



ROSA DATA GOVERNANCE GUIDANCE

DATA SHARING RECOMMENDATIONS

ROSA
Data Governance Program

ROSA Data Governance Guidance

Last Update

v1.0 | August 2026

Suggested Citation

Responsible Offshore Science Alliance. (2026). ROSA Data Governance Guidance. Zenodo.
<https://doi.org/10.5281/zenodo.21224314>



How to Use ROSA's Data Governance Guidance

A practical reference for planning, managing, and publishing data from offshore wind fisheries research and monitoring efforts.

WHO THIS IS FOR

Researchers and data collectors developing a Data Management and Sharing Plan (DMSP)

Funders of offshore wind fisheries research and monitoring

Offshore wind developers creating or implementing Fisheries or Benthic Monitoring Plans

USE IT AT THESE MOMENTS

01

Scoping a new project

Start your Data Management and Sharing Plan using the joint RWSC & ROSA template on DMPTool.org.

Section 3.1

02

Tracking publishing deadlines

Reference suggested timelines for DMSPs, updates, and data and metadata release.

Section 2.2

03

Deciding where data should live

START HERE

Domain-Specific Repository

Always check for a domain-specific repository first. ROSA and RWSC maintain a curated list covering fish survey data, acoustic telemetry, eDNA, benthic biology, plankton, and more.

Section 3.2.1

IF NO DOMAIN-SPECIFIC OPTION APPLIES

Generalist Repository

Zenodo, Dryad, or Figshare are acceptable when no domain-specific repository fits. Going this route requires additional attention: you must supply rich metadata following ISO 19115 or EML standards, and data must be in non-proprietary formats.

Section 3.2.2

04

Handling sensitive or restricted data

Endangered species locations, proprietary fishing data, or Indigenous data — see access restriction guidance and CARE principles.

Sections 2.4 & 1.1

05

Updating previously published data

Follow version control practices so earlier versions remain accessible and citable.

Section 3.3

QUICK REFERENCE

- Data counts as "published" only when it is deposited in an approved repository and has a DOI or other persistent identifier.
- Every funded project should have a DMSP. It is the document that ties together format, timeline, repository choice, and sensitive-data decisions.

Table of Contents

Table of Contents.....	3
1. PURPOSE.....	4
1.1 Guiding Principles.....	4
1.2 Data Governance Glossary.....	5
2. DATA PUBLISHING TIMELINES.....	6
2.1 Data Policy.....	6
2.2 Suggested Timelines.....	7
2.3 Advance Metadata Publication.....	7
2.4 Sensitive Data and Access Restrictions.....	7
3. DATA PUBLISHING GUIDELINES.....	8
3.1 Data Management and Sharing Plans.....	8
3.2 Repositories & Metadata Overview.....	9
3.2.1 Recommended Domain-Specific Repositories.....	9
3.2.2 Generalist Repositories.....	11
3.2.3 Metadata Preparation.....	11
3.2.3.1 Minimum Metadata Requirements.....	11
3.2.4 Data Formats and Interoperability.....	14
3.3 Data Changes and Version Control.....	15

Acknowledgements

ROSA gratefully acknowledges the contributions of the many individuals and groups who informed the development of this Data Governance Guidance. We thank the Intertidal Agency and the members of the ROSA Data Governance Committee for their sustained oversight and review throughout this process. We also thank the Fishing Gear Working Group and the Image Data Working Group for their technical expertise in shaping the recommendations included in this guidance. We thank the Regional Wildlife Science Collaborative for continued partnership, and in particular for collaboration in developing an aligned Data Policy, Data Glossary, Data Management and Sharing Plan template and user guidance for funders. This document reflects the collective input of these groups. We thank the Northeast Regional Ocean Council and the Mid-Atlantic Regional Council on the Ocean for both their input and generous funding support.

1. PURPOSE

The Responsible Offshore Science Alliance (ROSA) is a nonprofit organization that provides for and advances regional scientific research, monitoring and understanding of fisheries and the interactions with blue economy activities in state and federal waters of the U.S. through collaboration and cooperation. ROSA's Data Governance Program supports that mission by building data stewardship capacity and collaboratively developing processes, policies, standards, and workflows that maximize data value and impact. This document is designed to provide guidance around data management for the fisheries research and monitoring community as they scope, plan, and execute research and publish, manage, and share data.

