NOAA Data Documentation Directive: Deprecated 10/01/2024

Table of Contents

I. Purpose

The purpose of this *NOAA Data Documentation Directive* is to ensure that NOAA environmental data are documented with machine-readable metadata to enable discovery, access, and use of these resources. This Directive supports NOAA Administrative Order (NAO) 212-15, *Management of Environmental Data and Information*[1] (2010), which states that "environmental data will be visible, accessible and independently understandable to users, except where limited by law, regulation, policy, … or security requirements."

This version of the *Data Documentation Directive* was issued by the NOAA Environmental Data Management Committee (EDMC) and supersedes the previous version (v.1, 2011).[2]

---

[1]   http://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_212/212-15.html

[2] https://nosc.noaa.gov/EDMC/PD.DD.php

II. Scope

1. This Directive requires **metadata**[3] for:
    1.1. **On-line environmental data**, including both original observations and derived data products, resulting from NOAA observing systems, numerical models, and intramural research (whether performed by federal personnel or by on-site contractors or affiliates).
    1.2. **Off-line environmental data holdings**, including data stored on paper, on CD-ROM, or on off-line tape, disk, or other media.
2. A **Waiver** or an **Extension** (see Sec IX) can be filed with EDMC under certain conditions.
3. This Directive does not apply to:
    3.1. Data produced by extramural grants and contracts, which are instead subject to the *Data and Publication Sharing Directive for NOAA Grants, Cooperative Agreements, and Contracts (2016).*[4] That directive requires metadata for grant- and contract-produced data.
    3.2. Third-party data purchased or acquired from external sources, which are instead subject to the *NOAA Recommended Practice for Use of External Data*.[5] However, derived products created by NOAA from third-party data are subject to this Directive unless the terms of the purchase or acquisition forbid making derived products accessible.
    3.3. Text-based warnings, advisories, forecasts, and similar textual products.
    3.4. Web server access statistics, forecast performance metrics, and other business process information that are about data but do not themselves constitute environmental data.

III. Authority

This Directive has been issued by the NOAA Environmental Data Management Committee, with the approval of the NOAA Chief Information Officer (CIO) Council and the NOAA Observing System Council (NOSC), pursuant to the authorities granted in NAO 212-15.

Relevant agency and national policy:
- NAO 212-15[1]
- NOAA Environmental Data Management Framework[6]

---

[3] See definitions in Section XI.

[4] https://nosc.noaa.gov/EDMC/PD.DSP.php

[5] https://nosc.noaa.gov/EDMC/RP.UED.php

[6] https://www.nosc.noaa.gov/EDMC/framework.php

IV. Directive

This Directive has five core requirements. The sub-sections below each one provide additional detail.

Note: Data already archived at the NOAA National Centers for Environmental Information (NCEI) were documented at or before submission, and are therefore considered to have met all requirements of this Directive. New submissions to NCEI, and data not submitted to NCEI, must independently satisfy the requirements below.

1. **Metadata requirement:** All NOAA environmental data, as defined in Section XI, shall be described by a formal metadata record.
    - 1.1. Metadata records shall be created for both on-line digital data and off-line data, including data on paper, CD-ROM, magnetic tape, and other physical media.
        - 1.1.1. Off-line data may be documented at the collection or facility level if necessary (e.g., one record summarizing a room-full of holdings).
        - 1.1.2. If an on-line copy of off-line data is available, the off-line version is not required to have a separate metadata record.
    - 1.2. Metadata records shall be made publicly accessible on-line in a Web Accessible Folder (WAF).[7]
    - 1.3. Metadata WAFs shall be registered with the NOAA Data Catalog[8] as described in the *NOAA Data Access Directive.*[9]
    - 1.4. Within one year of the effective date of this Directive:
        - 1.4.1. Metadata shall be produced and registered.
        - 1.4.2. Extension requests (Sec. IX) shall be filed with EDMC for datasets that require additional time for metadata completion.
        - 1.4.3. Waivers (Sec. IX) shall be filed with EDMC for legacy or experimental datasets that will not have metadata.
2. **Metadata contents:** Metadata records shall include, at minimum, the information necessary to enable potential users to discover, access, evaluate, and use the dataset, and may include references to supplementary documentation.
    - 2.1. Appendix A summarizes mandatory and optional content for dataset metadata records, and provides links to the Metadata Rubric tool to evaluate metadata completeness. Appendix B describes some available tools to create metadata records.
    - 2.2. Metadata records for data accessible on-line shall include links to machine-readable data-access methods as required by the *NOAA Data Access Directive.*[9]
    - 2.3. Metadata records for data not accessible on-line shall include the physical location (address) of the data, contact information for the data custodian, instructions for off-line access (if available), and any constraints on access.
3. **Metadata granularity:** Wherever possible, metadata records shall be provided for aggregations of closely-related observations. 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.
    - 3.1. Data from an ongoing time series, and collections of related data from multiple

