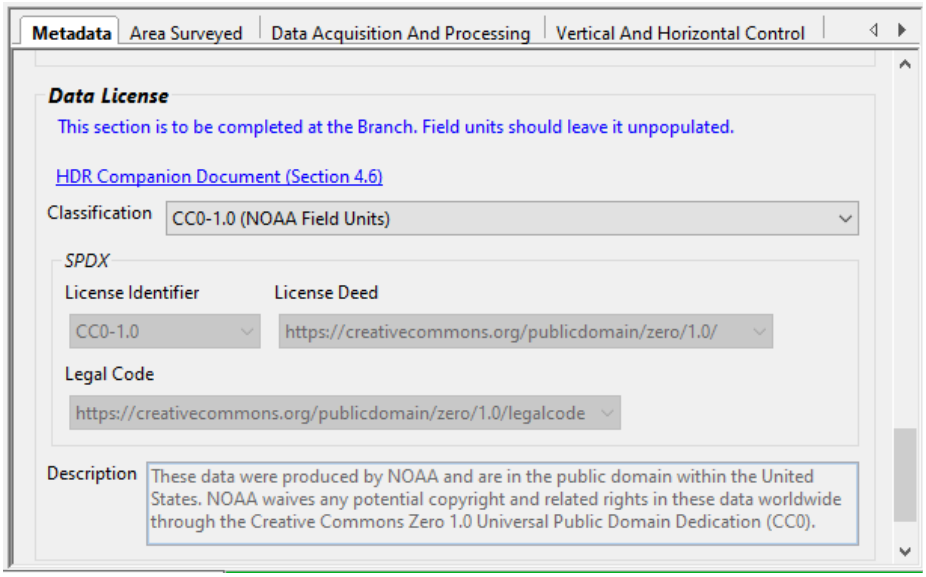
Full Text for Landing Pages: These data were produced by NOAA and are not subject to copyright protection in the United States. NOAA waives any potential copyright and related rights in these data worldwide through the Creative Commons Zero 1.0 Universal Public Domain Dedication (CC0 1.0).

Best Practices for Implementation:

Metadata

When creating metadata for NOAA internal data, the following statement should be used to identify the use of the CC0-1.0 license: "These data were produced by NOAA and are not subject to copyright protection in the United States. NOAA waives any potential copyright and related rights in these data worldwide through the Creative Commons Zero 1.0 Universal Public Domain Dedication (CC0-1.0)." The Best Practices for Implementation is based off the NOAA Open Data Licensing - Legislative, Policy, Directives and Implementation Guidance.

An example from the NOAA Office of Coast Survey of incorporating CC0-1.0 into metadata xml form is shown below. In a dropdown box (entitled “Classification”), the “CC0-1.0” option is selected (as is mandated for NOAA hydrographic field units). The license identifier (CC0-1.0) is populated automatically, as are the License Deed and Legal Code URLs from Creative Commons (although these URLs are not required to be included). Additionally, the Description field is populated with appropriate legal text.



Geospatial Data

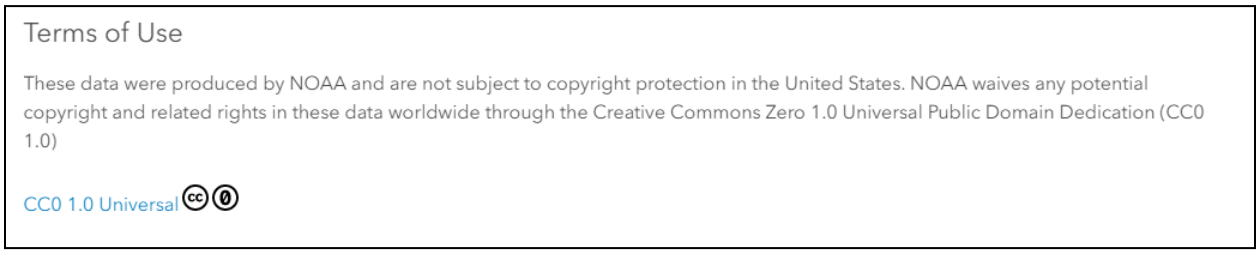
Users publishing NOAA data in the NOAA GeoPlatform (ArcGIS Online) should cite this data licensing information in the "Terms of Use" section. In order to preserve consistent formatting, please use the HyperText Markup Language (HTML) snippet below that can be added to the item in the NOAA GeoPlatform.

HTML to include in the “Terms of Use” Section:

<p>These data were produced by NOAA and are not subject to copyright protection in the United States. NOAA waives any potential copyright and related rights in these data worldwide

through the Creative Commons Zero 1.0 Universal Public Domain Dedication (CC0 1.0)</p>CC0 1.0 Universal

Example NOAA GeoPlatform Content Item with HTML snippet included:
https://noaa.maps.arcgis.com/home/item.html?id=3df790fa23984682b8a359eba1c87c0c



Example Implementation: NOAA Office of Coast Survey Data Licensing Website

The NOAA Office of Coast Survey adopted the well-known and internationally recognized licenses from Creative Commons. All data acquired by Coast Survey is formally dedicated to the public domain via the Creative Commons 1.0 Universal Public Domain Dedication (CC0), which removes all copyright from the data so that it may be used by anyone, for any purpose. This license ensures maximum use of our data, to spur and encourage exploration and innovation throughout the industry. For more details, please visit:
https://nauticalcharts.noaa.gov/data/data-licensing.html

To be added at a future date by the Data Licensing Task Team:
[CASE 2: External source data with CC BY 4.0 International]
[CASE 3: Sources that require a different type of license than CASE 1 or CASE 2]

Data Licensing Flowchart

The Data Governance Committee Data Licensing Task Team developed the following flowchart to represent the decision process related to deciding which license may be most appropriate for a specific kind of data.

