



NOAA Data Citation Procedural Directive

Table of Contents

I. Purposes.....	1
II. Scope.....	2
III. Authority.....	2
IV. Directive.....	2
V. Responsibilities.....	5
VI. Management and Ownership.....	5
VII. Intended Audience.....	6
VIII. Implementation Date.....	6
IX. Grandfather Exemption and Waiver Option.....	6
X. Performance Objectives and Measurements.....	6
XI. Definitions.....	6
XII. Frequently Asked Questions (FAQs).....	7
XIII. Approval.....	7
XIV. Appendices.....	8
Appendix A: Additional Guidance and Resources for Data Citation.....	8

I. Purposes

Data citation involves data stewards assigning unique and persistent identifiers to data in order that data users can reference those identifiers when data are used in research or as inputs to assessments, models, or derived products. Data citation is similar in many respects to the long-established convention of citing publications, and is an evolving but increasingly important scientific practice.

This Procedural Directive establishes the requirements for obtaining persistent identifiers for data archived at the NOAA National Centers for Environmental Information (NCEI),¹ the procedure for creating landing pages providing dataset information and access instructions, and the recommendations regarding citation of NOAA data by internal and external users. This Directive discusses the purpose and syntax of these identifiers, the appropriate level of granularity of a data collection, how to obtain an identifier, the content of the landing page, and the recommended data citation format.

¹ NCEI includes the organizations previously referred to as National Climatic Data Center (NCDC), National Geophysical Data Center (NGDC), and National Oceanographic Data Center (NODC).

The benefits of associating persistent identifiers with NOAA data are:

- To support traceability and scientific reproducibility through direct, unambiguous connection to the data used in derived products and results.
- To enable acknowledgement or credit for data producers, data distributors, and other contributors in the data production and stewardship process.
- To encourage submission of data and satisfactory metadata to the NOAA National Centers for Environmental Information for long-term preservation, as required by NOAA policy.
- To track data impact through references in scientific literature.
- To enable gathering of additional information regarding how data are being used.
- To improve NOAA's ability to catalog its data holdings.

II. Scope

This Directive applies to:

- Digitally formatted environmental data archived at NCEI.
- Publications archived at the NOAA Central Library.
- Data managers at NCEI.
- NOAA data producers or publication authors who wish to obtain an identifier.
- NOAA users of data to which an identifier has been assigned.

III. Authority

This Directive has been issued by the NOAA Environmental Data Management Committee (EDMC), with the approval of the NOAA Chief Information Officer (CIO) Council and the NOAA Observing System Council (NOSC), pursuant to the authorities granted in NOAA Administrative Order (NAO) 212-15, *Management of Environment Data and Information*.²

IV. Directive

1. Any interested party **may** request that a persistent identifier be assigned to a NOAA archival dataset.
 - 1.1. Requests to assign a Digital Object Identifier (DOI) shall be sent to the NOAA Data Citation Working Group at NOAA.Data.IDs@noaa.gov.
 - 1.2. Requesting an identifier is optional.
 - 1.3. Not all identifier requests will necessarily be granted.
 - 1.4. Identifiers may be assigned at a different level of granularity than requested.
 - 1.5. This Directive defines the syntax, eligibility, granularity, and other characteristics of NOAA

² http://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_212/212-15.html

dataset identifiers and their citation.

2. **Identifier syntax:** The DOI standard (ISO 26324³) shall be used for NOAA data identifiers.⁴
 - 2.1. NOAA shall maintain an agency-wide license from a DOI registrar to mint DOIs, and shall renew the license annually until this provision is rescinded. The annual cost of the NOAA DOI license shall be borne by NESDIS as the Line Office supporting the NCEI.⁵
 - 2.2. NOAA DOIs 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.
 - 2.2.1. NOAA DOIs shall not be structured to contain the names of organizations, programs, observing systems, or data.
 - 2.2.2. Only DOIs with the prefix 10.7289 shall be considered NOAA DOIs.⁶
3. **Eligibility:** Long-term preservation and sufficient documentation shall be required prior to receiving a NOAA DOI.
 - 3.1. Data shall be archived at NCEI prior to receiving a NOAA DOI.
 - 3.2. Data shall have a comprehensive metadata record that complies with the NOAA Data Documentation Procedural Directive⁷ prior to receiving a NOAA DOI.
4. **Identifier assignment:** NCEI shall designate personnel from its facilities who shall have the authority to issue and manage DOIs and to decline requests for DOIs if eligibility criteria are not met. This group of persons is referred to as the NOAA Data Citation Working Group.
 - 4.1. When a DOI has been assigned, NCEI shall update the metadata record to include the DOI and the recommended citation text.
 - 4.2. When a DOI has been assigned, NCEI personnel shall notify relevant individual(s) or organization(s) listed as creators or providers of the data.
 - 4.3. Any DOIs issued by a NOAA project under a separate license, even if that license was funded by NOAA, shall be considered project-specific DOIs rather than NOAA DOIs.⁸
 - 4.4. Datasets submitted for archiving at NOAA that have previously been assigned a DOI by another organization shall be handled as described or referenced in Appendix A.
5. **Identifier scope:** The level of granularity at which a DOI is assigned shall be at the discretion of NCEI, and shall be established in collaboration with data producers to the extent possible.
 - 5.1. DOIs shall generally be assigned at as coarse a level as possible. That is, data comprising many small components shall generally be assigned a single DOI rather than one per component.