---

[7] A WAF is a web server directory at a specific internet address (Uniform Resource Location, or URL) that directly lists one or more metadata records without an intervening index page.

[8] https://data.noaa.gov/

[9] https://nosc.noaa.gov/EDMC/PD.DA.php

locations, shall be represented by an aggregated metadata record.

    3.2.   Other levels of aggregation shall be at the discretion of the producers and stewards of the data.

    3.3.   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.

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

    4.1.   The default encoding for NOAA metadata shall be the International Organization for Standardization (ISO) Geographic Metadata Standard.

        4.1.1.   The ISO version currently in effect at NOAA is ISO 19115-2 (2009)[10] conceptual model with ISO 19139-2 (2012)[11] XML schema.

    4.2.   Legacy metadata records in US Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata format shall be accepted if already developed for existing datasets, but documentation of datasets that are new or previously undocumented as of the effective date of this Directive shall be required to use the ISO Geographic Metadata Standard.

        4.2.1.   FGDC metadata record contents shall comply with requirement IV.2, above.

        4.2.2.   Producers of FGDC metadata shall report annually to EDMC a schedule for transition to ISO metadata or a statement of unmet funding requirements.

    4.3.   No metadata standards or formats other than those described in 4.1 and 4.2 are currently accepted for registration in the NOAA Data Catalog.

5. **Transition to revised ISO standard:** The NOAA Metadata Working Group (WG)[12] shall develop and implement a Transition Plan for supporting the revised ISO Geographic Metadata Standard.

    5.1.   The revised standard includes ISO 19115-1 (2014)[13] conceptual model, ISO 19157 (2013)[14] data quality information model, and XML schema ISO 19115-3 (in press, 2016) and ISO 19157-2 (committee draft, 2016).

    5.2.   Within one year of the effective date of this Directive, the Metadata WG shall submit to EDMC either a Transition Plan, or a draft Plan and a request for schedule extension.

    5.3.   The complete Transition Plan shall include a timeline for adoption of the revised ISO standard, a list of software tools and training materials that should be modified or created, and an estimate of resource requirements.

    5.4.   Within one year of the approval of the Transition Plan by EDMC, the Metadata WG shall either develop and issue tools and training for beta-testing, or report to EDMC a schedule for completing that development and issuing the tools and training.

    5.5.   Within one year of the issuance of the tools in 5.4, the revised ISO standard shall become the default metadata standard for NOAA, and requirement 4.1.1 shall be updated accordingly.

    5.6.   Groups interested in migrating to 19115-1 more rapidly may contact the NOAA Metadata WG[12] for guidance.

---

[10] http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=39229

[11] http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=57104

[12] The Metadata WG can be contacted at noaa.enterprise.metadata@noaa.gov.

[13] http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=53798

[14] http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=32575

V. Responsibilities

1. **NOAA Programs[15] that produce or maintain on-line or off-line environmental data:** Ensure data are properly documented with metadata, and that metadata are registered with the NOAA Data Catalog, and that Extension or Waiver requests are filed if needed. NOAA Program Managers, or their designees, shall enforce the provisions of this Directive for the data they produce or maintain. Due date: See §IV.1.4.
2. **NOAA Catalog Working Group:** Support registration and harvesting of Metadata records. Due date: Ongoing.
3. **NOAA Metadata Working Group:** Develop Transition Plan for migration to revised ISO Geographic Metadata Standard. Due date: See §IV.5.2.
4. **NOAA Metadata Tool developers:** Update and issue for beta-testing tools which support revised ISO Geographic Metadata Standard. Due date: See §IV.5.4.
5. **NOAA EDMC:** Support implementation, tracking, and maintenance of this Directive. Compute metrics in §X. Review Transition Plan for migration to revised ISO Geographic Metadata Standard. Due date: Ongoing.