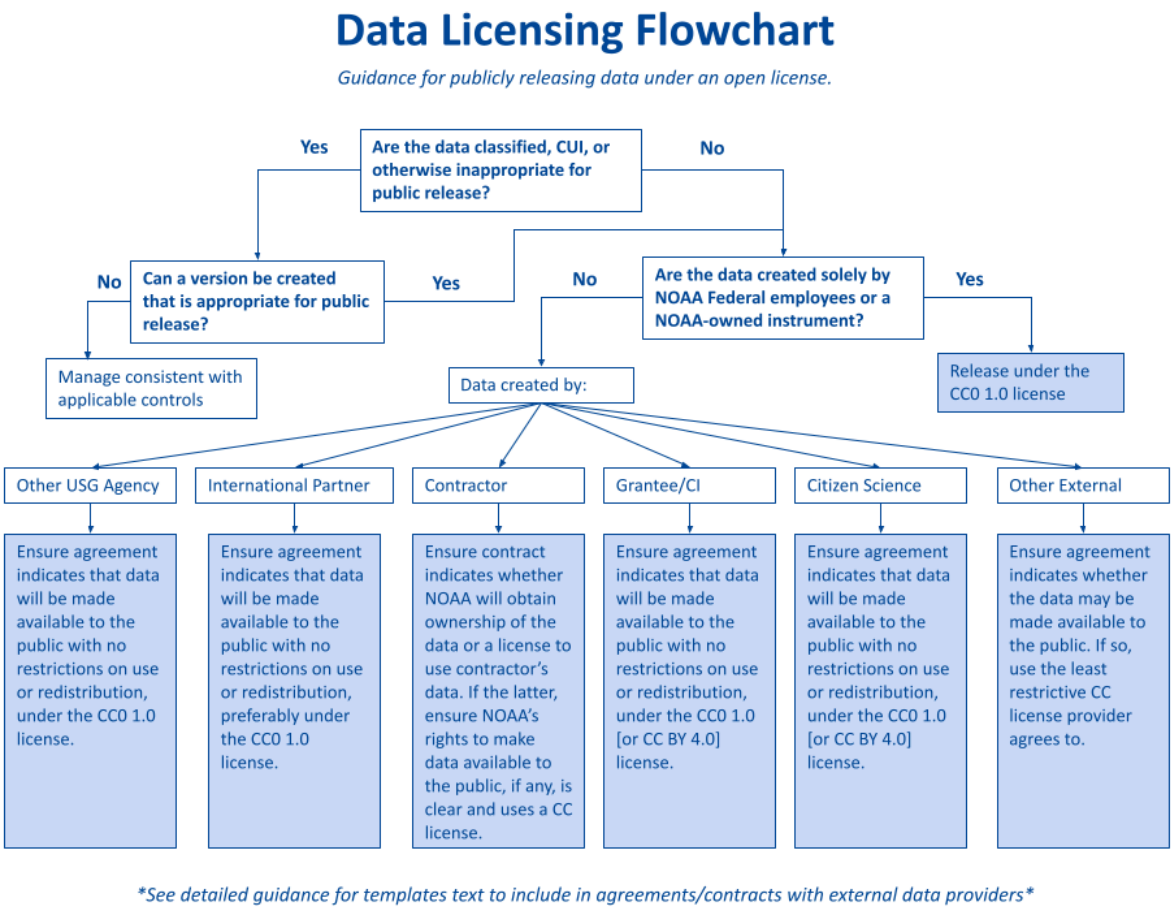


Figure 3: Data Licensing Flowchart

Appendix F - Data Documentation Details

Part 1 - Mandatory and Optional for Metadata Records

NOAA environmental metadata are structured to be consistent with the ISO 19115-2 Geospatial Metadata standard and the ISO 19139 xml representation of ISO 19115-2. NOAA has several tools that can assist with creating robust descriptive metadata for environmental data (see Part 2 below for additional information about some of these tools). The following sections describe requirements for mandatory and preferred/recommended metadata elements that describe all NOAA environmental data. The list is not exhaustive: **More metadata are always better than less metadata.**

Structured metadata (i.e., ISO xml file) is efficient for machine-to-machine communications, but the need for human-readable metadata are also of great value. Providing a 'data dictionary' or other textual description of detailed information about data are invaluable for understanding the described data. For example, explicit information about each column in a spreadsheet (e.g., variable name, units of measurement for that variable, missing value indicator, range of acceptable values) or using 'self-describing formats' (e.g., netCDF) that embeds this metadata in the data files are preferred for long term use and reuse.

1. Mandatory metadata elements. The following descriptive metadata elements are required for all NOAA environmental data. In metadata parlance, a "resource" is described using the identified elements.
    a. Title
    b. Data description (aka 'Abstract')
    c. Earliest and latest observation date
    d. Discovery keywords. Whenever possible, discovery keywords should be based on a standard keyword source or thesaurus, such as the NASA Global Change Master Directory thesauri [or other example...].
        i. Place name(s)
        ii. Instrument name(s)
        iii. Variables measured
        iv. Units of measurement used to measure each variable
2. Preferred but optional metadata elements. The following descriptive metadata elements are preferred/recommended for all NOAA environmental data but are not required. However, if metadata elements are known, the time to record them is now -- when the resource being described is most well understood. The goal of descriptive metadata is to ensure that data are independently understandable for the long-term so that others may make use of the described resource later.
    a. Grant/contract identifier(s)
    b. Funding source(s)
    c. Purpose
    d. Related published works ('cross-references' to other data or publications)

e. Platform name(s)
f. Project name(s)

Part 2 - Metadata Creation Tools

NOAA has several tools that can assist with creating robust descriptive metadata for environmental data. The following list is not exhaustive, but represents current tools available to NOAA environmental data managers for documenting environmental data.

- Collection Metadata Editing Tool (CoMET)
  - Multifunctional tool for documenting a collection or group of related data. CoMET also supports generating a Data Management Plan (DMP) and a Data Stewardship Maturity Report which is currently limited to only NCEI staff. Each user is required to have a userID/password.
- InPort
  - Used primarily by NMFS and NOS environmental data managers. Each user is required to have a userID/password, generated after a login is created by a librarian. InPort also supports generating a Data Management Plan (DMP).
- Send2NCEI (S2N)
  - Used to document relatively small, non-repeating data (e.g., single cruise or survey) that is offered to National Centers for Environmental Information (NCEI) for inclusion in NCEI long term archival repository. Each user is required to have a userid/password.
- Advanced Tracking and Resource tool for Archive Collections (ATRAC)
  - Used to document relatively large and/or repeating data (e.g., long term time series of data from multiple monitoring stations, large volume video data collections) that is offered to NCEI for inclusion in NCEI long-term archival repository. Each user is required to have a userID/password.
- Scientific Data Information System (SDIS)
  - Used exclusively for documenting Ocean Acidification Program (OAP) data.

Part 3 - Description of High-Level Sections of Metadata Records

NOAA data managers developed a Metadata Completeness Tool (aka, 'the rubric') that can be used to assess the presence or absence of metadata elements in an ISO descriptive metadata record. The rubric only assesses presence or absence of content in a metadata element, not the quality or accuracy of that content. It can be used in conjunction with metadata authoring tools, identified above in Part 2, to evaluate how complete the descriptive metadata record is. Tools such as InPort and CoMET automatically compute rubric scores for entered metadata.

The rubric XML is currently held publicly here and continues to be evaluated by the Enterprise Metadata Working Group and will be updated as appropriate.

