

Data Submission Agreement

between the

Integrated Ocean Observing System (IOOS)
Great Lakes Observing System (GLOS)

and the

NOAA National Centers for Environmental Information (NCEI)

Revised: 2024-02-29

Table of Contents

Table of Contents	2
Introduction	3
Purpose and Scope	3
Document Change Management	3
Approval Date	4
Approval Date of the SA	4
Contacts	4
A. Provider Contacts	4
B. NCEI Contacts	4
Data Overview	6
File Transfer Interface	7
A. Transfer Interface for All Files	7
File Submissions for Data Ingest	8
A. [Dataset Name or File Type Name]	8
Submission Manifest Files	8
A. Submission Manifest Files	9
Archival Storage	9
A. [Dataset Name or File Type Name]	9
Metadata Management	10
A. [Dataset Name or File Type Name]	10
Data Access Services	10
Web Content	10
A. Web Content Information	10
References	10
A. References	10
General Terms and Additional Information	10
A. General Terms and Additional Information	10
Appendix A: Document Change Log	14

Introduction

This document represents the agreement that the Integrated Ocean Observing System's (IOOS) Great Lakes Observing System (GLOS) (the "**Provider**") and the NOAA National Centers for Environmental Information ("**NCEI**") have reached for submitting the Provider's data, the Integrated Ocean Observing System (IOOS) Great Lakes Observing System (GLOS), to NCEI for long-term data preservation and stewardship. This document is a joint effort between the Provider and NCEI to accurately document the agreement and the expectations between the two entities.

This archive system comprises the NCEI's Archive Management System. The Archive Management System (AMS) is a NOAA information technology archive system that supports long-term, secure preservation and standards-based archiving of environmental data collections and information.

NOAA created NCEI by merging NOAA's National Climatic Data Center (NCDC), National Geophysical Data Center (NGDC), and National Oceanographic Data Center (NODC), including the National Coastal Data Development Center (NCDDC), per the Consolidated and Further Continuing Appropriations Act, 2015, Public Law 113-235. NCEI launched publicly on April 22, 2015.

Purpose and Scope

This Submission Agreement (SA) establishes a plan by which to operate the data submission, public access and long-term archive of the data. It is neither a comprehensive data management plan nor a record of the decision to archive the data. The technical data submission model defined herein has been negotiated, reviewed and tested by both parties prior to its implementation.

This SA document is effective upon the completion of the readiness review approval involving both parties. In order to ensure the quality and integrity of the data record in the archive, the Provider and NCEI agree to maintain this Agreement with accurate and up-to-date information throughout the lifespan of the specified data submission. Either party may terminate this Agreement with a written notice to the other party.

Document Change Management

Baseline submission agreements are maintained by NCEI in a document management system. For changes to this Agreement, either from a coordinated review or an ad hoc request, the one initiating the wanted change is responsible for submitting a change request to the document manager, at abdolib.ncei@noaa.gov. All changes must be reviewed and approved by both NCEI and the Provider before the baseline document can be modified.

Approval Date

The date that this agreement becomes effective.

Approval Date of the SA

Approval Date	Archive Project ID	Event Description
2024-03-07	SNAPP-570	Approved by the NCEI Data Submission Readiness Review Board (DSRRB). Refer to the DSRR documents for details.

Contacts

Contact information is provided for the identified functional roles.

A. Provider Contacts

Provider Function	Provider Contact
Primary Data POC	Sneha Bhadhade Senior Technical Advisor GLOS +(91) 890-542-6393 sneha@glos.org
Data Transfer POC	Joeseeph Smith Cyberinfrastructure Engineer GLOS 734-972-8096 joe@glos.org