This guidance can be cited by reference in contracts and/or data. It may be reused and remixed in contracts and scopes of work. It is CC-BY licensed.

This guidance applies to data generated through fisheries research and monitoring efforts conducted in U.S. Atlantic Waters. This includes, but is not limited to: biological sampling data, fisheries image data, socioeconomic and human dimensions data, acoustic telemetry data, environmental DNA (eDNA), plankton and oceanographic data, and derived data products (e.g., models, maps, and analytical outputs). It covers data collected across the lifecycle of a research or monitoring effort, from raw field data through to processed, analyzed, and synthesized outputs intended for publication or long-term preservation.

In the absence of existing guidance for this region, ROSA developed this document to provide researchers and data collectors with a comprehensive resource for managing and sharing these data types.

1.1 Guiding Principles

ROSA supports and encourages support for these guiding principles for data governance:

Open Science: Open science principles promote transparency, reproducibility, and broad accessibility of research outputs. Under these principles, data should be made "as open as possible, as closed as necessary" to balance the benefits of open access with legitimate needs for privacy, security, intellectual property protection, and ethical considerations.

FAIR Principles (Findable, Accessible, Interoperable, Reusable): The FAIR principles support data discovery, access, integration, and reuse by both humans and computational systems. This includes:

- Findable: Data must be registered in searchable repositories; assigned globally unique and persistent identifiers (e.g., DOIs); and described with rich metadata.
- Accessible: Metadata must remain accessible and access protocols described even when data are restricted; it is preferable that access should be open, free, and universally implementable where possible.

- Interoperable: Data must use formal, broadly applicable formats and vocabularies that follow community standards to enable integration with other datasets and analytical workflows.
- Reusable: Data must be richly described with accurate metadata, clear usage licenses, and provenance information to facilitate replication and combination in different contexts.

Indigenous Data and CARE Principles: Tribal Nations and Indigenous communities hold sovereign authority over data pertaining to their lands, waters, resources, and cultural heritage, and are essential partners in discussions around offshore development. ROSA is committed to supporting Tribal data sovereignty and partners should engage with Tribal Nations early, directly, and on an ongoing basis regarding any research, monitoring, or data collection that may affect Tribal interests, resources, or Indigenous Knowledge. ROSA intends to further engage Tribal Nations about data management and data sovereignty as it relates to fishery research.

ROSA encourages partners to be familiar with the [CARE Principles](#) (Collective Benefit, Authority to Control, Responsibility, Ethics):

- Collective Benefit: Data ecosystems shall be designed to enable Indigenous Peoples to derive benefit from the data.
- Authority to Control: Indigenous Peoples' rights and interests in Indigenous data must be recognized, and their authority to control such data must be supported.
- Responsibility: Those working with Tribal or Indigenous data have a responsibility to share how data are used to support the data partners' Indigenous Peoples' self-determination and collective benefit.
- Ethics: Data use must align with Indigenous ethical frameworks, minimize harm, and promote justice.

The following practices can support implementing the CARE principles:

- Consult with affected Indigenous communities early in project planning.
- Obtain Free, Prior, and Informed Consent for data collection and use.
- Recognize Indigenous Peoples' authority over data that concern them.
- Support Indigenous participation in data governance and decision-making.
- Ensure data use supports Indigenous self-determination and collective benefit.
- Respect Indigenous data protocols and ethical frameworks.
- Provide attribution and benefit-sharing as appropriate.

1.2 Data Governance Glossary

Clarity and standardization of terms used in data governance is vital to mutual understanding. ROSA has developed a [Glossary for Data Sharing/Curation/Documentation](#) to serve as a useful reference for terms that may not be familiar to all partners.

Below are a few terms used in this document:

Data Repository: A persistent, findable, searchable entity that provides infrastructure for long-term storage and access to data.

Data Standard: A technical specification that details the structure, organization, documentation, and format of data. Using data standards allows for the consistent collection of data, and aids in data aggregation, sharing and reuse, and interoperability of that data across different systems, sources, and users.