The rubric tool currently evaluates the presence or absence of content in a standard ISO-compliant metadata record in the following sections:

- *Identification information* provides content needed for basic discovery. It includes the title, a description of the dataset ("abstract"), theme keywords, point of contact, and scope.
- *Access information* describes the data formats, access points, distribution contacts, access constraints, and disclaimers.
- *Coverage information* describes the extent of the data, such as temporal range of content, geographic bounds of content, and place name keywords.
- *Content information* lists the parameters, attributes, variables, or features of the data or provides reference to other documents with this information.
- *History information* identifies the instruments and platforms used to collect the data and/or describes how the data was processed.
- *Quality information* reports on completeness and accuracy of the data or provides references to other documents with this information.
- *Connections* ensure that links have meaningful information associated with URLs, such as name or description and the function.
- *Metadata information* provides the identifier of the record, metadata standard in use, metadata contacts, and any relevant metadata notes.
- *Associated Resource(s)* provide identifiers or citations to related papers, sister resources, parents, or programs related to the data.

Attribution provides information that can be used to cite the data, such as creator, publisher, and DOI.

Appendix G - Data Preservation Details

Part 1 - Components of an Open Archival Information System (ISO 14721)

The basic components of an Open Archival Information System (OAIS, ISO 14721) is represented by the OAIS Reference Model (RM). This diagram (Ken Casey, 2016, pers. comm.) shows the top-level components of an OAIS RM compliant archival system.

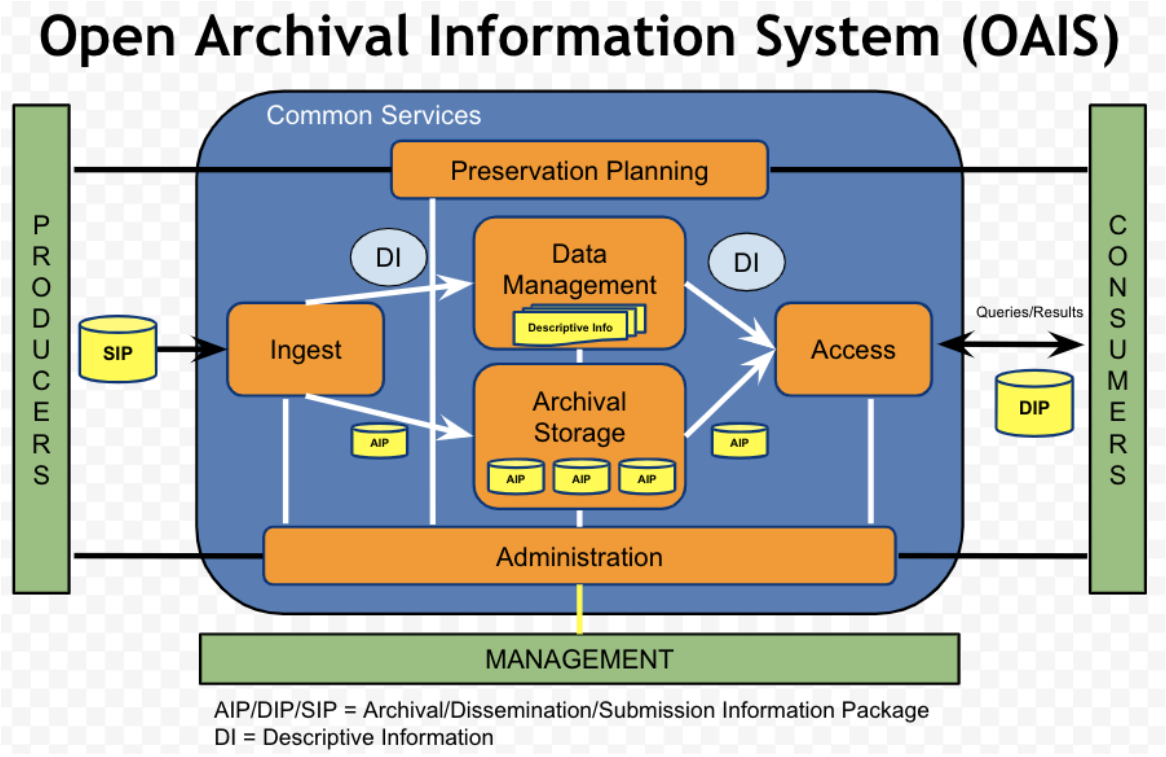


Figure 4: Open Archival Information System

Part 2 - Characteristics of Acceptable non-NCEI Repositories

An accurate list of current repositories that may meet NOAA, NARA and other federal regulations for repository management characteristics is outside the scope of this Handbook. The following types of repositories are recognized by NCEI as acceptable non-NCEI repositories if they meet any of the following criteria:

- Operated by a national government
- Operated by a U.S. state government
- Operated by a public or private university
- Operated continuously for at least twenty years

- Certified as a Core Trustworthy Data Repository
- World Data Service Member Organization

Part 3 - Detailed Archive Appraisal Procedures

The primary purpose of Part 3 is to define the requirements for a formal archive appraisal process for any NOAA data that may be archived by NCEI. A formal appraisal is necessary for requests to archive large, repeating, and/or complex data collections or datasets, e.g., data from a multi-year satellite mission, data from a multi-year networked ocean observing system. The procedure is flexible in that it allows for expeditious decisions regarding scientific records that are known to be within the legal mandates and scope of the NOAA mission and also allows for a lengthier, more formal appraisal process for complex archive requests. However, in either a formal or informal appraisal process, sufficient information to fully appraise the data is required for NCEI to make an informed decision about whether the request to archive will be accepted.

STEP A. Identify Scientific Records for NOAA Appraisal
The first process step involves identifying records that should be appraised for inclusion in or disposal from a NOAA facility's collection.

A.1 Initiation of Request
A Data Producer identifies scientific records and initiates a request that scientific records be added to a NOAA archive or that existing records be removed (disposed of) from a NOAA archive. It is expected that the Data Producer will contact NCEI early in the planning process for appraisal of new scientific records to develop a data management plan that ensures that adequate resources are available to archive the records, especially if the scientific records will be a large volume or complex collection of data and information.

For appraising new records, the Data Producer can be either a NOAA agency or any other national or international organization or agency, or individual. NCEI staff may initiate the process of evaluation in cases where the identified records have no willing or able entity to serve in the Data Producer role.

For appraising existing records, which were previously accepted and are currently contained within a NOAA archive, the Data Producer and NCEI may be the same entity. This occurs when NCEI is determining as to whether existing (previously accepted) records can be disposed of in concert with a NARA disposition schedule.

A.2 Receipt of Request
The request by the Data Producer should be sent to NCEI using one of the established NCEI applications for creating a 'request to archive', e.g. Send2NCEI, ATRAC, or a project-specific submission request process such as the Ocean Acidification Program or Paleoclimatology team process. The request to archive should I identify basic facts about the scientific records that describe the records (metadata), such as title, dataset description or data abstract, volume/size

of the collection, spatial and temporal extents, records format, recording media, and other descriptive information that may be considered important. Examples of other descriptive information to establish NOAA Mission Relevancy can be found in Appendix G Part 5 under Section 1. The Data Producer will receive, within 30 days of NCEI's receipt of the request, acknowledgement of the request and the expected duration of the process, which will return a decision to the requester.