B. NCEI Contacts

NCEI Function	NCEI Contact
Lead Division POC Data Content Manager	Jonathan Jackson Research Associate NCEI / COGS / OSB 228-688-1261 jonathan.Jackson@noaa.gov
Submission Agreement Author Metadata Management	Rob Persons Archive Specialist NCEI / DSD / AB 1 828-271-4711

	robert.persons@noaa.gov
Data Officer (For IAF Automations)	Sheri Phillips Physical Scientist NCEI / DSD / AB +1-828-989-9202 sheri.phillips@noaa.gov
Ingest (Receipt of Transfer) Archival Storage	John Relph IT Specialist NCEI / DSD / SDB +1-301-713-4914 john.relph@noaa.gov
Archival Storage	Eric Ogata IT Specialist NCEI / DSD / SDB +1-301-713-4871 eric.ogata@noaa.gov
Access Operations	Danny Brinegar Data Access Operations NCEI / DSD / DAB 1-828-271-4773 ncei-dsd.daops@noaa.gov
Information Services	NCEI Information Services NCEI / COGS / CoSB 1-866-732-2382 ncei.info@noaa.gov
Submission Agreement Document Management	Jewel Ward Data Archive Specialist NCEI / DSD / AB 1-828-271-4344 abdolib.Ncei@noaa.gov

Data Overview

This Submission Agreement covers the Great Lakes Observation System (GLOS). GLOS is one of 11 Regional Associations established nationwide through the NOAA Integrated Ocean Observing System (IOOS). IOOS coordinates the multiagency, cooperative effort to routinely collect real time data and manage historical information based on a continuously operating

network of buoys, ships, satellites, underwater vehicles, and other platforms.

These are netCDF format, in-situ buoy data collected and managed by GLOS. These data are needed for many purposes which include rapid detection and prediction of changes in our nation's ocean, Great Lakes, and coastal waters. These data include both meteorological and oceanographic parameters. These buoys are deployed seasonally as dictated by weather conditions.

This dataset has been archived at NCEI since July, 2014 which include measurements of water temperature, wind direction and many other elements.

File Transfer Interface

Identification of the Provider's system and the file transfer protocol for the data submissions to NCEI.

A. Transfer Interface for All Files

Interconnection Type	NCEI Connection: External to NCEI NESDIS Connection Type Code: NCEI-to-NOAA
Provider Organization	GLOS
Provider Server Name	seagull-api.glos.org IP Address redacted
Ingest Organization	National Centers for Environmental Information, NOAA/NESDIS
Ingest Server Host Name	ministry.ams.ncei.noaa.gov
File Transfer Protocol	HTTPS Pull by NCEI from the Provider
Username	N/A
File Path	Data Provider file path: https://seagull-api.glos.org/api/v1/ncei_inventory https://seagull-api.glos.org/api/v1/ncei_file/
Data Availability Time Period	The amount of time the Provider is able to keep the data available after it is initially staged for NCEI. The Provider can keep the data available for no less than 30 days.
Additional Information	None

File Submissions for Data Ingest

This section describes the files submitted by the Provider for the NCEI data ingest system.

A. IOOS GLOS

Dataset ID	gov.noaa.nodc:IOOS-GLOS
File Name Pattern	<station-ID>_<data-type-name>_DDD-dd_YYYY-MM-DD_YYYY-MM-DD.nc DDD-dd = Depth in meters and decimal fraction (can also be DDD-d) e.g., 1.2 would be 001-2 1.25 would be 001-25 YYYY-MM-DD_YYYY-MM-DD are start and end observation dates.
File Name Examples	45013_relative-humidity_0-0_2023-10-27.nc
Has Submission Manifest File	YES See Submission Manifest Files for details.
File Format	netCDF-4
File Compression	None
File Size (Range)	30KB - 350KB (average 158K)
Static or Ongoing File Submission	ONGOING
File Submission Frequency	Monthly
File Count	5 - 100 files per station, > 20 stations
Annual Total	158K/file x 50 files/station x 25 stations x .7 years x 1MB/1000KB $\approx$ 138MB/yr
Notes	None

Submission Manifest Files

The Submission Manifest File is a companion file that represents each file submitted to the NCEI data ingest system in order to ensure the data file integrity.

