Assessment Information

[CoreTrustSeal Requirements 2017–2019](#)

Repository:	World Data Service for Paleoclimatology
Website:	https://www.ncdc.noaa.gov/data-access/paleoclimatology-data
Certification Date:	17 February 2020

This repository is owned by: NOAA National Centers for Environmental Information

World Data Service for Paleoclimatology

Notes Before Completing the Application

We have read and understood the notes concerning our application submission.

True

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

Background & General Guidance

Glossary of Terms

BACKGROUND INFORMATION

Context

R0. Please provide context for your repository.

Repository Type. Select all relevant types from:

National repository system; including governmental

Brief Description of Repository

The National Centers for Environmental Information (NCEI) is the core repository of the United States National Oceanic and Atmospheric Administration (NOAA) for environmental data, and is one of the US Government's core repositories for geophysical data more broadly. NCEI's data holdings extend "from the surface of the sun to the bottom of the oceans, and from over 60 million years ago into the future".

NOAA/NCEI is a component of the United States Department of Commerce (DOC), which is the responsible funding agency at the US departmental level.

Within NCEI, there are two primary data and science centers, the Center for Weather and Climate (CWC) and the Center for Coasts, Oceans, and Geophysics (CCOG). These centers are funded to provide user-oriented scientific product development and data center functions and services across a broad range of scientific and general user communities that produce and utilize weather, climate, and other geophysical information.

NCEI holds paleoclimate data assets from many sources under the auspices of the Paleoclimate Program within CWC's Climate Science Branch (hereafter "NOAA WDS-Paleo"). Approximately 50% of the data collections are from US National Science Foundation funded research projects, and the rest are from a wide range of national and international contributors whose funding comes largely from national or international (e.g., the European Union) research councils. Data products produced by NOAA WDS-Paleo research and value-added products it produces are also housed within the repository.

Brief Description of the Repository's Designated Community.

The designated community for NOAA WDS-Paleo consists of a wide range of users of primary paleoclimate proxy data, derived paleoclimate reconstructions, and (more limited) paleoclimate models. Users are scientific researchers, decision makers, national, state, and regional/local authority organizations (e.g., ecosystem and water managers), along with education, media, consultant, and other users. NOAA WDS-Paleo has a close working relationship with PAGES (the Past Global Changes programme of Future Earth, formerly the International Geosphere-Biosphere Programme), which is the key international coordinating body for paleoclimate research and information dissemination. PAGES and NOAA WDS-Paleo were formed in conjunction with each other in the early 1990's, and there is a formal "Science Partner" agreement between our organizations. NOAA WDS-Paleo also produces federated searches for data housed at NEOTOMA, the leading paleoecological data repository whose products can also be used for paleoclimate research, and at PANGAEA, one of the world's richest repositories of geophysical data.

Reference:

1. Science Partner Agreement:

https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/PAGES-WDS-Paleo-NCEI-Partnership-2019.pdf

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

Level of Curation Performed. Select all relevant types from:

D. Data-level curation – as in C above; but with additional editing of deposited data for accuracy

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

Comments

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

Outsource Partners. If applicable, please list them.

Not applicable.

Other Relevant Information.

NOAA WDS-Paleo operates the World Data Service for Paleoclimatology (WDS-Paleo) under formal agreement between NOAA and the World Data System (WDS) of the International Council for Science (ICSU). The CTS certification is a component of NOAA WDS-Paleo's recertification with the WDS.

WDS-Paleo provides the world's primary facility for ingest, quality control, archive, and state-of-the-art search and retrieval of paleoclimate proxy information, along with climate reconstructions derived from the proxy data. These products and capabilities are used on an ongoing basis by thousands of professional scientists, decision makers, environmental managers, educators, media, consultants, and the members of the general public, as noted above. In this context, WDS-Paleo holdings act as a source of critically important scientific data for inclusion in future science projects, programs, applications, systems, products, and decision support systems. WDS-Paleo data facilitate a large number of scientific peer-review processes each year.

NOAA WDS-Paleo serves the same functionalities for the International Tree Ring Data Base (ITRDB), which is integrated within the WDS-Paleo system overall, as is the International Multi-Proxy Paleo-Fire Database (IMPD). WDS-Paleo also acts as the primary data hub for the Last 2000-Year Project of PAGES.

There is a range of laws, protocols, standards, policies, strategic directives, management plans, etc. under which NOAA WDS-Paleo operates, which underpin the functions and services it delivers. These are outlined in detail in requirements 1, 2, 3, 5, 6, 10, 15, and 16. WDS-Paleo operations encompass all elements in the Open Archive Information System (OAIS) model, and are focused on conformance with the "FAIR"(3) data management principles (Findable, Accessible, Interoperable, Reusable).

These arrangements highlight the centrality of the NOAA WDS-Paleo in relation to NOAA's and NCEI's missions to provide the highest quality information on climate variability. Paleoclimate data are the only source of such information prior to the deployment of widespread instrumental measurement beginning in the late 19th century and satellite

measurement beginning in the second half of the 20th century. Thus paleoclimatology forms an indispensable "fourth leg" of NOAA's climate measurement "stool" (along with satellite, surface and air instrument, and model data): effort to archive and facilitate science discovery via paleoclimate data products and to improve the geographic and temporal coverage, skill, and precision of paleoclimate information represent fundamental data management and applied science at the heart of NOAA's and NCEI's missions.

NOAA WDS-Paleo additionally develops primary paleoclimate science in support of these missions: in particular, highlighting hydroclimate and related temperature variability in light of the sensitivity of human and ecological systems to water availability and usage constraints in support of NOAA's hydroclimate variability mission focus.

The primary web page for NOAA WDS-Paleo can be found at https://www.ncdc.noaa.gov/data-access/paleoclimatology-data , and the search and retrieval capabilities page is at https://www.ncdc.noaa.gov/paleo-search/ .

References:
1. https://www.ncdc.noaa.gov/data-access/paleoclimatology-data
2. https://en.wikipedia.org/wiki/Open_Archival_Information_System
3. https://www.nature.com/articles/sdata201618

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

ORGANIZATIONAL INFRASTRUCTURE

I. Mission/Scope

R1. The repository has an explicit mission to provide access to and preserve data in its domain.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

NOAA/NESDIS/NCEI falls under the US Department of Commerce (DOC), a cabinet-level department of the Executive Branch of US Government. NOAA Administrative Order 212-15 provides the highest level policy for NOAA data management. NAO 212-15 states that it is the responsibility of the NOAA Environmental Data Management Committee (EDMC) to identify, develop and approve Procedural Directives associated with 212-15.