VI. Management and Ownership

This Procedural Directive is issued and managed by the NOAA Environmental Data Management Committee. The Directive and its Appendices will be reviewed and revised as needed on an occasional basis by the EDMC or a designated work team comprising members of the NOAA data management community.

VII. Intended Audience

The audience for this Directive includes data producers, data managers, program managers, service providers, archivists, metadata experts, and data center personnel.

VIII. Implementation Date

This Directive shall be effective as of 2017 Jan 01.

Due dates are listed in sections IV.1.4 and IV.5.

IX. Exemptions, Extensions, and Waivers

1. **Extensions** are temporary (up to 3 year) requests for additional time (beyond the 1 year period granted by Sec IV.1) to produce metadata for data that are in scope per Sec II.1. NOAA Programs may file an Extension Request with EDMC using the form available at https://nosc.noaa.gov/EDMC/PD.DD.php. A single waiver may be filed at the Program level to cover all relevant data within that Program.
2. *Waivers* are requests to permanently exclude certain classes of data from compliance with this Directive, even though such data would have been in scope per Sec II.1. NOAA Programs with legacy or experimental data as described below may file a Waiver Request with EDMC using the form available at https://nosc.noaa.gov/EDMC/PD.DD.php. A single waiver may be filed at the

---

[15] "NOAA Programs" is used here to refer to NOAA Programs, Projects, Activities, or other relevant sub-units at which budgets and responsibilities are tracked below the Line Office level.

Program level to cover all relevant data within that Program.[16] Metadata waivers are for:

    a. Legacy data that are no longer collected, <u>and</u> were never archived, <u>and</u> were not used to support the conclusion of a peer-reviewed publication or NOAA-issued publication, <u>and</u> were produced by NOAA Projects which no longer exist.

    b. Experimental or test data that have no utility beyond the immediate project <u>and</u> are not used to support the conclusion of a peer-reviewed publication or NOAA-issued publication.

3. <u>Exemptions</u>: Data ruled out of scope in Section II.3 are exempted from this Directive (no Waiver necessary).

X. Performance Objectives and Measurements

Three performance objectives have been established in order to ensure the completeness of NOAA's public data inventory is complete, 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 1: All NOAA datasets, whether on-line or off-line, have a publicly-accessible ISO Geographic Metadata record.
Metric 1.1: Total number of metadata records recorded by NOAA Data Catalog.
Metric 1.2: Percent of records describing off-line data.
Metric 1.3: Percent of records in FGDC format instead of ISO.
Metric 1.4: Number of requests filed for Extensions and Waivers filed per Sec. IX.

Objective 2: Metadata records in the NOAA Data Catalog include all required information and all relevant optional information (as defined in Appendix A).
Metric 2.1: Metadata completeness scores computed by assessment software (see Appendix A).
Metric 2.2: Additional statistics (to be determined) regarding metadata format, online access URLs, aggregation level, etc.

Objective 3: Metadata records in the NOAA Data Catalog identify the Program responsible for the record.
Metric 3.1: Number of records with Program affiliation specified.
Metric 3.2: Number of records per NOAA Program.

XI. Definitions and Abbreviations

Terminology

Environmental data: Recorded and derived observations and measurements of the physical, chemical, biological, geological, and geophysical properties and conditions of the oceans, atmosphere, space environment, sun, and solid earth, as well as correlative data such as socio-economic data, related documentation, and metadata. Numerical model outputs are included in this definition, and relevant model outputs should be made publicly accessible even if they are not archived. Digital audio or video recordings of environmental phenomena (such as animal sounds or undersea video) are included.