A. Submission Manifest Files

Submission Manifest File Description	NCEI performs a file integrity check using a companion manifest file from the Provider. The checks are intended to ensure that each submitted file has been successfully transferred and ingested. The manifest file contains 2) the submitted file checksum value, and 1) the submitted file name, in the format created by the linux “md5sum” tool. The submitted file name field is optional. Content format of the values in a submission manifest file: <file_checksum_value> <file_name>
Checksum Algorithm	The checksum algorithm used for the manifest file. MD5
File Name Pattern	<file_name>.md5
File Name Field Definitions	<file_name> = Full name of the submitted file represented by the manifest file “.md5” = manifest filename extension indicating use of md5sum
File Name Examples	45013_relative-humidity_0-0_2023-10-27.nc.md5
File Format	md5sum
File Compression	None
File Size (Range)	<150 bytes
File Count (Rate)	One for each submitted file

Archival Storage

Archive attributes define the identification and organization of the data in the NCEI archival storage system. Administration rules apply to the management and operation of data in the archive. An individual AMS accession will be defined by the buoyID within the AMS. Archival information packages will consist of files with the same buoyID.

A. IOOS GLOS

Dataset ID	gov.noaa.nodc:IOOS-GLOS
Archive Storage System	NCEI AMS
Archive File Name Pattern	<station-ID>_<data-type-name>_DDD-d_YYYY-MM-DD_YYYY-MM-DD.nc

Archive File Path	N/A
Tar File Inventory	N/A
CLASS CS Collection Registration Information	N/A
Collection Metadata File ID	N/A
Data DOI	Not Planned
Notes	None

Metadata Management

A. IOOS GLOS

Collection Manager Record Group WAF	https://data.noaa.gov/waf/NOAA/NESDIS/ncei/accessions/iso/
-------------------------------------	---

Data Access Services

A. IOOS GLOS - netCDF-4

Dataset ID	gov.noaa.nodc:IOOS-GLOS
Data Access Services	• HTTPS • THREDDS • FTP
Data Access Locations	• https://www.ncei.noaa.gov/archive/accession/XXXXXXX where XXXXXXX is the accession number
Data Access Configurations	• HTTPS ◦ https://www.ncei.noaa.gov/data/oceans/ioos/glos/ • THREDDS ◦ https://www.ncei.noaa.gov/thredds-ocean/ioos/glos/catalog.html • FTP ◦ ftp://ftp-oceans.ncei.noaa.gov/pub/data.nodc/ioos/glos/
Notes	None

Web Content

The Web Content section describes the published information NCEI supports for the data.

A. Web Content Information

Product Category	N/A
Web Page Location	N/A

References

The cited documents are applicable to or referenced from this Agreement.

A. References

#	Reference Document
1	NOAA Records Schedules, Chapter 1400 – National Environmental Satellite, Data and Information Services (NESDIS) Functional Files. Revised: October 2017.
2	NESDIS Policy and Procedures for Determining Minimum Documentation Requirements for System Interconnections, Version 2.1, NESDIS Chief Information Division, September 28, 2012.
3	NCEI Standard Operating Procedures (SOP) for archive automation of Great Lakes Observing System data. https://intra.ams.ncei.noaa.gov/twiki/bin/view/NODC/GLOSnnonFedBuoy

General Terms and Additional Information

A. General Terms and Additional Information

Data Ingest and Archival Storage	- NCEI's operational systems are designed to run continuously. However, the NCEI resources to monitor, identify and resolve operational issues are limited to Federal Government business hours (M-F, 8AM-5PM). - While NCEI makes every reasonable effort to ensure submitted data are ingested into its archive, NCEI makes no performance guarantee for data ingest prior to the data being archived. NCEI cannot assume liability for any data loss caused by errors, omissions, system downtime or other issues which may affect a data submission.
---	---

