

# Drift South Expedition: Data Management Strategy

#### Overview

Drift South Expedition (DSE) is committed to collecting high-integrity environmental data throughout our 2,350-mile Mississippi River transect (September–November 2025). Our data management strategy ensures that information gathered in the field is reliable, traceable, secure, and useful for both internal analysis and external collaboration.

This strategy outlines the protocols we will follow for data collection, verification, storage, metadata handling, sharing, and long-term access.

# 1. Guiding Principles

- **Scientific Rigor:** All data is collected using standardized protocols adapted from professional water quality monitoring methods.
- **Transparency:** Clear documentation accompanies every data point—ensuring usability by downstream partners or researchers.
- Redundancy & Reliability: Dual data collection and backup systems reduce the risk of data loss.
- Accessibility: Our dataset will be structured for ease of use by both academic and community partners.



• **Ethical Stewardship:** We respect the privacy of individuals and cultural sensitivities in river communities and ensure no sensitive personal data is collected.

#### 2. Field Data Collection

#### 2.1 Standard Parameters

- Secchi depth (turbidity)
- Temperature (air & water)
- pH
- Nitrate/Nitrite concentration
- GPS location
- Date/time stamp
- Weather conditions
- Observational notes (land use, algal blooms, visible pollutants)

#### 2.2 Optional Parameters (Partner-Dependent)

• Our experienced field technicians have the capability to record what you need. Please reach out (contact info at the bottom of document).



# 3. Data Recording Protocols

To ensure data integrity, each field measurement will be recorded via:

- **Primary:** Field Datasheets (Waterproof Rite-in-the-Rain notebooks or laminated data forms)
- **Secondary:** Digital Entry via tablets or phones (Google Sheets + offline backup tool)
- Tertiary: Manual photo documentation of field sheets and sampling setups

Each measurement includes:

- Unique site ID
- Crew member initials
- Equipment used (model or DIY type)
- Calibration notes (if applicable)
- Environmental metadata (qualitative observations)

# 4. Redundancy & Quality Control

To reduce error and protect against data loss:

For questions about this document or to contribute to the protocol, please contact: charlesrlampman@gmail.com



- **Dual Entry:** Data is entered manually and digitally at the point of collection. Each evening, field notes are reviewed and compared.
- Data Audit: Weekly audits by a designated data manager ensure consistency, flag anomalies, and track missing data..
- **Time-Stamped Archiving:** Each day's datasets are archived with version control in both cloud (Google Drive) and external hard drives.

## 5. Data Storage & Access

- **Short-Term Storage:** Google Workspace (with daily backups) + physical copies in crew document kits.
- Long-Term Archival: Upon expedition completion, datasets will be exported in CSV and JSON formats for open-access storage via:
  - GitHub repository (public, read-only)
  - Partner academic or nonprofit data portals
  - DSE website (downloadable data viewer) (under construction)
- Metadata & Codebooks: Every dataset will be accompanied by a detailed README including variable definitions, unit explanations, equipment types, and site maps.

## 6. Data Sharing & Attribution



- All non-sensitive datasets will be openly licensed under a Creative Commons Attribution 4.0 License (CC-BY 4.0).
- Any partner-provided equipment or data integration will include attribution clauses as agreed upon.
- **Early-Access Sharing:** Core partners will receive weekly data snapshots and a mid-expedition summary.
- Post-Expedition Report: Comprehensive datasets and a methods appendix will be included in the public DSE final report (Winter 2025–2026).

# 7. Future Use & Expansion

This data management framework is designed for iteration. Future Drift South Expedition projects (or other river-based expeditions) can adopt or scale this protocol. We also encourage community groups to replicate this system using our open templates and training materials.



# **Appendices**

- Appendix A: Field Data Sheet Template (print & digital)
- Appendix B: Metadata & Data Dictionary
- Appendix C: Partner Attribution Guidelines
- Appendix D: File Naming & Archiving Conventions



# **Appendices**

#### Appendix A: Field Data Sheet Template (print & digital) (EXAMPLE)

- Site ID: Unique location code
- Date / Time: Format YYYY-MM-DD / HH:MM
- GPS Coordinates: Decimal degrees (WGS 84)
- Water Temp (°C): Calibrated thermometer
- Air Temp (°C): Ambient air reading
- pH: Probe reading with calibration notes
- Nitrate/Nitrite (mg/L): Test strip
- Secchi Depth (cm): Turbidity measurement
- Weather: Brief summary of conditions
- Observations: Land use, algal blooms, pollution, etc.
- Crew Initials: Data collector identification
- Equipment & Calibration: Type/model and recent calibration info
   Type/model + recent calibration info |

#### Appendix B: Metadata & Data Dictionary

- Site ID: Unique code representing river mile or local reference.
- Date/Time: ISO 8601 format (e.g., 2025-09-21T14:30).
- GPS: Decimal degrees (WGS 84).
- Temperature (Water/Air): Degrees Celsius, calibrated thermometers.
- pH: Measured with handheld probe; includes calibration log.
- Nitrate/Nitrite: mg/L, test strips or lab results noted.
- Secchi Depth: cm, average of two consistent readings.
- Weather Conditions: Cloud cover, wind, precipitation.
- Observations: Freeform notes on nearby land use, visible algae, pollution, human activity.



#### **Appendix C: Partner Attribution Guidelines**

- Include full organization name and logo (if provided) in public reports.
- Cite specific equipment, data contributions, or analysis support.
- Shared datasets: list partner name in metadata README.
- Academic contributions may be acknowledged via co-authorship or citation, per mutual agreement.

#### **Appendix D: File Naming & Archiving Conventions**

- File Naming Format:
  - DSE\_YYYYMMDD\_SITEID\_PARAMETER.csv
  - Example: DSE\_20251005\_0855\_pH.csv
- Folder Structure:
  - o /raw\_data/YYYYMMDD/
  - o /processed\_data/
  - o /photos/YYYYMMDD/
- Version Control: Include date and change notes in README file for each dataset.