Metadata: Data about the data. The information that describes the data, both at the file-level (column headings, etc.) and the project level (data creator, geographic location, year, etc.), enables someone to know enough about the data to be able to reuse it.

Data Products: Models, maps, predictions, analysis results, visualizations and other syntheses that are produced from raw and/or derived data.

2. DATA PUBLISHING TIMELINES

Explicit timelines for data sharing support effective project management, coordination across research teams, and compliance with funder requirements. Clear expectations around data milestones help project managers track progress and allocate budget and staff time for data collection and curation activities. They also facilitate coordination among funders and researchers working on related or adjacent projects, enabling earlier data discovery and reducing duplication of effort. Explicit timelines and requirements support the release of data that might otherwise remain private, advancing the open science and FAIR principles that guide this program.

2.1 Data Policy

The [ROSA Data Policy](#) was designed to help inform other policies, contracts, and agreements so that fisheries, and offshore energy development research can advance knowledge, be used in decision-making, and have wide utility to researchers, government agencies, and others. ROSA encourages adoption of this language to standardize data availability of fisheries and offshore energy development research.

The ROSA Data Policy includes timelines for sharing data and essential metadata within one year of the end of data collection. ROSA also requires the development of a Data Management and Sharing Plan (DMSP; Section 3.1).

Data shall be considered published or shared if it:

- Has been submitted to one of the repositories described in this document, and
- Has a Digital Object Identifier (DOI) or other Persistent Identifier (PID).

Some organizations may have open data policies that specify data publication timelines. Funders - including industry, government agencies, and regrating partners - may require timelines for data management milestones and data publishing that flow down into contracts. Before agreeing on timelines for data activities and deliverables, data collectors should collect and review any additional applicable requirements from funders and partner institutions to ensure compliance with all deadlines.

2.2 Suggested Timelines

Deliverable	Deadline
Data Management and Sharing Plan (DMSP)	Draft submitted at proposal stage; Final version 6 months after signed project agreement
DMSP Updates	Annually, or with major project changes; Project End
Metadata and Data	Within one year of the last date of data collection, or at least annually for ongoing or multi-year data collection programs.

2.3 Advance Metadata Publication

Publishing comprehensive metadata records in advance of complete datasets is encouraged. Early metadata publication improves data discoverability, supports coordination among researchers working on related projects, and allows others to anticipate and plan for upcoming data availability while quality assurance and processing are ongoing.

Data collectors should identify their target repository early in a project and publish metadata records as soon as key descriptive information is available.

In cases where sensitive data cannot be made public on the timelines recommended above, metadata should be published to establish a record that the data exist and support their discoverability by others.

2.4 Sensitive Data and Access Restrictions

Some data collected may require access restrictions due to issues such as:

- Protection of endangered species locations
- National security concerns
- Proprietary business information (e.g., revenue data, fishing locations)

- Indigenous data sovereignty
- Personal privacy

When access must be restricted:

- Metadata should remain openly accessible describing what data exist
- The nature and duration of restrictions should be clearly documented in the DMSP and clearly stated in the public access point (e.g. metadata on a repository). and documented in the DMSP and repository
- Consideration should be given to providing aggregated or anonymized versions
- Restrictions should be reviewed periodically and lifted when no longer necessary or after a specified time period (e.g., 1-2 years after collection)
- Procedures for requesting access should be documented

3. DATA PUBLISHING GUIDELINES

3.1 Data Management and Sharing Plans

Data collectors should develop, maintain, and implement a Data Management and Sharing Plan (DMSP) to define, guide and illuminate data management strategies for project managers. ROSA and the Regional Wildlife Science Collaborative (RWSC) have developed a joint DMSP template to support coordinated, high-quality data stewardship across U.S. Atlantic offshore research projects.

The [U.S. Atlantic Offshore Research \(RWSC & ROSA\)](#) DMSP template is designed for any organization to organize and communicate how offshore data will be generated and made available from studies of wildlife, fisheries, and ocean ecosystems in U.S. Atlantic waters. The template integrates guidance described in this document to promote consistency across projects.