STEP B. Appraising Scientific Records to Determine NOAA Archival Value
The second process step involves appraising the scientific records.

B.1 Designate an Appraisal Team
The NCEI director will appoint a standing Appraisal Team that will perform appraisal of the records.

B.2 Preliminary Records Appraisal
The Appraisal Team evaluates the basic facts and any other descriptive information received from the Data Producer to determine if the request is within the scope of NCEI collection requirements and resources and if there is a legal mandate that requires archiving. If the Archive Appraisal Team determines this is the case, a formal appraisal process is not required and the Appraisal Team will assemble the recommendation package (step B.5). The Appraisal Team and Data Producer will iteratively negotiate what scientific records are to be archived as the background information is gathered and analyzed. These negotiations can also occur outside of the procedure defined in this document.

B.3 Formal Records Appraisal
When needed, the Appraisal Team will meet with the Data Producer to assemble detailed background information about the specific scientific records. The NOAA Scientific Records Appraisal Criteria Questionnaire (Appendix G Part 4) will be used to help gather this background information and can be used as a basis for a future submission agreement for new records approved for a NOAA archive. Either the Data Producer or the Appraisal Team, or both performs the first iteration of gathering background information. As needed, the Appraisal Team will have follow-up discussions with the Data Producer to verify/confirm all background information that has been gathered. The Appraisal Team and Data Producer will iteratively negotiate what scientific records are to be archived as the background information is gathered and analyzed. These negotiations can also occur outside of the procedure defined in this document.

B.4 External Science Review
When needed for more complex archive decisions, the Appraisal Team may request that an External Science Review Team, composed of external to NOAA scientists or users who have expert knowledge of data management and/or the scientific records, assist in reviewing or gathering additional information and provide recommendations. The Appraisal Team can either arrange for their own external science review or, in the case of particularly important decisions, ask the DGC to arrange for such a review. The External Science Review Team will be provided with the results of the Appraisal Criteria Questionnaire and other information as requested to aid

in developing the science team's recommendations. The additional information gathered and recommendations from any External Science Review Team will be documented. All recommendations made to NOAA will be used in accordance with the Federal Advisory Committee Act.

B.5 Assemble a Recommendations Package

The Appraisal Team assembles a recommendation package. The recommendation package will contain an approve, disapprove, or no decision recommendation along with a narrative that explains the decision. Approval recommendations are either "Accept" for new scientific records to be transferred to NCEI or "Retain" for existing scientific records. Disapprove recommendations are either "Reject" for new scientific records to be transferred to NCEI, “Accept with Condition” for new scientific records to be transferred to a non-NCEI archive, or "Dispose of" for existing scientific records. If no decision is provided, the narrative will describe the reasons with a recommendation of further actions that should be taken in order to render a decision. The recommendation package will contain all the background information gathered which includes all information assembled in the Preliminary Records Appraisal (Step B.2), the Formal Records Appraisal (Step B.3), and the External Science Review (Step B.3), when the latter two are conducted.

STEP C. Decision/Approval Process for Accessioning and Disposing of Scientific Records

The third broad process step is a multi-level approval process.

C.1 Receipt of Recommendations Package

The Appraisal Team provides the recommendation package to the NCEI director or the director's designee.

C.2 Office Director Decision/Approval

The NCEI director (or designee) will: a) approve/disapprove the recommendation as is, b) refer the recommendation back to the Appraisal Team for further background information, c) ask the Appraisal Team to conduct an External Science Review or ask DGC to coordinate this external DGC review, or d) coordinate with a NOAA LO for additional assistance when needed or when required by the LO policy. The NCEI director will notify the DGC of all decisions that result in a) scientific records being removed from a NOAA archive orb) scientific records being added to a NOAA archive that have also gone through a formal records appraisal process.

C.3 Coordination with NARA

Decisions to remove scientific records from an existing NOAA archive or to add scientific records to a NOAA archive will be performed by NCEI according to NARA-approved disposition schedules as contained in the NOAA Records Disposition Handbook before the decision is implemented.

C.4 Public Comment and Appeal Period

Any decision that results in a) existing scientific records being removed from a NOAA archive or

b) newly acquired scientific records being added to a NOAA archive that have also gone through a formal records appraisal process will be advertised for public comment and appeal by the NCEI director using their LO's procedure for implementing the "NOAA Policy on Partnership in the Provision of Environmental Information." Before the decision is implemented, any public comments and appeals received will be considered by the Appraisal Team for possible revisions to the recommendations package. The time period for public comments and appeals is prescribed in the Line Office's Partnership Policy. The NCEI director can also use other means of advertising decisions as a result of this procedure to the affected user community such as society journal articles, constituent meetings, newsletters, etc.

C.5 Archive Appraisal Reports and Process Tracking

NCEI will track all decisions and document the steps that result in a) scientific records being removed from a NOAA archive or b) scientific records being added to a NOAA archive that have also gone through a formal records appraisal process. NCEI will also provide a summary of archive decisions to the DGC to include reports, annual program planning, and budgeting processes. These updates are especially important for scientific records that have been approved for archive but where resources may not exist or are insufficient to support the archive.

STEP D. Implementing the Decision

The fourth broad process step is implementing the decision. Before implementing any decision, the NCEI director or designee will ensure that coordination with NARA and any public comments and appeals have been adequately addressed. Adequate resources must exist before a decision is implemented.

D.1 Decision Implementation for NEW Records

The NCEI director will notify the Data Producer of the decision. If the decision is to not archive at NOAA, the NCEI director will make a recommendation to the Data Producer as to where the scientific records could be archived when possible.

For records that have been approved for inclusion in a NOAA archive, the process that establishes a Submission Agreement will begin, or if already begun, will proceed to the establishment of a formal agreement between the Data Producer and NCEI. Data Producer responses to the Appraisal Questionnaire will be used as the initial basis of documentation in the Submission Agreement.

D.2 Decision Implementation for Existing Records

If the decision is to dispose of records and the coordination with NARA (Step C.3) has been completed and any public comments and appeals that were received (Step C.4) have been adequately addressed, the NCEI director will attempt to donate the records to interested agencies or individuals to include the original records creator when applicable. The agencies or individuals contacted will be documented. If there is no interest, the records can be destroyed in

accordance with established NARA Records Disposition Schedule requirements. The NCEI Director has the discretion to retain these records for a longer period of time.