³ ISO 26324:2012, Information and documentation -- Digital object identifier system (http://www.iso.org/iso/catalogue_detail?csnumber=43506)

⁴ 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.

⁵ Current cost \$1500/yr as of 2014 with California Digital Library EZID registrar. The license allows minting an unlimited number of DOIs per year. If the license is not renewed, existing DOIs remain valid but no new DOIs can be issued.

⁶ Example of a NOAA DOI already assigned: 10.7289/V5CR5R8P, wherein 10.7289 is the NOAA DOI prefix and V5CR5R8P is the suffix assigned to a particular dataset.

⁷ <https://www.nosc.noaa.gov/EDMC/PD.DD.php>

⁸ Data producers are strongly encouraged to use the NOAA DOI license in collaboration with NCEI.

- 5.2. Data components that differ only in time or in space (latitude, longitude, elevation, and depth) shall generally be aggregated under a single DOI (e.g., a single DOI would be assigned to measurements that are archived as individual monthly files).
- 5.3. 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.
- 5.4. Additional guidance referenced in Appendix A regarding granularity shall be followed when assigning DOIs.
6. **Landing pages:** NCEI shall ensure that each NOAA DOI resolves 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.
 - 6.1. Dataset DOIs shall not resolve directly to numeric data (that is, the DOI will reference a human readable landing page, which will then reference the actual numeric data).
 - 6.2. DOIs shall be resolvable using the mechanism established by the International DOI Foundation (<http://www.doi.org>).⁹
 - 6.3. NCEI shall update the landing page address with the DOI registrar when data are moved.
 - 6.4. Dataset landing pages shall be automatically produced from the metadata record by NCEI.
 - 6.5. Landing pages shall include a link to the actual metadata record required by 3.2.
 - 6.6. Landing pages shall include links or instructions for accessing the data (unless data have been retired from the archive).
 - 6.7. Landing pages shall include guidance on how to cite the data.
7. **Identifier permanence:** NOAA DOIs shall persist in perpetuity (more specifically, for as long as the NOAA National Centers for Environmental Information, or their successor organizations, exist).
 - 7.1. NCEI shall maintain the corresponding landing page in perpetuity.
 - 7.2. Data that have been retired from the archive shall continue to have a landing page, which shall include an explanation of the data removal.
8. **Publication identifiers:** The NOAA Central Library may, at their discretion and in consultation with document authors, issue DOIs for NOAA Technical Publications or other documents archived at the Library.
 - 8.1. A document must have sufficient bibliographic metadata, as defined by NOAA Central Library, prior to receiving a NOAA DOI.
 - 8.2. Document DOIs shall be assigned the object type "text" during the DOI registration process.
 - 8.3. Documents which have already received a DOI from another publisher (journal articles in particular) shall not be assigned a NOAA DOI.
 - 8.4. Document DOIs shall resolve to the actual document or to a landing page describing and linking to the document.
 - 8.5. NOAA Central Library shall ensure the resolvability and persistence of DOIs they assign.
9. **Citing data:** NOAA individuals and projects that use NOAA data, such as in a research paper, a

⁹ Example: the NOAA DOI 10.7289/V5CR5R8P can be resolved by following the URL <http://doi.org/10.7289/V5CR5R8P>, which leads to the current address of the landing page for that dataset.

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

- 9.1. If a DOI has been assigned, 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.¹⁰
- 9.2. If a DOI has not been assigned, the citation shall include at least the data name, data creator, data publisher, date of access, and subset used.¹¹
- 9.3. Non-NOAA users of NOAA data are strongly encouraged to cite the data used in the manner described above.

V. Responsibilities

- NOAA Data Citation Working Group: receive requests for DOI assignment; respond to questions; assign identifiers to selected data; notify data providers of newly-assigned DOIs;
- NOAA National Centers for Environmental Information: ensure long-term maintenance and availability of landing pages, metadata, and data.
- NOAA Central Library: Assign identifiers to selected publications; collect metrics on citation of NOAA DOIs in publications.
- NOAA data producers: If requesting dataset identifier(s), understand requirements in this Directive and receive assigned identifier(s).
- NOAA users of NOAA data: Cite data used in projects and results, including the assigned DOI if available. Inform data stewards (producers and NCEI) of data without identifiers about this Directive.
- NOAA users of non-NOAA data (to the extent possible): Cite external data in a manner similar to NOAA data, including identifier if available.
- NOAA Environmental Data Management Committee (EDMC): support implementation, tracking, and maintenance of this Directive.
- NESDIS headquarters: fund annual license for DOI issuance.