The EDMC mission statement as presented on its official website is that the committee coordinates the development of NOAA's environmental data management strategy and policy, and provides guidance to promote consistent implementation across NOAA, on behalf of the NOAA Observing Systems Council (NOSC) and NOAA Office of the Chief Information Officer (CIO) Council. Environmental data management is an end-to-end process that includes acquisition, quality control, validation, reprocessing, storage, retrieval, dissemination, and long-term preservation activities. The goal of the EDMC is to enable NOAA to maximize the value of its environmental data assets through sound and coordinated data management practices.

The EDMC Data Management Planning Procedural Directive effective 01 Jan 2015 mandates all NOAA repositories manage their data holdings in accordance with documented lifecycle management. This directive specifically states its applicability to NCEI (ref Sec 2.0). Responsibilities of NCEI under this directive include securing funding for its data management mission.

While the EDMC provides functional oversight of the NCEI data preservation mission, administratively, NCEI is a budget line office of NOAA's National Environmental Satellite, Data and Information Service (NESDIS). The NESDIS 5-Year Strategic Plan, dated 2016, states its mission to provide secure and timely access to global environmental data.

NESDIS further documents its commitment to sound stewardship of the nation's environmental data holdings by the publication of the NESDIS Environmental Data Management Planning Policy (NDP 6010.0A), dated 10 Feb 2017 with the signature of the NOAA Assistant Administrator for Satellite and Information Services. Section 3.2 of NPD 6010.0A states the NOAA National Center for Environmental Information (NCEI) constitutes the official NOAA archive that maintains, processes, distributes, and provides long-term stewardship for most of NOAA's environmental and geospatial data, and provides a broad range of user services. While NCEI is operated by NESDIS, the center performs data preservation and stewardship on behalf of the entire agency. Section 5.9 of NDP 6010.0A further states NCEI is responsible for all Archival functions under the directive.

NCEI hosts three World Data Service members, including WDS-Paleo, and as such shares the overall objectives of

ICSU-WDS, which are defined in its Constitution as follows:

- Enable universal and equitable access to scientific data, data services, products and information.

- Ensure long-term data stewardship - Foster compliance to agreed-upon data standards and conventions.

- Provide mechanisms to facilitate and improve access to data and data products.

References:

1. NOAA Administrative Order 212-15:

https://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_212/212-15.pdf

2. NOAA Environmental Data Management Committee (EDMC) website: https://nosc.noaa.gov/EDMC/

3. NOAA Observing Systems Council (NOSC): https://nosc.noaa.gov/

4. NOAA Office of the Chief Information Officer (CIO) Council: https://www.cio.noaa.gov/

5. EDMC Data Management Planning Procedural Directive: https://nosc.noaa.gov/EDMC/PD.DMP.php

6. NESDIS 5-Year Strategic Plan:

https://www.nesdis.noaa.gov/sites/default/files/asset/document/the_nesdis_strategic_plan_2016.pdf

7. NESDIS Environmental Data Management Planning Policy:

https://www.nesdis.noaa.gov/sites/default/files/asset/document/npd_6010_01a.pdf

8. NCEI World Data Service Members: https://www.ncdc.noaa.gov/customer-support/world-data-centers

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

II. Licenses

R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

NCEI, as an agency of the US Government that is a signatory of the World Meteorological Organization (WMO) protocol, adheres to WMO Resolution 40 that governs the exchange of meteorological and related data. Therefore, the only condition for use of NCEI data as depicted on its Customer Service webpage is: 'NCEI data and products that contain international data may have conditions placed on their international commercial use. They can be used within the United States or for noncommercial international activities without restriction. Redistribution of these data by others must provide this same notification. The non-U.S. data cannot be redistributed for commercial purposes.'

However, restrictions on data access and use are rare. As stated in para 4.9 of NESDIS Environmental Data Management Planning Policy (NDP 6010.0A) to the maximum extent possible, the repository will ensure that all its data are publicly discoverable and accessible pursuant to the NOAA Data Access Procedural Directive.

References:

1. WMO Resolution 40: www.wmo.int/pages/prog/hwrp/documents/wmo_827_enCG-XII-Res40.pdf

2. NCEI Conditions of Use Policy: https://www.ncdc.noaa.gov/customer-support

3. NESDIS Environmental Data Management Planning Policy:

https://www.nesdis.noaa.gov/sites/default/files/asset/document/npd_6010_01a.pdf

4. NOAA Data Access Procedural Directive: https://nosc.noaa.gov/EDMC/PD.DA.php

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

III. Continuity of access

R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

Since the passing by the US Congress of the Federal Records Act of 1950, NCEI along with its forerunners has been designated by the US National Archives and Records Administration (NARA) as the official US Government repository for environmental information with the requirement under law to provide access, preservation and storage of its holdings. For in-situ and remotely sensed environmental data, NARA requires preservation of "master data" (original data) for 75 years according to NOAA Records Schedule 1406-01.b.

Administrative oversight of NCEI falls under its parent organization, NESDIS. As depicted in the NESDIS Environmental Data Management Planning Policy and NESDIS 5-Year Strategic Plan, NESDIS is committed to the medium (3-5 years) and long term (>5 years) funding of the NCEI data preservation mission and does not anticipate cessation or significant decrease of funding to this mission as mandated by law.

References:

1. Federal Records Act of 1950: https://en.wikipedia.org/wiki/Federal_Records_Act

2. NOAA Records Schedule 1406-01.b: https://www.corporateservices.noaa.gov/audit/records_management/schedules/

3. NESDIS Environmental Data Management Planning Policy:
https://www.nesdis.noaa.gov/sites/default/files/asset/document/npd_6010_01a.pdf

4. NESDIS 5-Year Strategic Plan:
https://www.nesdis.noaa.gov/sites/default/files/asset/document/the_nesdis_strategic_plan_2016.pdf

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

IV. Confidentiality/Ethics

R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical

norms.

Compliance Level:

0 – Not applicable

Response:

Data archived with NOAA WDS-Paleo are fully public upon completion of data and metadata processing and hosting on the WDS-Paleo webservice. See https://www.ncdc.noaa.gov/customer-support/world-data-centers , and R10.

Disclosure risk is not an applicable consideration. The ethical context is full availability including identification of data authors and contributors.

V. Organizational infrastructure

R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Response:

The physical repository infrastructure on which NOAA WDS-Paleo is reliant is 'base-funded'. The US Congress commitment to the NCEI preservation mission is evident in past base funding appropriations for NCEI as follows:

Fiscal Year 2015: $68.0 million
Fiscal Year 2016: $59.0 million
Fiscal Year 2017: $59.2 million
Fiscal Year 2018: $60.6 million
Fiscal Year 2019: $60.4 million

Beyond FY 2019, however, it should be noted that the US Department of Commerce generally doesn't present funding requests beyond the next fiscal year.