Part 4 - NOAA Formal Records Appraisal Questions

Summary of NOAA Formal Records Appraisal Questions

Used for more complex archival review and appraisal processes

Records Appraisal – Short Question	**Records Appraisal – Detailed Questions**
Mission Relevant?	1. Where in NOAA's mission? 2. Environmental or geospatial? 3. Legal mandates?
Uniqueness, Provenance?	4. Unique or duplicated elsewhere? 5. Relationship to other NOAA data? 6. Authentic, reliable, unaltered, and usable? 7. Original purpose? New purpose? 8. Records value (now, future)?
Nature of the data?	9. Volume (bytes), growing or static? 10. Temporal and/or spatial extent? 11. Data format? 12. Solely digital or does analog exist? 13. Physical condition? Deterioration? 14. Is information retrievable? 15. Records location? Science center?
Metadata?	16. Does it exist? 17. Does it conform to standard format?
Processing level?	18. Completeness and quality? 19. "Raw" observational records? 20. If not "raw", level of processing? 21. Multiple versions?
R&D records?	22. Unprocessed or processed? 23. NOAA funded research?
Externally reviewed?	24. Evaluation or peer reviewed?
Restrictions?	25. Proprietary, sensitive, classified, CUI?
Intrinsic value?	26. Historic, aesthetic, artistic?
Resources?	27. Cost of long-term maintenance? 28. Resources for data stewardship?

Part 5 - NOAA Scientific Records Appraisal Criteria Questionnaire

Using the Appraisal Criteria Questionnaire: The Appraisal Team will use the background information collected from the questions found below to make decisions about scientific records currently within, or requested to be included, in a NOAA archive. This will result in more consistent appraisal decisions that can be readily explained both within NOAA and to its constituents. Additional appraisal questions and answers can be added by the Appraisal Team when needed and can be used in the appraisal process. A web-based system is anticipated for this questionnaire that will provide for easier data entry and analysis.

The appraisal questions were developed from guidelines produced by National Archives and Records Administration (NARA), from National Research Council (NRC) reports that contained recommendations to NOAA on data management, and appraisal processes used by other Federal Agencies. All questions have one or more references from the NARA and/or NRC reports indicating the origin or basis for the question. These references contain the actual wording in italics extracted from the NARA/NRC reports. The intent for this is twofold. The actual reference wording will: 1) further explain the question to those answering the questions and 2) provide context to the Appraisal Team during their evaluation of the answer to the question.

As described by NARA (2007), applying these questions to specific scientific records "... is not a mechanical process akin to adding up points or checking boxes. The questions should be considered together, rather than in isolation." Finally, it is not the intent that this is a static list of questions. Rather, it is expected that these questions will evolve over time based upon experience gained by using the Criteria Questionnaire tool and by incorporating new information from future assessments of NOAA data management activities.

Outline of the Appraisal Criteria Questionnaire:

Section 0: Administrative Metadata
Section 1: NOAA Mission Relevancy
Section 2: General Facts
Section 3: Physical Facts
Section 4: Metadata Facts
Section 5: Record Processing Level Facts
Section 6: Research and Development (R&D) Records
Section 7: External Records Review Processes
Section 8: Records Restrictions
Section 9: Records with Intrinsic Value
Section 10: Resources
Section 11: References for Questionnaire

SECTION 0: ADMINISTRATIVE METADATA

Collection Name, Date of Submission, and Date of Review:

Information Provider Name and Organization: Address:
Team Lead and Team Members: Lead Telephone & Email address:

Appraisal Team Information Organization: Address:
Team Lead and Team Members: Lead Telephone & Email address:

SECTION 1: NOAA MISSION RELEVANCY

1. **Where do these records fit within NOAA's mission?**
    a. **See *current NOAA strategic plan (www.ppi.noaagov/spo.htm)***
    b. ***NRC Principle #1 (2007):*** *Environmental data should be archived and made accessible*
2. **Are these scientific records Environmental Data or Geospatial Data as defined in NOAA Administrative Order (NAO-212-15) entitled *Management* of *Environmental and Geospatial Data and Information?***
   <www.corporateservices.noaa.gov/~ames/NA0s/Chap_212/naos_212_15.html>
    a. ***Environmental Data*** *- recorded observations and measurements of the physical, chemical, biological, geological, or geophysical properties or conditions of the oceans, atmosphere, space environment, sun, and solid earth, as well as correlative data and related documentation or metadata. Media, including voice recordings and photographs, may be included.*
    b. ***Geospatial Data*** *- information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the Earth. This information may be derived from, among other things, remote sensing, mapping, and surveying technologies. Statistical data may be included in this definition at the discretion of the collecting agency.*
3. **Do the records have legal mandates, which require their archiving?** If **yes, list them. Are there existing NARA dispositions schedules that pertain to these records?**
    a. **NRC (2007):** *NOAA must continue to archive and provide access to all data as required by law.*

SECTION 2: GENERAL FACTS

4. **Are the records unique?** If **not unique, where else do they exist?**
    a. ***NARA (2007):*** *Appraisals must be conducted in context with other records. The appraisal must determine whether the records under consideration are the only or are the most complete source for significant information. Records that contain information not available in other records (including other Federal records as well as files accumulated by state and local governments) are more likely to warrant permanent retention than records containing data that* is *duplicated in other sources. However, NOAA may decide to retain records that contain information available elsewhere in the case of records that are more complete or more easily accessible than the alternative source.*
    b. ***NRC Guideline (2007):*** *The most obvious candidates for reduced archiving requirements are data that are obsolete or redundant, that could be regenerated on demand, or clearly have only short-term uses. This includes older versions of reprocessed data and model output.*

c. ***NRC Guideline (2007):*** *NOAA should establish close partnerships with other national and international data holding institutions and engage these institutions as part of the archiving process. It* is *important to have clear agreement on which partner has what archival responsibility.*
d. ***NRC (1995):*** *For both observational and experimental data, the following retention criteria should be used to determine whether a data set should be saved: uniqueness.*

