NFSI Data Policy

This document establishes the NFSI Data Policy governing the management, access, sharing, and public release of data collected using NFSI instrumentation. It provides the default framework applicable to all NFSI-supported projects unless superseded by funding agency requirements or institutional security obligations. Project-specific Data Sharing Agreements shall be developed between NFSI and the participating Principal Investigators (PIs) before project commencement. These agreements shall implement this policy and document any project-specific provisions that are consistent with it.

All data collected using NFSI instrumentation will be archived at the NFSI Marine Data Repository (MDR). These data will be made available to Project PIs within 3 months post-cruise or post-security review if applicable. The Data Policy assumes that data in appropriate formats will be delivered electronically; if Project PIs desire hard copies of the data, this should be articulated at the solicitation of instrumentation, so these extra costs are budgeted.

1. Terms for Teams of Principal Investigators (PIs)

1.1 Funding agency (e.g., CFI and end-user grant policies) requirements and institutional security obligations on data access shall take precedence over the NFSI data policy.

1.2 Prior to project commencement, NFSI and the participating Principal Investigators shall execute a Project Data Sharing Agreement. The agreement shall implement this policy and document project-specific provisions, including any approved deviations from the default terms established herein.

1.3 The agreement will include provisions for:

- 1.3.1 Exclusive access rights and embargo period (Section 2).
- 1.3.2 Restricted post-embargo data access (Section 3).
- 1.3.3 Unrestricted post-embargo data access (Section 4).
- 1.3.4 Respecting the data policies of the respective funding agencies and journals where they publish using the data.
- 1.3.5 Procedures for resolving disputes or ambiguities regarding data sharing. Unless otherwise specified in the Project Data Sharing Agreement, disputes shall first be addressed through consultation between the Project PIs and the NFSI-MDR Team. If unresolved, the matter shall be referred to the *NFSI Executive Committee*[1] for review and recommendation.

1.4 Any changes to the agreed-upon data policy after the project starts will require unanimous written consent from all PIs and approval by the *NFSI Executive Committee*. Grounds for declining changes to the data-sharing terms include but are not limited to:

- 1.4.1 Excessive embargo extension request.
- 1.4.2 Excessive request for extension of restricted post-embargo data access.
- 1.4.3 Undocumented or unjustified changes to data-sharing terms.

2. Exclusive Access Rights and Embargo Period

The default data-access timelines described in Sections 2–4 are summarized in Figure 1.

2.1 Embargo period begins at data delivery. In the absence of specific funding agency requirements, the default embargo period is 3 years. If the data are not collected through a funding agency grant or the data are collected for industry, the embargo period is 10 years.

- 2.1.1 For multi-year experiments, each Leg will have its own embargo period.
- 2.1.2 During the embargo period, the team of PIs (and their mentees) has exclusive access rights to the dataset. Basic metadata (e.g., StationXML) can be released in advance with PI(s) approval.

2.2 Requests from PI(s)' collaborators (and their mentees) for data access during the embargo period will require an application to be reviewed by the Project PI(s). The application should include:

- 2.2.1 The requestee's name, institution, and statement of use.
- 2.2.2 A statement of adherence to NFSI data-sharing policies and guidelines.
- 2.2.3 Endorsement from the collaborating PI.
- 2.2.4 Approved applications will result in a legally binding contract issued by NFSI. The contract will specify the terms for data delivery, usage, storage, and sharing.

3. Restricted Post-Embargo Data Access

3.1 At the conclusion of the embargo period, and for datasets consisting of eight or more stations with at least 300 days of data each, the NFSI-MDR will enable data access to a minimum of 60-day data from one station via FDSNWS with no restrictions. These criteria apply to the individual legs of multi-year experiments.

Datasets that do not satisfy these eligibility criteria shall remain under the applicable embargo conditions until unrestricted public release under Section 4, unless otherwise provided in the applicable Project Data Sharing Agreement.

- 3.1.1 Stations exhibiting unrecoverable data, clock synchronization errors, or instrumental malfunctions that materially compromise data quality shall not be counted toward the minimum of eight qualifying stations. If eight or more qualifying stations remain after these exclusions, Sections 3.1.2–3.1.4 apply.
- 3.1.2 Project PIs will, within two months prior to the conclusion of the embargo period, inform the NFSI-MDR Team of the following:
  which station is selected for partial access, and
  which 60 days (consecutive or nonconsecutive) will be made open access for the above selected station.

3.1.3 Project PIs are expected to select a station and time period that reflects the average data quality and signal characteristics of the dataset.
3.1.4 NFSI-MDR will enable selected data access within 3 months from receiving PIs notification.
3.1.5 In the absence of PIs notification, NFSI-MDR will select the station and time period of dataset to release.