---

[16] Data whose very existence cannot be made public based on law, regulation, security classification, or NOAA policy must instead be covered by an Accessibility Waiver per the *NOAA Data Access Directive.[9]*

On-line data: Data that are accessible over the Internet, either directly or through an ordering mechanism (including near-line robotic tape drives).

Off-line data: Data recorded on paper, CD-ROM, magnetic tape, disk, or other media not accessible over the Internet.

Metadata: Descriptive information, structured according to agreed-upon standards for content and machine-readable encoding, about digital environmental data or research artifacts.

Abbreviations
DOI: Digital Object Identifier.
EDMC: NOAA Environmental Data Management Committee.
FGDC: US Federal Geographic Data Committee.
ISO: International Organization for Standardization.
NAO: NOAA Administrative Order.
NCEI: NOAA National Centers for Environmental Information.
NOAA: National Oceanic and Atmospheric Administration.
WAF: Web-Accessible Folder.[7]
XML: Extensible Markup Language.

XII. Frequently Asked Questions (FAQs)

Questions about this directive may be sent to edmc @ noaa.gov. Answers will be posted on the Data Documentation Directive page[17] of the NOAA Environmental Data Management (EDM) Wiki.

XIII. Approval

This Directive was approved on 2016 Nov 30 by the NOAA Environmental Data Management Committee.

_____

Dr. Jeff de La Beaujardière, Chair, NOAA EDMC

---

[17]   https://geo-ide.noaa.gov/wiki/index.php?title=Data_Documentation_PD

XIV. Appendices

Appendix A: Metadata Contents and Assessment Software

A complete metadata record typically includes information on the following topics:

- **Identification information** provides content needed for basic discovery. It includes the title, an abstract, theme keywords, point of contact, and scope.
- **Access information** describes the data formats, access points, distribution contact, 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 Digital Object Identifier (DOI).[18]

Whether specific fields are required or not depends in part on the nature of the data. For example, on-line access information would be omitted for research artifacts or data that are still in the planning stage; geographic bounds would be omitted for data that are not explicitly geospatial (e.g., solar and space weather observations). The NOAA Metadata Rubric described below is used to measure completeness of metadata records.

NOAA Metadata Assessment Tool

The ISO Geographic Metadata Standard provides a general content model and encoding. However, very little content is required to create a minimalist metadata record that is valid according to the ISO XML schema, so metadata elements beyond minimal ISO compliance are needed to properly support data discovery, use, and long-term stewardship. NOAA has therefore developed a metadata assessment tool as part of the Enterprise Metadata Management Architecture (EMMA)[19] project. The tool (sometimes referred to as a "Rubric") assesses the presence or absence of particular metadata fields to compute a score for each record. This helps metadata authors provide more thorough information about their data.

---

[18] NOAA Data Citation Directive (https://nosc.noaa.gov/EDMC/PD.DC.php) describes how to obtain and cite DOIs.

[19] http://www.ngdc.noaa.gov/metadata/emma/

Furthermore, scores are recorded in the EMMA database to help assess improvements over time in NOAA metadata. See the Metadata Rubric[20] page on the NOAA EDM Wiki for more information.

Examples of Complete ISO 19115-2 Metadata Records

When creating metadata it is often helpful to have an example to follow. As a guide to implementers, examples of high-scoring records will be posted on the Data Documentation Directive page[16] of the NOAA EDM Wiki.

More Information

The content within these documentation sections are the result of many conversations within the NOAA Metadata Working Group, [12] which regularly updates and documents its recommended practices within the ISO Explorer[21] on NOAA's Environmental Data Management Wiki. NOAA staff are welcome to contribute use cases and examples.

Appendix B: Metadata Support

The Metadata Tools[22] page on the NOAA EDM Wiki lists metadata support resources including information about training and tools.

---

[20]   https://geo-ide.noaa.gov/wiki/index.php?title=Completeness_Rubric
[21]   https://geo-ide.noaa.gov/wiki/index.php?title=Category:ISO_Explorer
[22]   https://geo-ide.noaa.gov/wiki/index.php?title=Category:Metadata_Tools