Within NCEI, NOAA WDS-Paleo is staffed with 6 Full Time Equivalents (FTE) fully devoted to its mission. The FTEs cover both data management and science, along with administration, and are a 'blended' workforce meaning it is a mix of civil servants (with tenure), contract employees and onsite university affiliates. Labor costs for the core team are approximately $965K/year, of which approximately $635K is base-funded within NCEI with the remainder externally funded by the NOAA Office of Oceanic and Atmospheric Research (OAR). This external OAR funding source has been consistent over the past several years and is expected to remain so in the future. Apart from the core NOAA WDS-Paleo team, there is a partial FTE from the NCEI general staff of 0.3 devoted to data management that is base-funded.

The NOAA WDS-Pleo staff is competently trained for its mission of data stewardship of a paleoclimatology archive. All of the staff hold academic science or computer science degrees including two domain PhDs. In addition, the staff is technically trained in areas conducive to maintaining a digital repository. Technical training includes advanced Oracle Application Express (https://apex.oracle.com/en/) and Javascript Object Notation (JSON). The staff is also trained in the ISO metadata standards specified by NOAA. Lastly, two of the staff received official CTDRR training in 2017.

Reviewer 2

Comments:
Accept

VI. Expert guidance

R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either inhouse or external, including scientific guidance, if relevant).

Compliance Level:

4 – The guideline has been fully implemented in the repository

Response:

The repository is linked to a wide network of expertise for securing guidance and feedback in the areas of technology and science. Foremost guidance is provided by the NOAA Environmental Data Management Committee (EDMC). As stated on its website, the EDMC coordinates the development of NOAA's environmental data management strategy and policy, and provides guidance to promote consistent implementation across NOAA, on behalf of the NOAA Observing Systems Council (NOSC) and NOAA Chief Information Systems Officer Council (CIO). Environmental data management is an end-to-end process that includes acquisition, quality control, validation, reprocessing, storage, retrieval, dissemination, and long-term preservation activities. The goal of the EDMC is to enable NOAA to maximize the value of its environmental data assets through sound and coordinated data management practices.

NOAA WDS-Paleo maintains connections with objective expert advice beyond NOAA via its membership in, or coordination with, the following entities:
- World Data System (https://www.icsu-wds.org/)
- United States National Science Foundation (NSF) (https://www.nsf.gov/)
- Past Global Changes (PAGES) (http://pastglobalchanges.org/)
- Research Data Alliance (https://www.rd-alliance.org/)
- International Tree-Ring Data Bank (https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/tree-ring)

- International Multi-Proxy Paleo-Fire Database
(https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/fire-history)
- Linked Earth Project (http://linked.earth/)

WDS-Paleo consults with related WDS member repositories PANGAEA and Neotoma, and constitutes ad-hoc expert advisory panels for guidance on specific matters as needed. An example of the output of the latter is the new Paleoenvironmental Standard Terms (PaST) Thesaurus, developed with the input of 25 subject matter experts across paleoclimatology disciplines. PaST was partially supported by the NSF through a grant to NCEI affiliate partner the University of Colorado, to provide expert guidance in library science for the thesaurus' logical structure and specialist informatics programming expertise.

References:

1. NOAA Environmental Data Management Committee: https://nosc.noaa.gov/EDMC/
2. NOAA Observing Systems Council: https://nosc.noaa.gov/
3. NOAA Chief Information Systems Officer Council: https://www.cio.noaa.gov/cio_council_committees.html
4. PANGAEA: https://www.pangaea.de/
5. Neotoma: https://www.neotomadb.org/
6. Paleoenvironmental Standard Terms (PaST) Thesaurus:
https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/past-thesaurus

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

DIGITAL OBJECT MANAGEMENT

VII. Data integrity and authenticity

R7. The repository guarantees the integrity and authenticity of the data.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer 1

Comments:
3 – The repository is in the implementation phase

Reviewer 2

Comments:
3 – The repository is in the implementation phase

Response:

WDS-Paleo is in the implementation phase with regards to fixity checks for Submission Information Packages (SIPs) and data access. In regards to archival storage, the WDS-Paleo Submission Agreement details the policy followed with regard to the implementation and use of MD5 checksums for the archival storage WDS-Paleo data. Access to this agreement is restricted due to the information technology security information it contains and is not allowed to be shared outside the US Government.

Provenance data are maintained in the metadata. Metadata for every dataset of the WDS-Paleo repository in multiple formats (JSON, DIF, ISO) are located here: https://www1.ncdc.noaa.gov/pub/data/metadata/published/paleo/

The identity of depositors is authenticated during the ingest process via familiarity with data producers through NOAA WDS-Paleo's ongoing user engagement and collaboration with the paleo community, subject matter expertise in the field, and peer-reviewed publications associated with the datasets.

Publications associated with data contributions are used to authenticate the identity of the contributor by cross referencing the dataset. For examples of this practice, please reference the publication cited in the following data contribution landing pages:
https://www.ncdc.noaa.gov/paleo/study/1902
https://www.ncdc.noaa.gov/paleo/study/1870

Changes to data and metadata for individual records are documented in the metadata for that record, in the "Study Notes" field. A more general tabular format for documenting all such changes in one place is in the implementation phase.

Reviewer 1

Comments:
Accept.

The procedures should be made publicly available at the next certification to reach Compliance Level 4.

Reviewer 2

Comments:
Accept

VIII. Appraisal

R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

Data contributions are appraised against the NOAA/NESDIS/NCEI Data Collecting Policy. This policy defines the type of data that will be acquired for long-term preservation and states limitations that may affect acceptance of environmental data at NCEI. NCEI archives data in accordance with the NOAA Procedure for Scientific Records Appraisal and Archive Approval and the scope of the NCEI archives. Priority is given to original environmental data over any products derived from such data. NCEI considers the following characteristics in its archive decisions:

- NOAA or NESDIS mandate

- Relevance and criticality to NOAA or NESDIS mission

- Archive-ready data (well structured and adheres to established NCEI standards)

- Broad user base

- Relevance to NCEI development of use-inspired products

- Multiple users (i.e., research, monitoring, relevant to a variety of sectors/users)

- Uniqueness

- Completeness

- Intrinsic value

As per the policy, Paleoclimatology is among the two dozen information product categories acquired by NCEI.

References:

1. NOAA/NESDIS/NCEI Data Collecting Policy: https://www.nodc.noaa.gov/submit/data-policy.html

2. NOAA Procedure for Scientific Records Appraisal and Archive Approval:

https://www.ngdc.noaa.gov/wiki/images/0/0b/NOAA_Procedure_document_final.pdf

WDS-Paleo quality control checks include assessment of the completeness and understandability of data and metadata contributed as well as compliance with required data/metadata formats. Quality control is performed by WDS-Paleo's staff of three Data Managers, who are subject matter experts in the 18 types of paleoclimate data stewarded by WDS-Paleo.