3.2 Following the embargo period, the data will be made available upon request to any researcher, institution or organization working within, or for, the country(ies) that provided direct or indirect funding for the project. The request must include:
3.2.1 A statement of adherence to NFSI data policies and guidelines.
3.2.2 A proposal outlining the intended use of the data.

3.3 Researchers and organizations from countries that did not provide direct or indirect funding may request access to the data by submitting a formal application to the PIs and the NFSI. The application must include:
3.3.1 A statement of adherence to NFSI data policies and guidelines.
3.3.2 A detailed proposal outlining the intended use of the data.
3.3.3 A data sharing and collaboration plan, including at least one collaborator from the team of PIs.

3.4 The PIs and *NFSI Data Review Committee*[2] will review applications. Approved applications will result in a legally binding contract issued by NFSI. The contract will specify the terms for data delivery, usage, storage, and sharing. Approval of requests under Sections 3.2 and 3.3 remains subject to applicable security, export-control, and institutional requirements.

3.5 Further data sharing beyond the approved parties will require a separate written agreement reviewed and approved by the PIs and the *NFSI Data Review Committee*.

4. Unrestricted Post-Embargo Data Access

4.1 Unless otherwise required by the applicable funding agency or Project Data Sharing Agreement, the entire dataset shall be released to unrestricted public access ten years after data delivery. For projects subject to a ten-year embargo (e.g., industry or non-grant-funded projects), unrestricted public access begins immediately upon expiry of the embargo, unless otherwise specified in the Project Data Sharing Agreement.

4.2 These data are made available publicly using the NFSI-MDR FDSNWS, where all access passwords are disabled.

5. General Data-Sharing Principles

5.1 All data users must comply with ethical guidelines, confidentiality agreements, and legal restrictions.

5.2 Proper acknowledgment of the data's origin and the funding source is mandatory in all instances of data usage. PIs will share with NFSI publications resulting from data usage.

5.3 Misuse of data or violation of collaboration agreements will result in penalties, including potential revocation of access and other legal actions.

6. Ownership and Intellectual Property

NFSI acts as the custodian of data archived within the NFSI Marine Data Repository. Ownership of the data, intellectual property rights, and rights to analyses, interpretations, publications, and derived research products remain with the originating investigators or their institutions, subject to applicable funding agreements and institutional policies. Nothing in this policy transfers ownership of data or intellectual property to NFSI.

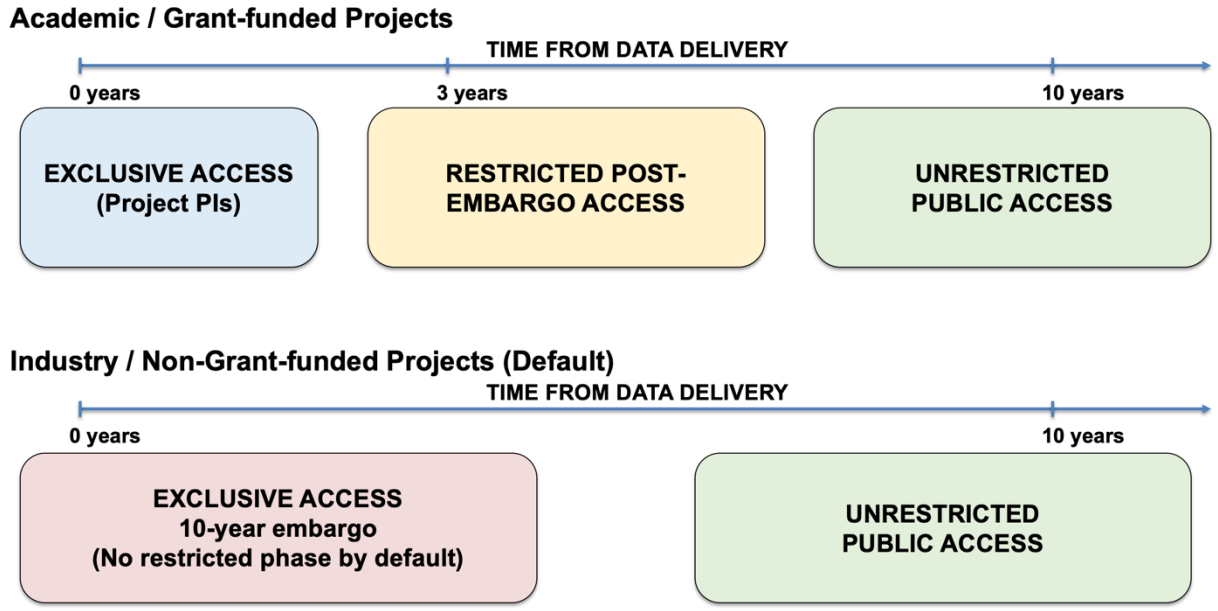


Figure 1. NFSI Data Access Timeline (Default)

1. **NFSI Executive Committee**: responsible for governance and policy decisions.
2. **NFSI Data Review Committee**: responsible for evaluating requests for restricted data access.

