

**NOAA Technical Information Series NESDIS
DSMR-00303 Version 1.0**

doi: 10.25923/3h7r-b967



Data Stewardship Maturity Report for NCEI/WDS Natural Hazards Image Database

Table 1 Legend				
Level 1	Level 2	Level 3	Level 4	Level 5
Ad Hoc	Minimal	Intermediate	Advanced	Optimal
Little or no management	Limited Management	Defined Management, partially implemented	Well-defined Management, fully implemented	Full Management, audited, measured, controlled

Table 1. Scores for the Nine DSMM Key Components at a Glance		
Preservability - 5	Accessibility - 1.66	Usability - 2.5
Production Sustainability - 3.66	Data Quality Assurance - 2	Data Quality Control/Monitoring - 2
Data Quality Assessment - 1	Transparency/Traceability - 2.5	Data Integrity - 1

NOAA National Centers for Environmental Information January 2020



U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Environmental Satellite, Data, and Information Service

Cover Image: Data Stewardship Rating Diagram for NCEI/WDS Natural Hazards Image Database

Shades of green are used to represent level 1 through level 5 ratings; denoting Ad Hoc, Minimal, Intermediate, Advanced, and Optimal stages for each of the nine key components, respectively. The dark green level indicates all the practices are completely satisfied. The lighter green levels indicate only some of the practices are satisfied. The lightest green level indicates none of the practices are satisfied.

The stewardship maturity of NCEI data product, NCEI/WDS Natural Hazards Image Database, is assessed based on a reference stewardship maturity framework. The current maturity ratings of NCEI/WDS Natural Hazards Image Database are at Level 1 or higher for all nine key components with three Level 1, four Level 2, one Level 3, zero Level 4, and one Level 5 key components.

The National Environmental Satellite, Data, and Information Service (NESDIS) manages the Nation's civil Earth-observing satellite systems, as well as global national data bases for meteorology, oceanography, geophysics, and solar-terrestrial sciences. From these sources, it develops and disseminates environmental data and information products critical to the protection of life and property, national defense, and the national economy, energy development and distribution, global food supplies, and the development of natural resources.

Publication in the NOAA Technical Memorandum series does not preclude later publication in scientific journals in expanded or modified form. The NESDIS series of NOAA Technical Reports is a continuation of the former NESS and EDIS series of NOAA Technical Reports and the NESC and EDS series of Environmental Science Services Administration (ESSA) Technical Reports.

Copies of earlier reports may be available by contacting NESDIS Chief of Staff, NOAA/ NESDIS, 1335 East-West Highway, SSMC1, Silver Spring, MD 20910, (301) 713-3578.

ASSESSMENT REVISION HISTORY

Revision	Description	Date
V01r00	Initial Release	02/02/2022

NOAA Technical Information Series NESDIS DSMR-00303

Version 1.0

doi: 10.25923/3h7r-b967

Data Stewardship Maturity Report for NCEI/WDS Natural Hazards Image Database

Frederick Gunther

NOAA's National Centers of Environmental Information (NCEI)

151 Patton Avenue, Asheville, NC 28801, (828) 271-4800

Recommended Citation

Frederick Gunther. (2021), Data Stewardship Maturity Report for NCEI/WDS Natural Hazards Image Database, NOAA Technical Information Series NESDIS DSMR-00303 Version 1.0, 19pp., doi: 10.25923/3h7r-b967

Table of Contents

List of Tables	7
Preface	8
1. Introduction	9
2. Results	10
3. Acknowledgment	14
4. References	15
Appendix I The Scientific Data Stewardship Maturity Matrix (DSMM)	16

List of Tables

Table 1. Scores for the Nine DSMM Key Components at a Glance	1
Table 2. Dataset and Data Stewardship Maturity Assessment Metadata	11
Table 3. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the Dataset	12

Preface

In response to the President's Open Government Initiative and related policies, NOAA has committed to providing improved public access to all of its environmental information, to enable research and commercial innovation through ease of data discovery and use [Casey, 2016].

OneStop supports NOAA's efforts by leveraging existing access technologies and infusing specific innovations to provide improved discover, access, and visualization services for NOAA's data. Also, OneStop is viewed by a NESDIS as a pathfinder effort with an initial focus on selected high-priority datasets from NESDIS and other program data meeting OneStop standards, but eventually scalable across NOAA's data. Lastly, OneStop is implementing the USGEO Common Framework for Earth Observation Data and leveraging/supporting the NOAA Big Data Project (BDP) and Big Earth Data Initiative (BEDI) [Casey, 2016].