- NCEI cannot verify file integrity without a companion submission manifest file.
- The NCEI Archive contact may provide an inventory of the files ingested and archived to the Provider via email as requested by the Provider.
- NCEI and the Provider are expected to reconcile identified errors within reasonable limits.
- Any duplicate file replaces a previously received file with the same file name.
- The Provider is responsible for notifying NCEI of any wanted unscheduled replacements or corrected submissions of data in the archive.
- A submission of updated data requires adequate notification to NCEI prior to the data submission.
- A new corresponding submission manifest is required for files re-submitted by the Provider.
- Any data ingested to NCEI servers may be subject to vulnerability scans.
- The Provider is responsible for verifying the success of an SFTP/ FTPS/ FTP Push session and reporting connection problems to the NCEI Ingest POC, listed in the [Contacts](#) section. Including the associated dataset ID number listed in the [Archival Storage](#) section i.e., DSI number.
- IT Security representatives for NCEI and the Provider are responsible for exchanging the necessary information to establish the network connection for the data transfer.
- NCEI is responsible for removing files in the directory received from an SFTP/ FTPS/ FTP push. The data provider should not remove any files successfully transferred via SFTP / FTPS/ FTP, unless the files were sent erroneously.
- NCEI is responsible for verifying the success of an SFTP / FTPS/ FTP Pull session and reporting connection problems to the Provider Transfer Contact.
- NCEI recommends the Provider maintain the data availability for 14 days in order to minimize the chance of Ndata loss, with a minimum data availability of 7 days. There is a certain probability for data loss over the course of a year if the data are available for less than 7 days.
- In the case when the SA lead does not have / provide the IP address within the SA, IT Security representatives for NCEI and the Provider are responsible for exchanging the necessary information to establish the network

	connection for the data transfer. • When transferring any ad hoc or unscheduled files to NCEI, the Provider must notify NCEI Ingest POC located in the Contacts section. Including the associated dataset ID number listed in the Archival Storage section (e.g., a DSI or Accession number). • This submission is subject to the record schedule governing retention of all records of this type in the archive. • NCEI encourages the use of appropriate file format extensions in all data, documentation, and source code file names. • For aggregated file formats, NCEI recommends setting the owner and group of a tar file to 0 so that the originating system information is not available to external users. NCEI also recommends setting the tar format to POSIX. • Submitted aggregate files should have complete contents, containing all files as specified in this SA. • The aggregate files should not contain any symbolic links, zero-byte size files, and file names with special characters. • The aggregate files may contain tar files or subdirectories. • For CDR source code: The NCEI IT Security contact may perform an IT security scan on the received source code as required prior to the submission. Upon a successful scan, the NCEI processing contact is responsible for packaging the source code, including the results of the IT security scan, and submitting it to the ingest system for archiving. • Correct submission manifest files are required for a successful ingest of the associated data file. • Submission manifest files are transferred via the same protocol and method as the associated data file. • Submission manifest files are transferred immediately following a successful transfer of the associated data file. • A compliant submission manifest file generator may be available from NCEI upon request. • Submitted tar files should not contain any symbolic links, zero-byte size files, or file names with special characters. • The file counts in this table do not account for the companion submission manifest files.
--	---

Data Access and Web Services	• For data archived on HDSS tapes: No more than 1,000 files or 500GB per order from the NCEI tape archive. • For data archived on CLASS tape: No more than 1,000 files or 366 days per order from the CLASS tape storage. Order limitations may also vary depending on file sizes and CLASS user account settings. • Access latency: Data archived within the NCEI AMS is normally checked and available online within 30 days of data receipt. • NCEI can only support public access to data that have been ingested and archived. • For web access metrics: please contact the Data Operations Branch at ncei-dsd.daops@noaa.gov • All data submitted to the archive may be publicly shared unless otherwise stated in this document or another agreement. • NCEI may adjust titles, abstracts, and other identification information about the data to promote data access and discoverability on the web. • NCEI can only provide public access to data after it has been ingested and archived. • NCEI may give the Provider statistics on the data access upon request. • NCEI may issue a digital object identifier (DOI) for citing the data once the data have been archived. • Additional terms and conditions may be applicable and specified in related documentation such as user agreements for the data. • Product announcements and notifications may be disseminated to users via communication channels including the NCEI website and social media. • The Provider is responsible for providing metadata information about the data to NCEI. NCEI is responsible for ensuring that the provided metadata comply with NOAA requirements following ISO standards. This metadata allows users to discover, access and understand the data available from NCEI. The ISO standard metadata is published by NCEI for use by public data catalogs.
-------------------------------------	--

Appendix A: Document Change Log

Below is a change record for this Submission Agreement.

Version/Date	Change Authority	Change Description	Section(s) Affected
V0.1A / 22 FEB 2024	NCEI	Initial Draft for preliminary review	All
V0.2A / 29 FEB 2024	NCEI	Draft for group review:	All
V1.0 / 07 Mar 2024	NCEI	Initial Baseline: AB-SA-03735 R0	All