Appendices

A. Metadata, and data formats

A.1. All digital data will be provided in standard formats.

A.1.1. Passive-source data will be archived in miniSEED format.

A.1.2. Active-source data will be archived in SEG-Y format.

A.1.3. Metadata and instrumentation response will be provided in StationXML format.

A.1.4. Shot-data (timing, location) will be provided in Shot Tables in ASCII files (most commonly .txt or .sht format).

A.2. All metadata, logs, location surveys, and data quality summaries associated with the project will be generated and archived by the NFSI-MDR Team, with input from Project PIs.

A.3. NFSI will provide a documentation of experiment logistics, instrument metadata, including but not limited to date and times, location calculations, time correction and clock drift, instrument tilt, etc. Cruise reports will be compiled by the PIs.

A.4. Passive source metadata and data

A.4.1. The NFSI-MDR Team and the cruise Chief Scientist(s) will be jointly responsible for assembling all necessary metadata, including instrument deployment details, recovery locations and times, all auxiliary channel metadata.

A.4.2. The NFSI-MDR Team and Project-PIs will identify, prior to project start, who is responsible for registering the FDSN Network Code for the experiment. The identified party will register the FDSN Temporary Network code and report it to the Project PIs and the NFSI.

A.4.3. Project PIs will provide, before the deployment experiment, the final station names to the NFSI-MDR Team, in a clear and concise format, that enables the NFSI to match OBS Instrument IDs with station names.

A.4.4. All channels of data recorded by the Aquarius OBS instruments will be archived and made accessible to the Project-PIs. These include, but are not limited to, the time-series broad-band data, hydrophone or absolute pressure gauge, system status channels, environmental status channels, etc.

A.4.5. Data will be archived using a standard directory and file-naming structure agreed upon between the Project PIs and the NFSI-MDR Team. Default directory structure follows the SeisComP Data Structure (SDS).

A.4.5.1. Data Structure: YYYY/NET/STN/CHN.D/NET.STN.LOC.CHN.D.YYYY.JJJ

A.4.5.2. Auxiliary non-Data Structure YYYY/NET/STN/CHN/NET.STN.LOC.CHN.YYYY.JJJ

A.4.6. The standard dataset provided by the NFSI will have the default time-correction applied to the time series, based on the clock calibrations carried out at the deployment and recovery of the instrumentation.

A.4.6.1. For stations within the dataset for which one of the time calibrations is missing or failed, or for data where the time calibration is outside of

standard tolerances, the data for these stations will be provided as is, uncorrected, with this clearly identified in supporting documents.

A.4.6.2. The NFSI-MDR Team will maintain an archive of all uncorrected time-series data, available upon special request by Project PIs, if necessary.

A.5. Active source metadata and data

A.5.1. The OBS data are provided in SEG-Y format with each component (X, Y, Z and hydrophone) of each instrument as a separate file. File naming includes the OBS number and component, and line number. Metadata (shot location, time, source-receiver offset, etc) are usually also provided within the corresponding trace headers, with the SEGY file header having information about the project name, line number, etc.

B. Data Access

B.1.1. Data access during the embargo and post-embargo period (but prior to Open Access) will be administered by the NFSI-MDR Team through user authentication. This may include the use of username/password, authentication tokens, etc, to be determined by the NFSI-MDR Team.

B.1.2. Data will be made available to project PIs as a complete download once it is prepared following the recovery cruise. This will be made available within three months following completion of the recovery cruise or completion of any required security review, whichever occurs later, or a specific timeline to be negotiated between Project PIs and the NFSI-MDR Team.

B.1.2.1. The complete download will contain all required metadata information, quality control summaries, logs and experiment summaries, as well as the SDS archive of all channels of data, at their original sampling rates.

B.1.2.2. This will be provided via protected FTP, Torrent, direct download, or other method as arranged between the PIs and the NFSI-MDR Team.

B.1.3. Ongoing waveform access will be made available to the Project PIs via authenticated access to the NFSI-MDR FDSNWS.

B.1.3.1. The NFSI-MDR FDSNWS will provide the following services at a minimum: fdsnws-station (metadata queries), fdsnws-dataselect (data queries).

B.1.3.2. Data will remain available on the NFSI-MDR FDSNWS to Project PIs in perpetuity. If the authenticated FDSNWS service is scheduled for permanent discontinuation, NFSI shall provide Project PIs with not less than **12 months' written notice** and shall make available a complete export of all archived project data and associated metadata in the standard archive formats defined in this policy.

B.1.3.3. Once the project moves to Open Access after the post-embargo period concludes, the data will remain available on the NFSI-MDR FDSNWS with all authentication requirements removed.