5. **Are the records related to other records in a NOAA archive, *i.e.,* extensions, new versions, improved quality, *etc.?*** If **yes, to what degree do the records add value to other records held by NOAA or others?**
    a. ***NARA (2007):*** *Other things being equal, records that add significantly to the meaning or value of other records already appraised as permanent are more likely to warrant retention than records lacking such a relationship. Records that are chronological continuations of records already in the NOAA archive are likely to warrant permanent retention, particularly if the older segments of the records are subject to high reference use.*
6. **Are the records judged to have authenticity, reliability, integrity and usability (see Definitions Section)?**
    a. ***NARA (2007):*** *To be appropriate for long-term temporary or permanent retention, observational data should possess authenticity, reliability, integrity, and usability (as defined in ISO 15489-1:2001 entitled "Information and documentation -- Records management." Intellectual linkage with the related metadata is essential.* ***http://www.whitefoot-forward.com!iso_l5489-1.pdf***
7. **What was the original purpose of the records? Do the records provide information and value beyond their original purpose and user community?**
    a. ***NARA (2007):*** *Records are more likely to be appraised as permanent if they not only can be used for scientific purposes but also for legal, commercial, educational, engineering, resource management, or other purposes.*
8. **What is the value (scientific, public, government) of the records to be archived in terms of current and anticipated future benefits and levels of service required to achieve these benefits?**
    a. **NARA (2007):** *The future research potential of records is the most difficult variable to determine. What* is *of relatively low research use today may become of great research use in the future. Perhaps even more important and difficult to predict are the issues and topics that will be considered of significance in the future. Nevertheless, it* is *important to consider this question in making appraisal decisions. It* is *necessary to consider the kinds and extent of current research use and to try to make inferences about anticipated use both by the public and by the Government.*
    b. **NRC (2007):** *Not all data sets are of equal value, and practical constraints prevent nil data from being archived and made readily accessible, so at some point certain data will need to be designated for reduced archiving and/or access requirements. Ideally, this decision would be made based on the current utility and potential future value of the data, but ... it is extremely difficult to assess even*

the current value of any particular environmental data stream. Likewise, it is *virtually impossible to anticipate its potential future uses. The decision-making process also needs to be ongoing, with data managers/stewards continually reviewing the data holding under their purview to determine the appropriate level of service for each data set given legal and mission requirements, user needs, and available resources.*

SECTION 3: PHYSICAL FACTS

9. **What is the volume of the records (archive storage size)? Is the record collection static or growing? If the collection is growing, what is the expected volume?**
    a. ***NARA (2007):*** *Volume will play a role only in the appraisal of records whose archival value* is *marginal. Other things being equal, records that are compact are more likely to be appraised as permanent than those that are voluminous.*
10. **What are the time period (temporal range) and location (spatial area) that are covered by these records?**
    a. ***NARA (2007):*** *Observational records covering a long time period tend to have more value because they enable long-term patterns to be identified and thereby increase confidence in the reliability of data and the conclusions drawn from them.*
11. **What is the data format of the records?**
    a. ***NARA (2007):*** *Some records may pose such technological challenges that extraordinary measures may be required to recover the information, while other records containing similar documentation (either electronic records or records in another format) may be usable with much less effort.*
    b. ***NRC (2007):*** *The best archive formats are those where the digital content of each data record can be described in elementary terms (for example, number of bytes, numeric type, character string, pixel, etc.). This* is *one feature of an open format standard that helps minimize software and computer operating system dependencies that could render the data inaccessible in the worst case. So-called proprietary formatted data (non-open format description) should in general not be considered as a good candidate for long-term archiving unless a plan and a process are in place to translate the data to an open format standard.*
12. **If these records are currently in electronic format, do these records still exist on other media *(e.g.,* paper, film)? If yes, is it required to maintain copies on other media?**
    a. **NARA *(2007):*** *Many data series now collected in electronic format were formerly created and maintained in other formats such as paper or photographs. Agencies may still maintain older data in such formats for use in conjunction with the related electronic data. Appraisers should extend their review of electronic systems to include any related data in other formats, as these older data may add to the usefulness of the electronic data if they are still in a usable format. All formats should be considered during the appraisal.*
13. **What is the current storage media for the records? How does the physical condition of the media affect their usability? Is the current storage media at risk?**

a. ***NARA (2007):*** *Some records may have deteriorated to the point that the information they contain* is *not readable.*
b. ***NRC (1995):*** *The appraisal process must apply the established criteria while allowing for the evolution of criteria and priorities, and be able to respond to special events, such as when the survival of data sets is threatened.*

14. **Does appropriate hardware and software technology exist to enable usability of the records? If yes, describe.**
    a. **NARA** ***(2007):*** *Some records may pose such technological challenges that extraordinary measures may be required to recover the information, while other records containing similar documentation (either electronic records or records in another format) may be usable with much less effort.*
    b. **NRC 1995:** *For both observational and experimental data, the following retention criteria should be used to determine whether a data set should be saved: availability of hardware to read the records*
15. **Have the records ever physically resided at a scientific data center or center of data where stewardship was provided? Where do they reside now? What scientific expertise would best provide stewardship for the records?**
    a. ***NARA (2007):*** *It is appropriate for many observational data of long-term temporary or permanent value to be maintained on a continuing basis by a scientific data center that possesses the necessary expertise to ensure preservation and access.*
    b. **NRC** ***Principle#*** **6** ***(2007):*** *Data and metadata require expert stewardship.*
    c. ***NRC Guideline (2007):*** *Good stewardship requires systematic, ongoing assessment and improvement of data and metadata.*

SECTION 4: METADATA FACTS

16. **What metadata exists and is the metadata sufficient to support the broad understanding of the scientific records?**
    a. ***NARA (2007):*** *Metadata should include information such as purpose and time period of data collection; location of collection site; methods and instrumentation used in collection; units of measurement, acceptable values, and error tolerance; data aggregation methods; processing history; and quality assessment. The types of metadata required vary with the nature of the data and their likely future uses*
    b. ***NRC Principle*** #5 ***(2007):*** *Metadata are essential for data management.*
    c. ***NRC Guideline (2007):*** *Metadata that adequately document and describe each archived data set should be created and preserved to ensure the enhancement of knowledge for scientific and societal benefit.*
    d. **NRC** *(1995): For both observational and experimental data, the following retention criteria should be used to determine whether a data set should be saved: adequacy of documentation (metadata). Complete metadata should define the content, format or representation, structure, and context of a data set.*
17. **Is the metadata in a standard format or can it be automatically translated into a standard format? What other important metadata exists that is not standardized?**

a. ***NARA (2007):*** *It is preferable for metadata, whenever possible, to conform to standards issued by such broad-based organizations as the Federal Geographic Data Committee (FGDC) and the International Organization for Standardization (ISO).*
b. **NRC *Guideline (2007):*** *The application and expansion of metadata and related standards are essential for good stewardship; NOAA and its partners should continue to expand their usage of standards and reference models.*
c. ***NRC (2007):*** *Metadata should be stored in similarly open formats and should be tightly coupled with and managed in conjunction with the data so both are always readily available to the user.*