VI. Management and Ownership

This Procedural Directive is issued and managed by the NOAA Environmental Data Management Committee. The Directive will be reviewed at least every 3 years and revised as needed by the EDMC or a designated work team.

¹⁰ *Example: dataset 10.7289/V5CR5R8P includes this citation guidance: "Cite as: National Geophysical Data Center, British Geological Survey (2009): World Magnetic Model 2010. National Geophysical Data Center, NOAA. doi:10.7289/V5CR5R8P [access date]."*

¹¹ Data users are encouraged to request assignment of a DOI to data of interest by contacting the NOAA Data Citation Working Group at NOAA.Data.IDs@noaa.gov

VII. Intended Audience

This directive applies to NOAA data producers, to NOAA users of NOAA data, and to the NOAA National Centers for Environmental Information.

VIII. Implementation Date

This EDMC Procedural Directive shall take effect on 2015 June 1.

Not all archived data will immediately receive an identifier on that date; instead, NCEI will gradually assign identifiers to their holdings as needed based on available resources and prioritization. DOI assignment will initially be done manually for individual datasets, but may become automated as a routine step in archival submission. The level of automation will be at the discretion of NCEI.

IX. Grandfather Exemption and Waiver Option

There is no exemption to the core requirements for receiving an identifier (namely, archive submission and satisfactory metadata record). However, NCEI shall have flexibility to prioritize data for identifier assignment and to exclude low-priority data.

Data producers are not required to seek identifiers for their data.

Publications published (or submitted for publication) prior to the effective date of this Directive 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.

X. Performance Objectives and Measurements

Metrics regarding usage and citation of NOAA-issued data identifiers will be collected and reported.

- Effective immediately upon issuance of this directive, the number of NOAA DOIs assigned¹² shall be reported at least quarterly to EDMC and to the NOAA Office of the Chief Information Officer for inclusion in agency Open Government metrics reporting.
- Effective no more than 1 year after issuance of this directive, statistics regarding citation of NOAA DOIs in publications shall be collected by NOAA Central Library and shall be reported at least quarterly to EDMC.

XI. Definitions

Terminology from NAO 212-15

Environmental data: Recorded and derived observations and measurements of the physical, chemical,

¹² Information available automatically from <http://search.datacite.org/ui?q=10.7289> and <http://search.datacite.org/api?q=10.7289>.

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. Digital audio or video recordings of environmental phenomena (such as animal sounds or undersea video) are included in this definition.

Data Stewards: individuals who are responsible for establishing, maintaining, and being accountable for the quality, integrity, documentation, and preservation of environmental data under their purview.

Other definitions

Identifier: unique combination of letters, numbers or symbols based on an agreed-upon system or pattern. A data identifier is analogous to an International Standard Book Number (ISBN) or a Universal Product Code (UPC).

DOI: Digital Object Identifier.

NESDIS: National Environmental Satellite Data and Information Service.

XII. Frequently Asked Questions (FAQs)

Questions about this directive may be sent to the NOAA Data Citation Working Group at NOAA.Data.IDs@noaa.gov.

Answers will be posted on the NOAA Environmental Data Management Wiki site at https://geo-ide.noaa.gov/wiki/index.php?title=Data_Citation_PD.

XIII. Approval

This Directive (version 1.0) was approved on 2015 May 15 by the NOAA Environmental Data Management Committee.

Dr. Jeff de La Beaujardière, EDMC Chair

XIV. Appendices

Appendix A: Additional Guidance and Resources for Data Citation

Technical information regarding NOAA data identifier assignment and usage is recorded in the NOAA Environmental Data Management Wiki site at

https://geo-ide.noaa.gov/wiki/index.php?title=Data_Citation_PD.

Community Resources

- ESIP: http://wiki.esipfed.org/index.php/Interagency_Data_Stewardship/Citations/provider_guidelines
- DataCite.org: <http://datacite.org/whycitedata>
- CODATA: <http://www.codata.org/taskgroups/TGdatacitation/>
- The Dataverse Network Project: <http://thedata.org/citation>
- DOI Citation Formatter (beta): <http://crosscite.org/citeproc/>
- The DOI System: <http://www.doi.org/faq.html>
- DCC : <http://www.dcc.ac.uk/resources/how-guides/cite-datasets#fn14x0>