Researchers can create DMSPs using the template via the [DMPTool.org](#) online click-through wizard. Funders can use the recommendations and instructions provided below in their future solicitations and research agreements.

- [Access the template via DMPTool.org](#)
- [Download RWSC's & ROSA's recommendations and instructions for funders](#)
- [More about DMPTool](#)

If not using the U.S. Atlantic Offshore Research (RWSC & ROSA) template, DMSPs should include:

- Expected data types, formats, and volumes
- Metadata standards and documentation practices to be used
- Quality assurance and quality control procedures

- Data storage, backup, and security during the project
- Timeline for data sharing and publication
- Repositories selected for long-term preservation
- Roles and responsibilities for data management
- If possible, detailed budget allocation for data management activities
- Approach to managing sensitive data or data with restricted access
- Evidence that plans align with the FAIR and CARE principles

3.2 Repositories & Metadata Overview

Project data and metadata, data products, and other research outputs should be published in public repositories. The repository should provide for data publication by data holders, as well as access for using/reusing data.

Data holders should consider domain-specific repositories first and only choose a generalist repository if a domain-specific repository is not applicable or available. Minimum metadata and accessible format recommendations are provided for generalist repository submission.

3.2.1 Recommended Domain-Specific Repositories

ROSA and the RWSC have developed a list of domain-specific repositories recommended for many ocean data types. Table 1 lists data types most relevant to the ROSA community.

Repositories for other data types are available at https://rwscolab.github.io/repositories-data-catalog/repositories_v2.html.

Table 1: Domain-specific data repositories recommended by ROSA and RWSC relevant to the ROSA community (updated August 2026).

Data Type	Recommended Repository
Fishing Gear & Fish Biological Sampling Data Fish survey data, including data collected with trawls, nets, traps, or, dredges e.g. species occurrences, abundance counts, age samples	Ocean Biodiversity Information System (OBIS)
Socioeconomic & Human Dimensions Data Human survey and interview data	Zenodo

Data Type	Recommended Repository
Image Data (Video & Photograph) Seafloor plan-view photos, sediment profile images, underwater video, baited remote underwater video (BRUV), HabCams, drop cameras and other methods	NROC-MARCO-RWSC Seafloor Data Repository; fisheries video and photograph data under development through ROSA Image Data Working Group
Code, Software, and Related Tools Data Products Models, scripts, and other material	Zenodo (shared via a version-controlled platform, e.g., Github)
Acoustic Telemetry Data Fish tagging data	Ocean Tracking Network (OTN) and compatible nodes: ACT-MATOS & FACT
Environmental DNA (eDNA) Data Raw sequences, derived observations, reference libraries for genetic typing of fish	National Center for Biotechnology Information (NCBI); OBIS; Global Biodiversity Information Facility (GBIF); Catalogue of Life
Biological Oceanography Plankton and chlorophyll data collected with nets, tows, and active acoustics	Environmental Data Initiative (EDI); Biological and Chemical Oceanography Data Management Office (BCO-DMO) (with prior permission)
Physical Oceanography; Chemical Oceanography Oceanographic data and oceanographic model simulation outputs	Environmental Data Initiative (EDI); Biological and Chemical Oceanography Data Management Office (BCO-DMO) (with prior permission)

Data Type	Recommended Repository
Benthic Biologic Data Data and observations of submerged vegetation, infauna or epifauna	PANGAEA ; OBIS ; GBIF ; EDI

3.2.2 Generalist Repositories

When the domain-specific repositories above are not applicable or available, data and other research outputs should be deposited in trusted generalist repositories that support long-term preservation and provide persistent identifiers. In these cases, it is critical to deposit the data with rich metadata, using standards such as [ISO 19115](#) or Ecological Metadata Language (EML).

Examples of such generalist repositories include:

- [Zenodo](#)
- [Dryad](#)
- [Figshare](#)

Key services that qualify these three repositories as acceptable include:

- Persistent identifiers (such as [Digital Object Identifiers: DOIs](#))
- Version control capabilities
- Open access to metadata
- Long-term preservation commitments
- Clear licensing frameworks

3.2.3 Metadata Preparation

3.2.3.1 Minimum Metadata Requirements

All shared research outputs should be described by comprehensive metadata following relevant community norms. Datasets published in public repositories should follow the metadata requirements of those repositories, and will meet or exceed minimum metadata elements. If publishing in a generalist repository, or otherwise communicating about expected metadata, data holders should follow the following minimum metadata guidance from [RWSC](#) (Table 2).