SECTION 5: RECORD PROCESSING LEVEL FACTS

18. **What is the completeness and quality of the scientific records and metadata?**
    a. ***NARA (2007):*** *Additional factors favoring long-term or permanent retention are the completeness and quality of observational data; quality and completeness of metadata*
19. **Describe the data processing level of the scientific records. For example, are the records "raw" or minimally processed, quality controlled or calibrated, *etc.?***
    a. ***NARA (2007):*** *Raw or minimally processed records are more difficult for anyone except the primary user(s) to understand and use but are essential for conducting a reanalysis, such as to verify findings or support a new hypothesis. These observational records are likely to be appraised as either long-term temporary or permanent. Unlike laboratory experimental data, observational records typically document phenomena that can never be repeated. Observational records establish a baseline to help determine future rates of change and frequency of occurrence of unusual events. Moreover, observational records frequently can be processed and used in novel ways, for example, to verify new scientific concepts.*
    b. **NRC *Guideline (2007):*** ***It*** is *especially important to save the most primitive useful forms of all environmental data.*
    c. **NRC *(1995):*** *As a general rule, all observational data that are non-redundant, useful, and documented well enough for most primary uses should be ... maintained.*
20. **If not "raw" or minimally processed, describe the data processing level of the scientific records.**
    a. **NARA *(2007):*** *Appraisal decisions should consider that the uses of data vary according to the level of processing. Processed records are more likely to have long-term value if they would be costly to recreate from the raw data. It may be warranted to appraise as permanent both a raw version and one or more processed versions of certain records. With each higher level of processing, records generally become easier to use but less subject to reanalysis. To facilitate future reanalysis, it is usually appropriate to preserve processed records at the lowest level of processing compatible with effective use.*
    b. **NRC *Guideline (2007):*** *It may be more cost-effective to regenerate certain kinds of environmental data on demand.*

c. ***NRC Guideline (2007):*** *The most obvious candidates for reduced archiving requirements are data that are obsolete or redundant, that could be regenerated on demand, or clearly have only short-term uses. This includes older versions of reprocessed data and model output.*

21. **If these records are processed, do multiple versions of the same processed records exist?**
    a. **NARA** ***(2007):*** *Processed data are more likely to have long-term value if they would be costly to recreate from the raw data. It may be warranted to appraise as permanent both a raw version and one or more processed versions of certain data.*
    b. **NRC** ***Guideline (2007):*** *It may be more cost-effective to regenerate certain kinds of environmental data on demand.*
    c. **NRC** ***(2007):*** *In situations where multiple versions of derived products have been generated, it would be helpful to have a defined process in place to determine which versions need to be archived. The following three questions, for example, could form the basis for such decisions. If the answer to all three questions is positive, then multiple versions should be archived:*
        i. *Is it feasible to retain multiple versions of the data?*
        ii. *Are the differences among the various versions sufficiently large and scientifically important to make it worth preserving multiple versions?*
        iii. *Is it too technically difficult to regenerate earlier versions?*

SECTION 6: RESEARCH AND DEVELOPMENT (R&D) RECORDS: RECORDS GENERATED AS A RESULT OF AN EXPERIMENT USING THE WORKFLOW PROCESS BASED UPON THE SCIENTIFIC METHOD. SKIP TO THE NEXT SECTION IF RECORDS ARE NOT R&D.

22. **Are the R&D records unprocessed (original or raw) or processed (compiled or analyzed products)?**
    a. ***NARA (2007):*** *Raw data are generated by an experiment, whereas processed data consist of raw data manipulated to help identify patterns in the data. Research data commonly have short-term value when they are narrow in scope and can be replicated by a new experiment if necessary. For data to be valuable over the long term, they should be unique, complete, valid, and accompanied by appropriate metadata. Data with long-term research value often are most appropriately maintained by the* R&D *agencies, which created them because the creating agencies usually possess the scientific expertise essential for providing effective access to the data.*
    b. **NRC** ***(1995):*** *Laboratory data sets are candidates for long-term preservation if there* is *no realistic chance of repeating the experiment, or if the cost and intellectual effort required to collect and validate the data were so great that the long-term retention* is *clearly justified.*
23. **If the R&D was performed by a non-NOAA entity, was the project funded by a Federal funding source contract or grant?**
    a. ***NARA (2007):*** *For projects funded by contracts, records specified in the contract as deliverables generally are Federal records and, in conformance with the*

contract requirements, may be maintained by either the contractor or the funding agency. By contrast, the primary records of grant-funded projects usually are not considered to be Federal records and are maintained by the grantee. Recordkeeping for collaborative projects is *affected by the diversity of funding sources and institutions (including nonfederal institutions) involved. Effective appraisal of these records requires a determination of which institutions have responsibility for the records and their disposition.*

SECTION 7: EXTERNAL RECORDS REVIEW PROCESSES

24. **Have the records undergone user evaluation and/or scientific peer review, been used extensively in publications, and/or subjected to other appraisal processes such as the NOAA Satellite Products and Services Review Board (SPSRB)?** If **yes, please describe.**
    a. ***NARA (2007):*** *In general, data are more likely to be appraised as permanent if the data have successfully undergone the scientific peer review process. This* is *especially true for processed records.*
    b. ***NRC Principle # 3 (2007):*** *Environmental data management activities should recognize user needs.*
    c. ***NRC Principle# 7 (2007):*** *A formal ongoing process, with broad community input, is needed to decide what data to archive and what data not to archive.*
    d. ***NRC Guideline (2007):*** *It* is *essential to solicit user input when making decisions on whether to archive or continue archiving a data set.*
    e. **NRC *Guideline (2007):*** *Because the decision to stop archiving* is *normally irrevocable, extra attention to community engagement* is *needed before final disposal of any data.*
    f. **NRC *(1995):*** *For both observational and experimental data, the following retention criteria should be used to determine whether a data set should be saved: evaluation by peer review. All stakeholders scientists, research managers, information management professionals, archivists, and major user groups should be represented in the broad, overarching decisions regarding each class of data.*

SECTION 8: RECORDS RESTRICTIONS

25. 25. **Do any restrictions apply to the records *(e.g.,* redistribution, proprietary, national security, classified, sensitive natural resource, others)?** If **yes, describe the restrictions.**
    a. **NRC *(2007):*** *There are some data for which access restrictions are clearly needed, such as the location of rare in situ specimens or data with national security implications. Some data sets NOAA will want to archive are proprietary in nature, particularly data derived from international and/or commercial sources. There should be provisions in data management systems for incorporating such data.*

b. **NRC** *(1995): Classified data must be evaluated according to the same retention criteria as unclassified data in anticipation of their long-term value when eventually declassified. Evaluation of the utility of classified data for unclassified uses needs to be done by stakeholders with the requisite clearances to access such data.*

SECTION 9: RECORDS WITH INTRINSIC VALUE

26. **Do the records have intrinsic value?**
    a. **NARA *(2007):*** *Records with intrinsic value are rare and possess one or more specific qualities or characteristics as defined by NARA. These include but are not limited to records in an original form that document an early media type (e.g., glass plate negatives, wax cylinder recordings, etc. - Note that only a representative sample would have intrinsic value and not the entire collection), Aesthetic or artistic quality (e.g., manuscripts; photographs; pencil, ink, or watercolor sketches; maps, etc.), Age (e.g., Generally, records of earlier date are of more significance than records of later date).*