The required use of recommended data and metadata formats, in the form of Data Contribution Templates, and the use of tree ring community-specific data formats, is documented in the WDS-Paleo Data Contribution Guidelines. In addition to general required metadata (provenance, identification, site, temporal and spatial characteristics, content URL, etc.), the templates encapsulate metadata that describe what was measured in terms of the PaST thesaurus/controlled vocabulary. This latter requirement and related search capabilities are new to WDS-Paleo as of 2018. Explanations and links to all data and metadata contribution requirements are accessible from the WDS-Paleo Data Contribution Guidelines.

As part of the WDS-Paleo workflow, quality control software, peer-reviewed journal manuscripts associated with the contribution, and visual inspection are used to assess data and metadata quality. Details about these checks are described in Section R12 of this document.

ISO 19115 metadata is the NOAA standard metadata format. WDS-Paleo generates an ISO 19115 metadata record for every WDS-Paleo dataset (in addition to NASA DIF and JSON standard formats). To ensure that metadata meets the NOAA standards for quality and completeness, the NCEI Collection Manager Metadata Tools are employed.

These tools help scientists and data managers create, manage, quality assure, display and improve their standards-compliant documentation. The tools are used to analyze the completeness and quality of all WDS-Paleo ISO 19138 metadata including:

- Diagnostic reports and visualization of the metadata metrics over time
- Measurements of completeness and quality of metadata content

The Collection Manager Metadata Tools also contain and employ the XML Schema for ISO19139 Metadata. Note: ISO 19139:2012 (International Organization for Standardization (2012-12-15)) provides the XML implementation schema for ISO 19115.

Large-scale synthesis contributions including metadata and data are becoming more prevalent, especially from the projects of our PAGES partner organization. To ensure the quality and completeness of metadata and data early on, as specified in the WDS-Paleo Data Contribution Guidelines* and PAGES/WDS-Paleo announcement*, WDS-Paleo requests that leaders of projects have their designated liaison contact WDS-Paleo staff early during research and development in order to coordinate on formats and requirements.

Preferred metadata and data formats are published in the NCEI Data Collecting Policy and the WDS-Paleo Data Contribution Guidelines*

References:

1. WDS-Paleo Data Contribution Guidelines: https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/contributing

2. Paleoenvironmental Standard Terms (PaST) Thesaurus:

https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/past-thesaurus

3. NCEI Collection Manager Metadata Tools:

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

4. XML Schema for ISO19139 Metadata: https://data.noaa.gov/resources/iso19139/schema.xsd

5. International Organization for Standardization (2012-12-15):

https://www.iso.org/obp/ui/#iso:std:iso:ts:19139:-2:ed-1:v1:en

6. Past Global Changes (PAGES) (http://pastglobalchanges.org/)

7. PAGES/WDS-Paleo announcement:

http://www.pages-igbp.org/news/all-news-items/9-latest-news/1937-announce-noaa-wds-paleo

8. NCEI Data Collecting Policy: https://www.nodc.noaa.gov/submit/data-policy.html

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

IX. Documented storage procedures

R9. The repository applies documented processes and procedures in managing archival storage of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

NESDIS Environmental Data Management Planning Policy, NPD 6010.01A (para. 3.2) specifies that NCEI define archival functions in accordance with OAIS-RM, namely Ingest, Archival Storage, Access, Data Management, Preservation Planning, Administration, and Data Stewardship.

Reference:
1. NESDIS Environmental Data Management Planning Policy:
https://www.nesdis.noaa.gov/sites/default/files/asset/document/npd_6010_01a.pdf

Relevant processes and procedures for the archive are maintained in a relational database document management system. These documents are under version control with any proposed change requiring approval by the appropriate governance board.

Incoming data are scanned in a 'DMZ' outside of the archive's security boundary firewall for threats. Also, there is always a security representative on any Archive Appraisal or Dataset Readiness Review to provide a risk assessment on any proposed new dataset. Security related processes and procedures are documented with a System Security Plan (SSP). The SSP is a sensitive document that cannot be shared outside the US Government.

Section 4 of NESDIS Environmental Data Management Planning Policy, reaffirms the requirement in NOAA Administrative Order 212-15 that data management is an end-to-end process requiring a long-term archival stewardship. The policy further declares NESDIS will provide sufficient development, operations and maintenance resources to support NCEI's archive function pursuant to the NOAA Environmental Data Management Framework.

References:
1. NESDIS Environmental Data Management Planning Policy:
https://www.nesdis.noaa.gov/sites/default/files/asset/document/npd_6010_01a.pdf
2. NOAA Administrative Order 212-15:
https://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_212/212-15.pdf
3. NOAA Environmental Data Management Framework: https://nosc.noaa.gov/EDMC/framework.php

The repository's backup strategy is to keep two copies of all data on tape. Copies are then transported to an offsite climate-controlled facility several kilometers away from the main physical repository. This process is depicted in NCEI Standard Operating Procedure 60-6-015 (https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/20160928-MSC_SOP_AppendixM-NCDC200-01-018_IT_Off-site_Backup_Procedures.pdf).

Documented data recovery provisions are in place as depicted in NCEI Standard Operating Procedure 60-6-015 (https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/20160928-MSC_SOP_AppendixM-NCDC200-01-018_IT_Off-site_Backup_Procedures.pdf).

The repository's backup strategy is to keep two copies of all data on tape. This two-tape system is the risk mitigation

strategy.

The repository loads a md5 checksum associated with the data file into the Oracle digital catalog. Any time a file is pulled either from primary or offsite, the file's checksum is checked against the catalog. Every month the repository runs scripts that compare file sizes between the primary and offsite to detect either mismatches or if one of the copies is missing. New ingest software is being phased in at the repository to further automate comparisons with more frequency.

The repository has dedicated Database Administrators (DBA) who monitor and manage tape deterioration through software associated with the tape library system that provides continuous metrics on tape health and degradation. There is also a periodic tape refresh every few years budgeted for in recurring Operations and Maintenance (O&M) costs.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

X. Preservation plan

R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

The repository has published a NCEI Digital Archive Data Management Plan dated 30 Aug 2017 (https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/ncei-data-management-plan-aug2017-ver-1.0.pdf) that

serves as the long-term preservation plan.

To define preservation levels, NCEI has developed and published Tiers of Data Stewardship as shown in Figure 2 of the NCEI Digital Archive Data Management Plan (https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/ncei-data-management-plan-aug2017-ver-1.0.pdf). These tiers range from basic long-term preservation and access (Tier 1) to the highest order of National Service or International Leadership (Tier 6) for the data.