As with any process of improvement planning, agencies need to find out where they are in terms of their compliance to the federal regulations and what they need to do if any areas of non-compliance are identified. To this end, a unified framework would be beneficial for assessing the current stage of stewardship practices applied to individual datasets and for providing a road map that will guide future investments towards enhanced stewardship of environmental datasets. The value and quality of a dataset depends in part on the stewardship practices applied after its development and production. Therefore, a unified framework providing a holistic view of the quality of stewardship practices applied to individual datasets is beneficial to data stewards and users [Casey, 2016].

The Data Stewardship Maturity Matrix (DSMM), jointly developed by domain (data management, technology, and science) subject matter experts from NOAA's National Centers for Environmental Information (NCEI) and Cooperative Institute for Climate and Satellites – North Carolina (CICS-NC), provides such a consistent framework [Peng *et al.*, 2016]. The DSMM, leveraging institutional knowledge and community practices and standards, defines a graduated maturity scale for each of nine key components of scientific data stewardship to enable a consistent assessment of the measureable stewardship practices applied to a given data set or product.

The NOAA Data Stewardship Maturity Technical Series captures stewardship maturity assessment results for individual datasets, provides consistent representation and citable documents of those assessments, ensures transparency, and allows better data quality information integration and content-based search and discovery of NOAA data.

Data Stewardship Maturity Report for NCEI/WDS Natural Hazards Image Database

1. Introduction

1.1 Purpose

The purpose of this document is to describe the results of stewardship maturity assessment for NOAA Climate Data Record for Mean Layer Temperature (Upper Troposphere & Lower Stratosphere from UCAR, Version 2, utilizing the Scientific Data Stewardship Maturity Matrix or DSMM [Peng, et al, 2016]. DSMM defines levels of stewardship maturity stages for Preservability, Accessibility, Usability, Production Sustainability, Data Quality Assurance, Data Quality Control/Monitoring, Data Quality Assessment, Transparency/Traceability, and Data Integrity key components. Each of these components is ranked from ‘Ad hoc’ to ‘Optimal’ (see Appendix I). This report is based on evaluation performed by NOAA OneStop metadata specialists working with Subject Matter Experts and utilizing the DSMM template [Peng, 2016].

1.2 Scope

Assessing stewardship maturity - the current state of how datasets are documented, preserved, stewarded, and made accessible publicly, is a critical step towards meeting U.S. federal regulations, organizational requirements, and user needs [Peng et al., 2016]. The goal of this document is to provide consistent and transparent stewardship maturity information to data users and decision-makers.

1.3 Dataset Abstract

Photographs and other visual media provide valuable pre- and post-event data for natural hazards. Research, mitigation, and forecasting rely on visual data for post-analysis, inundation mapping and historic records. Instrumental data only reveal a portion of the whole story; photographs explicitly illustrate the physical and societal impacts from an event. This resource provides high-resolution geologic and damage photographs from natural hazards events, including earthquakes, tsunamis, slides, volcanic eruptions and geologic movement (faults, creep, subsidence and flows). The earliest images date back to 1886. Each event also links to NCEI's Global Historical hazards databases, which provide details for these events.

1.4 Document Maintenance

This document is generated and maintained by NOAA's National Centers for Environmental Information. More on policy is available at <https://www.ncei.noaa.gov/>.

2. Results

The data stewardship maturity assessment information is summarized in Table 1. Each component is displayed along with its corresponding score in a color-coded table.