Metadata should be published and shared in a machine actionable format (e.g. JSON, XML) and preferably shared via international metadata standards like Dublin Core, ISO 19115, and Ecological Metadata Language. A human readable document (e.g. pdf) may be shared alongside the machine actionable version.

Table 2: Minimum Metadata Requirements

	Description	Equivalent fields in common metadata standards		
Essential Metadata	Common-language description of the essential metadata terms.	EML	ISO19115-2	DCAT-US
Title	Should be informative and include the geographic area, the study subject, and the date (e.g. Benthic data from the Southern Gulf of Maine from 2020-2023)	Title	Title	Title
Author(s)	The person(s) responsible for the intellectual input(s) to this research output	Creator	Author	Creator
Contributor(s)	The person(s) or organization(s) responsible for parts of the data package, such as metadata provider, data steward, custodian, owner, principal investigator	Metadata Provider; Associated Party; Protocol Creator	Resource Provider; Custodian; Processor	Contributor; Rights Holder
Contact	The person responsible for curating the research output, and answering questions and providing further information about it	Contact	Point of Contact	Contact Point
Persistent Identifier (PID)	A unique persistent identifier such as a Digital Object Identifier (DOI)	Identifier	Identifier	Identifier

	Description	Equivalent fields in common metadata standards		
Essential Metadata	Common-language description of the essential metadata terms.	EML	ISO19115-2	DCAT-US
License	Terms that describe how the research output (including data) can be shared and reused. Good practice is to explicitly state the license selected for the data. The recommended licenses are CC0 (public domain) or CC-BY .	License and Usage Rights	License	License
Date(s)	Date(s) of data collection, or the date the research output was created	Date Range	Date	Temporal
Time(s)	Time(s) of data collection, or the time the research output was created	Temporal Coverage	DateTime	Temporal
Geographic location	Description and geographic coordinates or bounding coordinates of the data collection or research output	Geographic Description; Bounding Coordinates	Geographic Bounding Box	Spatial
Publisher	The name of the repository or journal or other organization that published the research output	Publisher	Publisher	Publisher

	Description	Equivalent fields in common metadata standards		
Essential Metadata	Common-language description of the essential metadata terms.	EML	ISO19115-2	DCAT-US
Publication Date	Date the research output was published	PubDate	Publication Date	Issued
Funder	Name and persistent ID (e.g. ROR , if available) of the organization(s) funding this research	Funding	Funder	Qualified Attribution

3.2.4 Data Formats and Interoperability

Data should be shared and published only in non-proprietary, machine-readable formats that support long-term preservation and reuse. Table 3 below provides examples.

Table 3: Examples of preferred file types for specific data types

Data Type	Preferred Format	Alternative
Tabular data	CSV, TSV	Structured text
Spatial data	GeoJSON, GeoPackage	Shapefile
Time series	NetCDF (CF conventions)	CSV
Images	TIFF, PNG	JPEG
Documents	PDF/A	PDF
Genetic sequences	FASTA, FASTQ	—

If you must use proprietary formats (Excel, ArcGIS files, etc.):

- Always include an open-format equivalent
- Document which software/version was used
- Explain any features not captured in the open format

Proprietary formats may be included as supplementary files, but must be accompanied by open format equivalents whenever possible.

3.3 Data Changes and Version Control

Once published, data and metadata may need to be corrected or updated to address changes such as:

- Additional quality control measures
- Expanded documentation
- Corrected errors or inconsistencies
- Supplementary contextual information
- Updated processing methodologies

These changes must be documented in the metadata. Data repositories support versioning to track these improvements while maintaining access to earlier versions and ensuring proper citation practices.

Please work with your selected repository to make sure that when published data are changed:

- New versions should be assigned updated version numbers or new DOIs.
- Changes between versions should be documented.
- Previous versions should remain accessible.
- Citation guidance should be updated to reflect the current version.