Submission agreements (SAs) are signed between depositor and NCEI that clearly state the roles and responsibilities of each party.

All WDS-Paleo data are covered by a single encompassing SA with the main NCEI repository with stated mutual obligations between both parties throughout the life cycle of each dataset.

NCEI makes clear to depositors the repository's adherence to NOAA Administrative Order 212-15, which states that environmental data is of unrestricted access unless explicitly restricted by law, regulation, policy or explicit signed agreement between the depositor and NCEI.

Reference:
1. NOAA Administrative Order 212-15:
https://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_212/212-15.pdf

Figure 1 of NCEI's Digital Archive Data Management Plan depicts the process flow of identifying, appraising, implementing and maintaining data collections. Section 8 of this management plan specifies data preservation and protection policy and actions, in compliance with NOAA Environmental Data Management Committee policy (see paragraph immediately following). Additionally, preservation standards are found on the NCEI website at https://www.ncei.noaa.gov/archive.

Reference:
1. NCEI's Digital Archive Data Management Plan:
https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/ncei-data-management-plan-aug2017-ver-1.0.pdf
2. NOAA Environmental Data Management Committee policy:
https://nosc.noaa.gov/EDMC/documents/NOAA_Procedure_document_final_12-16-1.pdf

NOAA maintains a standing governing Environmental Data Management Committee (EDMC) to ensure continuing adherence to NCEI's long-term preservation responsibilities that are in accordance with NOAA policies and directives.

Reference:
1. NOAA Environmental Data Management Committee: https://nosc.noaa.gov/EDMC/

XI. Data quality

R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Response:

Data quality is ensured as part of the NOAA WDS-Paleo workflows described in the Section R12. For the wide variety of paleoclimate data deposited, quality assurance for metadata and data is performed by NOAA WDS-Paleo data managers in collaboration with subject matter experts (see also Sections R8 and R6).

External quality control of the data, consistent with NOAA guidelines, occurs in the peer-review publication process prior to the metadata and data deposition with the repository. In the cases of contributions from the International Tree-Ring Data Bank
(which comprise about 50% of NOAA WDS-Paleo's holdings), and the International Multi-Proxy Paleo-Fire Database, these communities have given explicit direction for additional quality control checks to be done. These checks are performed using tools and quality metrics that these communities have provided, which are detailed in Section R12.

References:

1. International Tree-Ring Data Bank: https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/tree-ring

2. International Multi-Proxy Paleo-Fire Database:

https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/fire-history

As documented in Section R12, ingest includes storing all dataset metadata in the NOAA WDS-Paleo Oracle metadata database. At the end of the ingest process for a specific dataset, and as changes to metadata for sets of datasets are needed, standard metadata formats are output from the WDS-Paleo Oracle metadata database to the public NOAA WDS-Paleo web-accessible folder. These metadata formats include: ISO 19115, and NASA DIF standard metadata formats, and a NOAA WDS-Paleo tailor-designed JSON metadata format (which contains all information contained in the ISO 19115 and NASA DIF formats, in addition to full PaST thesaurus terms which describe what was measured).

Automated assessment of ISO 19115 metadata is performed via the NOAA Collection Manager Metadata Tools as described in Section R8. These tools are used to analyze the completeness and quality of all WDS-Paleo ISO 19115 metadata including diagnostic reports and visualization of the metadata metrics over time and measurements of completeness and quality of metadata content

The Collection Manager Tools also contain and employ the XML Schema for ISO19139 Metadata. Note: ISO 19139:2012 provides the XML implementation schema for ISO 19115 (International Organization for Standardization (2012-12-15)).

References:

1. NOAA Collection Manager Metadata Tools:

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

2. XML Schema for ISO19139 Metadata: https://data.noaa.gov/resources/iso19139/schema.xsd

3. ISO 19115 (International Organization for Standardization (2012-12-15)):

https://www.iso.org/obp/ui/#iso:std:iso:ts:19139:-2:ed-1:v1:en

Since its inception in 1990, NOAA WDS-Paleo has fostered and maintains ongoing relationships with communities which distribute or plan to distribute data via NOAA WDS-Paleo. These collaborations focus on community-driven, data-centered activities involving data stewardship best practices, which include the design and assessment of metadata and data formats and content quality, highlights of which follow. Flagship (since the early 1990's) ongoing collaborative partnerships with The International Tree-ring Data Bank and Past Global Changes (PAGES) (as detailed in Section 2 of the Scientific Partnership between PAGES and NOAA WDS-Paleo) focus on dendrochronological data, and multi-disciplinary paleo data products respectively. The International Multi-proxy Paleo-fire Database collaboration, which began in 2004, focuses on multiple aspects of paleo fire event-based data. Two recently formed collaborations, Linked Paleo Data (LiPD) (since 2016) and Paleo-Event Data Standards (since 2017) are community frameworks being designed to simplify the sharing, reuse and analysis of paleoclimate data.

Linked Paleo Data (LiPD) is a community open-ended metadata and data format, which is a hierarchical container that encapsulates metadata and data, and is part of the Linked Earth Project (LE). The LiPD Playground (web-based software) includes tools to create and quality assure LiPD. WDS-Paleo is collaborating with LiPD investigators to develop a NOAA

WDS-Paleo LiPD Standard which meets the requirements of both NOAA WDS-Paleo and LE, and to incorporate it into the LiPD Playground, NOAA Ready (Beta) release. The goal is to allow users to simultaneously create, quality assure to NOAA WDS-Paleo and LE requirements, and generate both the required NOAA WDS-Paleo Template and the corresponding LiPD format files.

In the Fall of 2017, NOAA WDS-Paleo took part in the Past Global Changes Workshop: Paleo-Event Data Standards for Dendrochronology, collaborating with over a dozen dendrochronologists from five nations to address the Past Global Changes Data Stewardship Integrative Activity. The goal was to develop a general data model for dendrochronological-event data including data and metadata structures to promote best practices of data stewardship. In October, 2018 an application was submitted to propose a PAGES Working Group for Dendro Events, co-authored by workshop leaders from the University of Arizona and the US Forest service, NOAA WDS-Paleo, and with the support of the workshop participants and other international researchers.

References:

1. The International Tree-ring Data Bank:

https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/tree-ring

2. Past Global Changes (PAGES): http://pastglobalchanges.org/

3. Scientific Partnership between PAGES and NOAA WDS-Paleo:

https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/PAGES-WDS-Paleo-NCEI-Partnership-2019.pdf

4. International Multi-proxy Paleo-fire Database:

https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/fire-history

5. Linked Earth Project (LE): http://linked.earth/

6. LiPD Playground: http://lipd.net/playground

7. Past Global Changes Workshop: Paleo-Event Data Standards for Dendrochronology:

http://www.pastglobalchanges.org/products/11525

8. Past Global Changes Data Stewardship Integrative Activity: http://pastglobalchanges.org/ini/int-act/data-stewardship

Citations to works related to each NOAA WDS-Paleo dataset are documented in all of its metadata record formats (ISO 19115, NASA DIF, JSON), which are available from the NOAA WDS-Paleo Metadata Web Accessible Folder (WAF).