SECTION 10: RESOURCES

27. **What are the cost considerations for long-term maintenance of the records? Are resources available for archiving and providing access to these records?** If **pertinent to the appraisal decision, has a detailed cost/ benefit analysis of the records been completed *(e.g.,* USGS cost/benefit analysis located at: <http://eros.usgs.gov/government/ratool/view_questions.php>)?**
    a. ***NARA (2007):*** *This consideration should play a significant role only in marginal cases. In such cases, an appraisal should balance the anticipated research potential of the records with the resource implications of retaining them permanently. Other things being equal, records with low long-term cost implications are more likely to warrant permanent retention than those records that earn; high long-term costs.*
    b. ***NRC Principle #2 (2007):*** *Data generating activities should include adequate resources to support end to-end data management.*
    c. ***NRC Guideline (2007):*** *Archiving and access decisions are closely related. In general, when resources are limited, access to older or less commonly used data should be scaled back rather than removing data from the archive.*

28. **Are resources available for Data Stewardship that will enable activities that preserve and improve information content, accessibility, and usability of the records based upon technology changes and future discoveries that advance the understanding and knowledge of the records?**
    a. **NRC *Guideline (2007):*** *Good stewardship requires systematic, ongoing assessment and improvement of data and metadata.*

b. **NRC *Guideline (2007):*** *NOAA should establish and maintain data and metadata migration plans for all current and future long-term archive systems to adapt to information technology evolution.*

Appendix H - Data Citation Details

Part 1 - Requesting Reserve Data DOIs

Caveats to requesting a reserved data DOI from NCEI:

The following caveats apply when requesting a reserved DOI from NCEI.

- Without exception, data getting a NOAA DOI must be under archival control at NCEI prior to minting the DOI. The foundation of NAO 212-15B is that "data are archived at NCEI" prior to getting a DOI.
- Prior to creating a reserved DOI, the rationale for not providing access to data with a reserved DOI must be documented such that NCEI Archive Branch, Data Access Branch and Customer Engagement staff can understand and explain why data cannot be accessed.
- Unless data for which a reserved DOI is created meet one of the exempted classes identified in the Freedom of Information Act (FOIA), NCEI must make it clear to the data provider that data not being made publicly accessible will be released if NCEI receives a FOIA request for those data, with or without a DOI.
- Data from a Federal source cannot be withheld from access "solely for the purpose of being the first to publish" per the NOAA Data Access Section.

Part 2 - Additional Guidance and Resources for Data Citation

In general, NCEI and the NOAA Central Library will not assign an additional DOI for environmental data or a publication for which a DOI already exists. If a DOI is provided and maintained in another resource, that DOI must appear in the NOAA metadata for that dataset or publication.

Community Resources
Federation of Earth Science Information Providers: ESIP data provider guidelines
DataCite.org: DataCite's Why cite data?
The Dataverse Network Project: Dataverse Network Project data citation guidance
The Digital Object Identifier System: DOI System FAQ page
Digital Curation Center: Digital Curation Center data citation guidance

Appendix I - Data Access Details

Part 1 - NOAA Dissemination Requirements and Recommendations

The NOAA Data Governance Committee established a cross-Line Office task team to develop a current state analysis and future recommendations for NOAA's environmental data dissemination activities. The Final Report (2023) includes requirements and recommendations driven by user needs which fall into five categories: Data Access, Data Availability, Data Management & Technologies, Data Quality & Integrity, and Data Community (link will be added when the document is released).

Part 2 - Suggested Data Access Providers and Submission Tools

Suggested data access providers and submission tools: NOAA has several tools that can assist with making data publicly available. The following list is not exhaustive, but represents current tools available to NOAA environmental data managers for publishing environmental data. Additionally, the U.S. Geological Survey (USGS) maintains a list of alternate repositories that may be appropriate for data outside of the NCEI Archive Collecting Policy scope.

- Send2NCEI (S2N)
    - Used to document relatively small, non-repeating data (e.g., single cruise or survey) that is offered to National Centers for Environmental Information (NCEI) for inclusion in NCEI long-term archival repository. Each user is required to have a userID/password.
- Advanced Tracking and Resource tool for Archive Collections (ATRAC)
    - Used to document relatively large and/or repeating data (e.g., long term time series of data from multiple monitoring stations, large volume video data collections) that is offered to NCEI for inclusion in NCEI long-term archival repository. Each user is required to have a userID/password.
- Scientific Data Information System (SDIS)
    - Used exclusively for documenting Ocean Acidification Program (OAP) data.
- Comprehensive Large Array-Data Stewardship System (CLASS)
    - The Comprehensive Large Array-data Stewardship System (CLASS) is NOAA's information technology system designed to support long-term, secure preservation and standards-based access to environmental data collections and information, currently supporting POES, GOES, DMSP, JPSS (formerly NPOESS), RADARSAT, JASON, GNSS, and selected model reanalysis data.
- Environmental Research Division's Data Access Program (ERDDAP)
    - A data broker that provides access to configured gridded and tabular data in many file formats through a graphical user interface (GUI) and a machine-to-machine RESTful interface. ERDDAP uses various protocols for

accessing data: griddap for subsetting gridded datasets, tabledap for tabular datasets, a files system to allow users to browse datasets, and a Web Map Service to request images with associated data plotted on the image

- NOAA GeoPlatform
    - Used to browse public NOAA data via Esri's ArcGIS Online. Users can access various NOAA map services to visualize public NOAA datasets
- NOAA Open Data Dissemination (NODD)
    - NODD is considered a NOAA Enterprise Dissemination System, and all NOAA Line Offices can utilize NODD for dissemination of NOAA open data if it meets their dissemination requirements. Using its arrangements with commercial cloud service providers, NODD disseminates near real-time data, non-real-time data, and static data. The data must be quality assured, of value to the public, and without any use restrictions. User Registration is not required to access data. Data owners must provide appropriate metadata, update relevant catalogs, point to the cloud instance from their website, and comply with relevant LO governance. There are no egress costs for users or the agency.
- Digital Coast
    - Used to provide access to NOAA datasets that are key to managing our coasts. Digital Coasts has a range of tools, training and information needed to effectively use the content on their platform. No registration is required to download the data.

Appendix J - List of Metrics

Metric C.1.1: Number of datasets that are tracked.

Metric D.1.1: Percentage of datasets that are covered by a DMP.

Metric E.2.1: Percentage of datasets documented with metadata in the NOAA Data Catalog.

Metric H.1.1: Percentage of datasets that are accessible by the public.

Metric H.1.2 : Percentage of datasets that are restricted.