Table 2. Dataset and Data Stewardship Maturity Assessment Metadata

Dataset Title	NCEI/WDS Natural Hazards Image Database
Dataset Information URL	https://doi.org/10.7289/V5154F01
Data Provider POC (Name; Email; Affiliation)	National Centers for Environmental Information, NESDIS, NOAA, U.S. Department of Commerce 301-713-3277 NCEI. Info@noaa.gov
Dataset POC (Name; Email; Affiliation)	Hazards Data Manager; haz.info@noaa.gov; NOAA National Centers for Environmental Information
SMM Version (Document ID and Version Number)	NCDC-CICS-SMM_0001_Rev.1 12/09/2014
SMM POC (Name; E-mail; Affiliation)	Ge Peng, ge.peng@uah.edu, University of Alabama- Huntsville
SMM Template Version (Document ID and Version Numbers)	NCDC-CICS-SMM_0001_Rev.1 v4.0 06/23/2015
SMM Template POC	Ge Peng, ge.peng@uah.edu, University of Alabama- Huntsville
SMM Assessment Version (v<nn>r<mm>, e.g., v01r00)	V01r00
SMM Assessment Date (MM/DD/YYYY)	02/01/2022
SMM Assessment POC (Name; E-mail; Affiliation)	Frederick Gunther; frederick.gunther@noaa.gov; Riverside Technology
Stewardship Maturity Ratings (each key component) (kc1/kc2/kc3/kc4/kc5/kc6/kc7/kc8/kc9)	5/1.66/2.5/3.66/2/2/1/2.5/1
SMM Original Assessment Date (MM/DD/YYYY)	02/01/2022
SMM Original Assessment POC (Name; E-mail; Affiliation)	Frederick Gunther; frederick.gunther@noaa.gov; Riverside Technology
SMM Last Modified Date (MM/DD/YYYY)	02/01/2022
SMM Last Modification POC (Name; E-mail; Affiliation)	Frederick Gunther; frederick.gunther@noaa.gov; Riverside Technology
SMM Modified Date (MM/DD/YYYY)	02/01/2022
SMM Modification POC (Name; E-mail; Affiliation)	Frederick Gunther; frederick.gunther@noaa.gov; Riverside Technology

Table 3. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the Dataset.

DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
Preservability	<p>Level 5</p> <ul style="list-style-type: none"> ▪ Conforms to NCEI archive guidelines which are OAIS RM and NARA compliant. Conforms to ISO 19115-2 metadata standards. ▪ Plans in place to upgrade to newer ISO 19115-1 metadata standard. ▪ Products managed per the submission agreement (SA). ▪ Archive procedures and processes are managed and Trustworthy Digital Repositories (TDR) audit in place. ▪ Annual reviews per the SA. ▪ Future archiving standard changes planned <p>Comments:</p>
Accessibility	<p>Level 1.66</p> <ul style="list-style-type: none"> ▪ Collection level searchable online: https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ngdc:Hazard_Images_Database ▪ Granule level is searchable online ▪ https://www.ngdc.noaa.gov/hazardimages/#/ ▪ Additional search options available from collection level site <p>Comments:</p> <p>Enhanced data server performance Dissemination report metrics defined and implemented internally through requests and inquiries to haz.info@noaa.gov Granule/file level searchable</p>
Usability	<p>Level 2.5</p> <ul style="list-style-type: none"> ▪ NCEI data sets use community standard-based interoperable format and metadata ▪ Metadata landing page and data, products and services for natural hazards can be found here ▪ https://www.ncei.noaa.gov/metadata/geoportal/rest/metadata/item/gov.noaa.ngdc:Hazard_Images_Database/html ▪ https://www.ngdc.noaa.gov/hazard/ ▪ Dataset is also accessible from NCEI Site ▪ https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ngdc:Hazard_Images_Database ▪ Limited supporting documentation available here ▪ https://www.ngdc.noaa.gov/hazard/ ▪ and here https://www.ngdc.noaa.gov/hazel/view/about <p>Comments:</p>
Production Sustainability	<p>Level 3.66</p> <ul style="list-style-type: none"> ▪ The dataset is currently supported and up to date ▪ The project is maintained by NCEI and funded by NWS ▪ Project is long term and has institutional commitment through NCEI <p>Comments:</p>

Table 3. Stewardship Maturity Levels and Detailed Justifications for Each of Nine DSMM Key Components for the Dataset.	
DSMM Key Component	Stewardship Maturity Rating, Justification, and Comments
Data Quality Assurance	<p>Level 2</p> <ul style="list-style-type: none"> ▪ Data quality assurance (DQA) procedure ad hoc. not defined or documented <p>Comments: Limited data quality assurance within ISO metadata. Data Quality tag exists within XML</p>
Data Quality Control/ Monitoring	<p>Level 2</p> <ul style="list-style-type: none"> ▪ Limited product-specific metrics defined & implemented (lat/long, photographer credit, date/time, captions) <p>Comments: Limited data quality assurance within ISO metadata. Data Quality tag exists within XML</p>
Data Quality Assessment	<p>Level 1</p> <ul style="list-style-type: none"> ▪ No algorithm document or product assessment found <p>Comments:</p>
Transparency / Traceability	<p>Level 2.5</p> <ul style="list-style-type: none"> ▪ Product information available in literature: ▪ [April, 2019]. https://doi.org/10.1007/s00024-019-02191-7 ▪ [December, 2019] https://ui.adsabs.harvard.edu/abs/2019AGUF.MNH43F0992A ▪ Dataset DOI assigned: https://doi.org/10.7289/V5154F01 <p>Comments:</p>
Data Integrity	<p>Level 4</p> <ul style="list-style-type: none"> ▪ Data archive integrity verifiable - Checksum technology is available, each NCEI/WDS_NHID package is accompanied by a manifest in XML format containing hash digests generated using various algorithms, including MD5, SHA-1, SHA-384, etc. That includes checksums (.md5) for every file package. https://www.nodc.noaa.gov/archive/arc0025/0061664/0061664.1.1.xml ▪ Data authenticity is verifiable (since data can be downloaded via HTTPS and HTTPS uses certificates to prove site authenticity) ▪ NCEI-MD does not provide digital signatures for data dissemination <p>Comments:</p>