As noted in Section R8, metadata and data contribution is a collaborative endeavor between WDS-Paleo and the contributor. If metadata is insufficient or not contributed in an accepted format, the WDS-Paleo Data Manager works with the contributor, and where possible consults the associated publication to resolve the issues. Details about this are described in Section R12.

References:
1. ISO 19115: https://www1.ncdc.noaa.gov/pub/data/metadata/published/paleo/iso/xml/
2. NASA DIF: https://www1.ncdc.noaa.gov/pub/data/metadata/published/paleo/dif/xml/
3. JSON: https://www1.ncdc.noaa.gov/pub/data/metadata/published/paleo/json
4. NOAA WDS-Paleo Metadata Web Accessible Folder (WAF):
https://www1.ncdc.noaa.gov/pub/data/metadata/published/paleo/

XII. Workflows

R12. Archiving takes place according to defined workflows from ingest to dissemination.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Response:

NOAA Paleoclimatology has policies and procedures in place to cover the lifecycle of the data from the pre-ingestion phase to digital preservation.

Data ingest begins the NOAA WDS-Paleo "Contributing Data" web page and the PaST Thesaurus. All data acquisitions follow the contribution policy and use the PaST set of standardized terms to describe the variables measured.

Data contributors in the process of publishing articles based on the data contributed can request that their data be held off the public server until their article is accepted by a journal.

The data ingest workflow varies slightly by data type. Contributions to the International Tree-ring Data Bank (ITRDB) follow this workflow:
https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/impd-tree-based-workflow.txt

Contributions to the International Multiproxy Paleofire Database (IMPD) follow this workflow:

https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/impd-tree-based-workflow.txt

All other data types follow this workflow:

https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/Paleo-ingest-general-workflow.txt

After data templates for all data types have been validated, the metadata is entered into the WDS-Paleo Oracle database via the PINGMAN data entry application, or via

batch ingest. Using an ingest tool that connects directly to the database allows for the ingest process to be guided by dynamically generated lists of valid options and through the use of database constraints to check incoming entries. After the metadata ingest is complete, the WDS-Paleo database is used to create an ISO-19139 record for each study. This ISO record is then quality controlled using NCEI's Collection Manager Metadata Tools. The text template is put into a NCEI FTP directory for access by external users.

All data ingested into the NOAA/WDS Paleo database receive an internal unique identifier. DOIs are also being minted for legacy and incoming data sets housed at WDS-Paleo. Minting of DOIs for legacy data is 50% complete at this point. All identifiers assigned at the point of the original ingest will persist for the lifetime of the original data contribution within the WDS-Paleo archive.

Data are transferred to NCEI's Archive Branch and archived following NCEI's Archiving Guidelines (see also R9 and R10). The Archive Branch follows the OAIS-RM standard as described in this document published in the 2007 IEEE International Geoscience and Remote Sensing Symposium at https://ieeexplore.ieee.org/abstract/document/4423732.

References:

1. NOAA WDS-Paleo "Contributing Data" web page:

https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/contributing

2. Past Global Changes (PAGES) Thesaurus:

https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/past-thesaurus

3. International Multiproxy Paleofire Database (IMPD):

https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/fire-history

4. NCEI's Collection Manager Metadata Tools:

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

5. NCEI's Archiving Guidelines: https://www.ncdc.noaa.gov/atrac/guidelines.html

The methods and procedures for communicating with contributors are outlined in the Workflow section above.

Data contributors in the process of publishing articles based on the data contributed can request that their data be held off the public server until their article is accepted by a journal. There are no human subject data (see R4).

The different methods for quality control of incoming data are outlined in the Workflow section above.

Data are typically associated with the publication of a peer-reviewed article prior to metadata and data deposition with the repository. As noted in R11, and above in the Workflows/business process descriptions component of this response, contributions to the International Tree Ring Data Bank and the International Multi-proxy Paleoenvironmental Database are additionally guided by those communities, including additional quality control checks that are done. Data are also selected by: data managers with guidance from the NOAA/NESDIS/NCEI Data Collecting Policy which defines the type of data that will be acquired for long-term preservation; NOAA/WDS Paleo domain scientists; journal editors and reviewers; and other scientific user communities, such as Past Global Changes (PAGES).

References:
1. NOAA/NESDIS/NCEI Data Collecting Policy: https://www.nodc.noaa.gov/submit/data-policy.html
2. Past Global Changes (PAGES) Thesaurus:
https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/past-thesaurus

Data that do not fall within NOAA/WDS Paleo's mission are not archived. If possible, the data contributor is directed to a more appropriate archive for their data.

The different workflows for different data types are outlined in the Workflow section above.

Decisions on data handling are made by WDS-Paleo data managers and scientists in consultation with data contributors. Oversight of tree ring data is provided by the ITRDB Advisory Committee. Oversight on fire history data is provided by the IMPD Advisory Board.

Changes to the WDS-Paleo data workflow are handled through meetings of the Data Managers and the WDS-Paleo scientists. Any changes to the ingest process are then made on the WDS-Paleo "Contributing Data" web page and in the workflow description documents.

References:
1. NOAA WDS-Paleo "Contributing Data" web page:
https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/contributing

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

XIII. Data discovery and identification

R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

NCEI offers search and download of Paleoclimatic proxy and reconstruction data from the WDS-Paleo archives. Over 10,000 data sets are available, derived from natural sources such as tree rings, ice cores, corals, and ocean and lake sediments.

The WDS-Paleo Graphical User Interface (https://www.ncdc.noaa.gov/paleo-search/) and web service provide a General Search ("Google-type") box and also Advanced Search capabilities to improve filtering by investigators, locations, keywords, time periods, and PaST variable definitions. Data files for the entire result set, or a subset of it, can be bundled and downloaded as a single compressed file.

The WDS-Paleo Interactive Map allows users to select study locations by region, proximity to a point, or text-based attributes. Through a set of Google Earth Files (http://gis.ncdc.noaa.gov/kml/paleo.kmz), users can locate and download paleoclimatology studies using a Google Earth map interface.

NOAA OneStop, NOAA/NCEI Geoportal and Catalog Service for the Web international standard (OGC CSW) search facilities (as specified in the WDS-Paleo Data Management Plan section 6.3) provide access to all WDS-Paleo holdings (https://www1.ncdc.noaa.gov/pub/data/paleo/data_management/ncei-paleoclimatology-data-management-plan.pdf).