3. Acknowledgment

This work is supported by the NOAA OneStop Project.

We thank the dataset POCs for their valuable input, as well as the collaborative efforts of the OneStop teams, especially the Metadata team. We would also like to show appreciation to Ge Peng for her contributions.

The draft of this data stewardship maturity report is systematically generated by a tool created by Kieran Hodnett and populated with the stewardship maturity assessment done by the author(s) of this report. The tool was developed based on a Word template created collaboratively by Robert Partee II, Raisa Ionin, Paul Lemieux III, Ge Peng, Don Collins, and Sonny Zinn with helpful input from the NOAA Central Library and the NCEI Communication Team.

4. References

Casey, K. (2016), The NOAA OneStop data discover and access framework project, Version:June 3, 2016. <https://cdn.ioos.noaa.gov/media/2017/12/OneStop-IOOS-DMAC-03-June-2016.pdf>

Peng, G. (2015) The scientific data stewardship maturity assessment model template, Version: NCDC-CICS-SMM-0001-Rev.1 v4.0 6/23/2015. doi:10.6084/m9.figshare.1211954.

Peng, G., J.L. Privette, E.J. Kearns, N.A. Ritchey, and S. Ansari (2015), A unified framework for measuring stewardship practices applied to digital environmental datasets, *Data Science Journal*, 13, 231-253, doi: 10.2481/dsj.14-049.

Peng, G., J. Lawrimore, V. Toner, C. Lief , R. Baldwin, N. Ritchey, D. Brinegar, and S. A. Delgreco (2016) assessing stewardship naturity of the global historical climatology network-monthly (GHCN-M) dataset: use case study and lessons learned, *D-Lib Magazine*, 22, doi:10.1045/november2016-peng.

Arcos, N.P., Dunbar, P.K., Stroker, K.J. et al. The Impact of Post-tsunami Surveys on the NCEI/WDS Global Historical Tsunami Database. *Pure Appl. Geophys.* 176, 2809–2829 (2019)

Past, present, and future: Interface for the NCEI/WDS Global Historical Tsunami Database

Appendix I: The Scientific Data Stewardship Maturity Matrix (DSMM)

Table A1: This matrix (Version: NCDC-CICS-SMM-0001-Rev.1. 12/09/2014) describes the criterion used to evaluate data stewardship maturity for each of the nine DSMM key components [Peng *et al.*, 2015].

DSMM Component	Level 1 <i>Ad hoc</i> Little or no management	Level 2 <i>Minimal</i> Limited management	Level 3 <i>Intermediate</i> Defined management, partially implemented	Level 4 <i>Advanced</i> Well-defined management, fully implemented	Level 5 <i>Optimal</i> Full management, audited, measured, controlled
<i>Preservability</i> <i>(The state of being preservable)</i>	Any storage location Data only	Non-designated repository Redundancy Limited archiving metadata	Designated archive Redundancy Community-standard archiving metadata Conforming to limited archiving standards	Level 3 + Conforming to community archiving standards	Level 4 + Archiving process performance controlled, measured, and audited Future archiving standard changes planned
<i>Accessibility</i> <i>(The state of being searchable and accessible publicly)</i>	Not publically available person-to-person	Publically available direct file download (e.g., via anonymous FTP server) Collection or dataset level searchable online	Level 2 + Non-standard data service Limited data server performance Granule/file level searchable Limited search metrics	Level 3 + Community-standard data service Enhanced data server performance Conforming to community search metrics Dissemination report metrics defined and implemented internally	Level 4 + Dissemination reports available online Future technology and standard changes planned

<p>Usability</p> <p><i>(The state of being easy to use)</i></p>	<p>Extensive product-specific knowledge required</p> <p>No documentation online</p>	<p>Non-standard data format</p> <p>Limited documentation (e.g., user's guide online)</p>	<p>Community standard-based interoperable format & metadata</p> <p>Documentation (e.g. source code, product algorithm document, processing or/and data flow diagram) online</p>	<p>Level 3 +</p> <p>Basic capability (e.g., subsetting, aggregating) & data characterization overall/global,</p> <p>e.g., climatology, error estimates) available online</p>	<p>Level 4 +</p> <p>Enhanced online capability (e.g., visualization, multiple data formats)</p> <p>Community metrics of data characterization (regional/cell) online</p> <p>External ranking</p>
<p>Production Sustainability</p> <p><i>(The state of data production being sustainable and extendable)</i></p>	<p>Ad Hoc or Not applicable</p> <p>To obligation or deliverable requirement</p>	<p>Short-term</p> <p>Individual PI's commitment (grant obligations)</p>	<p>Medium-term</p> <p>Institutional commitment (contractual deliverables with specs and schedule defined)</p>	<p>Long-term</p> <p>Institutional commitment</p> <p>Product improvement process in place</p>	<p>Level 4 +</p> <p>National or international commitment</p> <p>Changes for echnology planned</p>
<p>Data Quality Assurance</p> <p><i>(The state of data quality being assured)</i></p>	<p>Data quality assurance (DQA) procedure unknown or none</p>	<p>Ad Hoc and random</p> <p>QA procedure not defined and documented</p>	<p>DQA procedure defined and documented and partially implemented</p>	<p>DQA procedure well documented, fully implemented and available online with master reference data</p> <p>Limited data quality assurance metadata</p>	<p>Level 4 +</p> <p>DQA procedure monitored and reported</p> <p>Conforming to community quality metadata & standards</p> <p>External review</p>

<p>Data Quality Control/Monitoring</p> <p><i>The state of data quality being controlled and monitored</i></p>	<p>None or Sampling unknown or spotty</p> <p>Analysis unknown or random in time</p>	<p>Sampling and analysis are regular in time and space</p> <p>Limited product-specific metrics defined & implemented</p>	<p>Level 2 +</p> <p>Sampling and analysis are frequent and systematic but not automatic</p> <p>Community metrics defined and partially implemented</p> <p>Procedure documented and available online</p>	<p>Level 3 +</p> <p>Anomaly detection procedure well-documented and fully implemented using community metrics, automatic, tracked and reported</p> <p>Limited quality monitoring metadata</p>	<p>Level 4 +</p> <p>Cross-validation of temporal & spatial characteristics</p> <p>Physical consistency check</p> <p>Conforming to community quality metadata & standards</p>
<p>Data Quality Assessment</p> <p><i>(The state of data quality being assessed)</i></p>	<p>Algorithm/method/model</p> <p>Theoretical basis assessed (methods and results online)</p>	<p>Level 1 +</p> <p>Research product assessed (methods and results online)</p>	<p>Level 2 +</p> <p>Operational product assessed (methods and results online)</p>	<p>Level 3 +</p> <p>Quality metadata assessed</p> <p>Limited quality assessment metadata</p>	<p>Level 4 +</p> <p>Assessment performed on a recurring basis</p> <p>Conforming to community quality metadata & standards</p> <p>External ranking</p>
<p>Transparency/Traceability</p> <p><i>(The state of being transparent, trackable, and traceable)</i></p>	<p>Limited product information available</p> <p>Person-to-person</p>	<p>Product information available in literature</p>	<p>Algorithm Theoretical Basis Document (ATBD) & source code online</p> <p>Dataset configuration managed (CM)</p> <p>Unique Object Identifier (OID) assigned (dataset, documentation, source code)</p> <p>Data citation tracked (e.g., utilizing Digital Object Identifier</p>	<p>Level 3 +</p> <p>Operational Algorithm Description (OAD) online, OID assigned, and under CM</p>	<p>Level 4 +</p> <p>System information online</p> <p>Complete data provenance online</p>

<p>Data Integrity</p> <p><i>(The state of data integrity being verifiable)</i></p>	Unknown or no data ingest integrity check	Data ingest integrity verifiable (e.g., checksum technology)	(DOI) system) Level 2 + Data archive integrity verifiable	Level 3 + Data access integrity verifiable Conforming to community data integrity technology standard	Level 4 + Data authenticity verifiable (e.g., data signature technology) Performance of data integrity check monitored and reported