International agreed standards are employed by NOAA-WDS Search capabilities as follows:

WDS-Paleo GUI, web service, and Interactive map searches (detailed above) return ISO and NASA DIF Metadata. NOAA

OneStop and Geoportal search facilities (detailed above) search and return WDS-Paleo's ISO metadata records. The NOAA/NCEI Catalog Service for the Web international standard (OGC CSW) (detailed above) returns WDS-Paleo's ISO metadata records.

All of these search capabilities employ NASA GCMD scientific and locational keyword vocabularies (https://earthdata.nasa.gov/about/gcmd/global-change-master-directory-gcmd-keywords).

As detailed in the NOAA WDS-Paleo Data Management Plan section 6.3, NCEI provides machine harvesting of WDS-Paleo metadata through the Web Catalog Service for the Web international standard (OGC CSW)

NOAA-WDS Paleo's archive holdings are included in the following registries:

PAGES recommends WDS-PALEO as an archive for the deposit of data produced for new PAGES projects in the PAGES Data and links web page: PAGES Data and Links web page (http://pastglobalchanges.org/my-pages/data).

Datacite Repository Selector (https://repositoryfinder.datacite.org) returns WDS-Paleo as a re3data (https://www.re3data.org/) repository and a repository which supports Findable, Accessible, Interoperable, Re-usable (FAIR) principals via the following respective web service API calls:
https://repositoryfinder.datacite.org/search?open=&pid=&query=paleoclimatology&sort=relevance&subject=
https://repositoryfinder.datacite.org/search?open=true&pid=true&query=paleoclimatology&sort=relevance&subject=34

NASA Global Change Master Directory Search (https://gcmd.gsfc.nasa.gov/KeywordSearch/Keywords.do?Portal=GCMD &KeywordPath=Parameters%7CHome&MetadataType=0&Columns=0) returns WDS-Paleo via the following web service API call:
https://gcmd.gsfc.nasa.gov/search/Titles.do?subset=GCMD&search=paleoclimatology&Portal=&KeywordPath=&Freetext =&MetadataType=#titles

ICSU World Data System returns WDS-Paleo as a Community Member via the following web service API call:
https://www.icsu-wds.org/community/membership/regular-members/@@member_view?fid=paleoclimatology-branch-noa as-national-climatic-data-center

Every WDS-Paleo dataset's metadata record and landing page (for example: https://www.ncdc.noaa.gov/paleo/study/10437) includes recommended citation information which provide credit and attribution to individuals who contributed to the creation of the dataset, and requests the user to "Please cite original publication, online resource, dataset and publication DOIs (where available),and date accessed when using downloaded data. If there is no publication information, please cite investigator, title, online resource, and date accessed".

For every WDS-Paleo dataset, a unique online resource URL, which resolves to the dataset's landing page, is formulated using a distinct identifier assigned by the Paleo Oracle Metadata Database (for example: https://www.ncdc.noaa.gov/paleo-search/study/21970 , where 21970 is the distinct identifier).

WDS-Paleo is in process of minting DOIs for every dataset. As of July 1, 2019, 51% of existing datasets have been assigned a DOI. Citations to the DOI will be included in dataset metadata records and landing pages (for example: https://www.ncdc.noaa.gov/paleo-search/study/search.json?xmlId=19948 , and https://www.ncdc.noaa.gov/paleo-search/study/21970).

The number of datasets for which DOIs have been minted is growing daily and the current status can be viewed at: https://www1.ncdc.noaa.gov/pub/data/metadata/published/paleo/doi/paleo-doi.csv

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

XIV. Data reuse

R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

Through collaborative partnerships with fellow international WDS Members PANGAEA and Neotoma, WDS-Paleo promotes re-use of paleoclimate data beyond those of the WDS-Paleo archive. Federated searches of the Neotoma and

PANGAEA archives can be accessed via the "Data Publisher" section of the WDS-Paleo search (https://www.ncdc.noaa.gov/paleo-search/), thus allowing the user to query for holdings of PANGAEA, Netoma, and NOAA WDS-Paleo archives in any combination simultaneously.

Required metadata are contributed via a WDS-Paleo recommended Data Contribution Template as defined in the WDS-Paleo Data Contribution Guidelines (https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/contributing). The content of the set of metadata required in each of the WDS-Paleo Data Contribution Templates is a synergistic union of content of the ISO 19115 metadata format required by NOAA as per the NOAA Archiving Guidelines https://www.ncdc.noaa.gov/atrac/guidelines.html , NASA DIF as a standard within the wider international earth science community, and the PaST Thesaurus to describe what was measured.

As described in a recent article in a PAGES Magazine special edition PaST is a new advance at WDS Paleo which promotes understanding and thereby re-use of the diverse holdings of WDS-Paleo beyond paleoclimate specialists. PaST-enhanced web-service search capabilities additionally enable aggregation of WDS-Paleo's small, long-tailed datasets into larger, standardized collections. This can facilitate large-scale data syntheses, which is a key thrust in Paleoclimatology (e.g., PAGES 2k Consortium (2017) Sci Data 4: 170088)

References:
1. PAGES Magazine special edition: http://pastglobalchanges.org/products/pages-magazine/12712
2. PAGES 2k Consortium (2017) Sci Data 4: 170088: https://doi.org/10.1038/sdata.2017.88

Community specific standardized data formats are provided by the International Tree Ring Data Bank, the International Multi-proxy Paleofire Database, and net CDF (Unidata). Specifics are described in the R12. Workflow section. These formats ensure data reuse, data interoperability, and the application of standardized quality assurance tools developed by the specific community. The NASA DIF format is a common standard used by the international paleo community including PAGES, PANGAEA, and WDS-Paleo, and has been used as a common metadata format for federated searches described above.

References:
1. International Tree Ring Data Bank: https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/tree-ring
2. International Multi-proxy Paleofire Database:
https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/fire-history
3. net CDF (Unidata): https://www.unidata.ucar.edu/software/netcdf

Metadata formats and content areas including discovery, geospatial, provenance, and the semantic web are evolving and this effort is ongoing across a number of standards development organizations.
(https://www.lter.uaf.edu/metadata_files/UnderstandingMetadata.pdf NISO Press, ISBN 1-880124-62-9)
To address the evolution of metadata over time, all metadata information is stored in the WDS-Paleo Oracle metadata database, and output via database processes, which incorporate XSLT, to specific standard formats as necessary.

Currently three metadata formats are output from the WDS-Paleo's Oracle metadata database (ISO, NASA DIF, and the WDS-Paleo JSON format), and the database model is poised to be enhanced and extended with new content areas, if and as necessary, in order to output new metadata formats in the future.

In keeping with the NCEI Archiving Guidelines, File Format section, WDS-Paleo uses open data formats maintained by standards organizations as opposed to proprietary or product-specific formats, in order to facilitate re-use, interoperability and long-term access (Linux information project: http://www.linfo.org/free_file_format.html).

References:

1. Metadata formats: https://www1.ncdc.noaa.gov/pub/data/metadata/published/paleo/

2. NCEI Archiving Guidelines, File Format section: https://www.ncdc.noaa.gov/atrac/guidelines.html

As detailed in R11. Data Quality, a NOAA-standard for the Linked Paleo Data (LiPD) format will enhance data reuse over time as LiPD is used more by the community.

WDS Paleo ensures understandability of its diverse data holdings via complete and high quality metadata documentation that enables discovery, understanding, and usage of the data; which is key to re-use and facilitates sharing and data citation. The process used by WDS-Paleo to ensure complete and high-quality metadata is detailed in R8.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

TECHNOLOGY

XV. Technical infrastructure

R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer 1

Comments:
3 – The repository is in the implementation phase

Reviewer 2

Comments:
3 – The repository is in the implementation phase

Response:

The Repository reference standard is ISO 19115 as specified in para 4.1 of NOAA Data Documentation Directive effective 2017-01-01. This directive is issued by the NOAA EDMC which reviews the standard at least annually.

References:

1. NOAA Data Documentation Directive:

https://nosc.noaa.gov/EDMC/documents/DataDocumentationPD-v2.0.0.signed_accessible.pdf

2. NOAA Environmental Data Management Committee (EDMC): https://nosc.noaa.gov/EDMC/

NOAA has a goal of having ISO 19115-1 be the administration's default standard by 2020. NOAA with NCEI in a leading role has developed a rough roadmap toward implementation that follows the Federal Geographic Data Committee (FGDC) ISO Geospatial Metadata Implementation Workflow Model.

Section 9 of the NOAA Data Documentation Directive does allow waivers to the ISO 19115 requirement for legacy or experimental datasets. However, for WDS-Paleo, ISO 19115 has been fully implemented. NCEI Collection Manager Tools have been used to test compliance and ISO 19115 metadata is available for all WDS-Paleo datasets.

References:

1. Federal Geographic Data Committee (FGDC) ISO Geospatial Metadata Implementation Workflow Model:

https://www.fgdc.gov/metadata/iso-implementation-model-workflow

2. NOAA Data Documentation Directive:

https://nosc.noaa.gov/EDMC/documents/DataDocumentationPD-v2.0.0.signed_accessible.pdf

3. ISO 19115 Metadata: https://www1.ncdc.noaa.gov/pub/data/metadata/published/paleo/iso/xml/

The NCEI Archive is part of its parent organization's (NESDIS) Ground System Enterprise (https://www.nesdis.noaa.gov/OSGS/about_osgs.html) under the system component Storage and Access. Part of the mission of the NESDIS Office of Satellite Ground Services (OSGS) is to develop the next generation of common enterprise ground architecture from which the NCEI Archive will directly benefit.

Beyond the intermediate plans of OSGS for infrastructure development, NCEI is a beneficiary of NOAA's Big Data Project (https://www.noaa.gov/big-data-project) created to explore the benefits of storing environmental data using cloud technologies.

NCEI maintains a documented baseline software inventory for its digital archive that is under configuration management. This software baseline covers operating systems, commercial-off-the-shelf (COTS) and internally-developed software. Since the repository is on a US Government network, system documentation is considered sensitive and, therefore, not releasable to the public.

There is no community software in use relevant to operating systems and core infrastructural software.

While WDS-Paleo isn't a real-time data stream, the NCEI repository does provide around-the-clock connectivity to data providers and users with a dedicated on-call staff to ensure this connectivity. NCEI ingests on average 2-3 TB of data per day into the archive with sufficient bandwidth. The Paleo-WDS contribution to this volume is too small to significantly affect this data stream.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

XVI. Security

R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The repository, as part of the NOAA IT infrastructure, complies with the Federal Information System Security Act of 2002 (FISMA) and operates in accordance with the National Institute of Standards and Technology (NIST) Special Publication (SP) 800-34, Rev. 1, Contingency Planning Guide for Information Systems. NIST SP 800-34 provides the repository clear guidance to help NCEI evaluate information systems and operations to determine contingency planning requirements and priorities.

Under the direction of NIST 800-34 (para 2.2.7), NCEI has documented a clear Information System Contingency Plan (ISCP) that includes an annually conducted Business Impact Analysis (BIA) which determines process and system criticality, identifies resource requirements, and identifies recovery priorities for the repository. The ISCP also delineates the roles and responsibilities and expected behavior of all individuals who access the system. The ISCP is considered a sensitive document and cannot be shared outside the US Government so therefore cannot be presented as public documentary evidence for this requirement.

The repository has a designated Information System Security Officer (ISSO) whose team performs daily security oversight and incident response. System IT security incidents are reported and handled in accordance with NOAA incident response procedures. An intrusion detection system monitors traffic with tools from the latest commercial software.

Under the requirements specified in Appendix F of NIST Special Publication 800-37, Guide for Applying the Risk Management Framework to Federal Information Systems, the repository system also maintains an Authorization to Operate (ATO) and authorization documentation and related artifacts are created and maintained in the Department of Commerce's Cyber Security Assessment and Management (CSAM) system to support the authorization to operate. Access to the ATO is restricted to official US government use only and cannot be made public.

WDS-Paleo data is backed up incrementally on a daily basis, with a full backup performed weekly. A complete set of backed up data is securely stored in an off-site storage facility. In the event of catastrophic loss, data can be recovered in accordance to the ISCP (as mentioned above in this section, this document is not publicly releasable as evidence due to US Government restrictions).

The NCEI repository does not outsource technical infrastructure and therefore is not required to assess risk in this area.

References:
1. Federal Information System Security Act of 2002 (FISMA):
https://en.wikipedia.org/wiki/Federal_Information_Security_Management_Act_of_2002
2. The National Institute of Standards and Technology (NIST) Special Publication (SP) 800-34, Rev. 1, Contingency Planning Guide for Information Systems: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-34r1.pdf
3. NIST Special Publication 800-37, Guide for Applying the Risk Management Framework to Federal Information Systems: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-37r1.pdf

APPLICANT FEEDBACK

Comments/feedback

These requirements are not seen as final, and we value your input to improve the core certification procedure. To this end, please leave any comments you wish to make on both the quality of the Catalogue and its relevance to your organization, as well as any other related thoughts.

Response:

No specific comments at this time. My organization has four WDS centers and WDS-Paleo is our first to attempt re-certification. We are interested in how requirements will change in 2020 since this will affect our subsequent certifications.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments: