

Observation Metadata Workbook

Give a future reader enough context to understand what a record can, and cannot, support.

Author: Theo Nguyen, digital product lead · Scientific editor: Amina Okafor · 17 October 2026 · Edition 1



Generated illustration. No instrument deployment, observation record, or expedition is claimed.

1. Start with the record's purpose

A number without its method is a difficult inheritance. Metadata are the account of who made a record, what it means, where and when it applies, how it was produced, and what restrictions accompany it. They should travel with the data, not remain in the memory of the person who exported a spreadsheet.

[NOAA NCEI's introduction to metadata](#) includes content, format, accuracy, provenance, resolution, responsible parties, citation, and licensing. This workbook translates that framing into a small working package. It is not an archive submission standard; use the receiving repository's requirements for a real deposit.

Before filling a table, write the decision or question the observations were intended to inform. Specify the target phenomenon, spatial coverage, and time window. Then write a use the record cannot support. A single near-surface temperature series, for example, does not describe the heat content of an entire water column.

Workbook rule. All sample identifiers, timestamps, and values in the exercises are fictional. Use them to practise documentation. Do not copy them into an environmental dataset or present them as Still North observations.

2. Make a collection passport

Separate collection-level context from row-level fields. The passport describes the project, deployment, or observation series as a whole. Individual records then point back to that context through stable identifiers. Repeating a paragraph in every row makes corrections harder, while hiding everything in a filename makes the record fragile.

Collection passport: write a specific answer, or mark the gap

Entry	Prompt
Identity and purpose	Title, stable collection ID, version, intended question, and unsuitable uses.
People and responsibility	Collector, processor, metadata editor, steward, and an appropriate contact route.
Place and time	Coverage, coordinate reference, time reference, start and end, and location precision.
Method	Instrument or observation protocol, settings, calibration evidence, and changes.
Access and continuity	Authority, restrictions, citation, review date, storage owner, and retention decision.

Location needs more than two numbers. Record coordinate order, units, reference system, and whether the position belongs to a station, instrument, sample, or moving platform. Keep precision honest. Extra decimal places do not turn an approximate location into an accurate one. Public location detail may also need to differ from the controlled record.

Time needs the same care. Distinguish the instrument clock, the observation time, the export time, and the processing date. Use an unambiguous representation such as 2026-06-12T14:00:00Z for a UTC timestamp when that reference is actually known. If a handwritten note says only “afternoon”, preserve that limitation rather than manufacturing a precise time.

For a fictional logger known, in the exercise, to run two minutes fast, keep its original timestamp and document the basis for a corrected timestamp. The correction is a transformation with an author and date, not an invisible replacement of history. If the offset is unknown, record “not established”.

Purpose / unsuitable use: _____ / _____.

3. Write the dictionary before adding rows

A data dictionary explains each field independently of the software used to open it. Use one variable per column and one defined observation unit per row. Describe whether a value is instantaneous, a mean over an interval, an accumulated total, or a category. Put the aggregation window in the dictionary or a dedicated field.

Fictional teaching dictionary for a three-record exercise

Field	Meaning and rule
record_id	Unique text identifier; never reuse it for a different observation.
time_utc	Fictional instant, UTC, written with the Z suffix. Not an interval average.
water_temp_c	Illustrative temperature in degrees Celsius. Empty means no value; zero is a value.
quality	One of unchecked, accepted, suspect, or missing. These are local teaching labels.
note	Reason for a flag or limitation; no personal or restricted information.

```
record_id,time_utc,water_temp_c,quality,note
EX-001,2026-06-12T14:00:00Z,2.4,unchecked,Teaching value only
EX-002,2026-06-12T15:00:00Z,,missing,No teaching value supplied
EX-003,2026-06-12T16:00:00Z,27.0,suspect,Requires investigation
```

Do not turn EX-002 into zero to make a graph continuous. Do not delete EX-003 merely because it looks surprising. A review might find a real event, a sensor problem, a unit error, or a transcription mistake. Until the reason is established, the flag preserves the question.

In your own dictionary, add data type, allowed categories, missing-value convention, measurement depth or height, and links to applicable method documents. If a sentinel such as -9999 is unavoidable in a supplied format, document it explicitly and prevent it from entering summary statistics as a measurement.

Reopen an exported file with a simple viewer when the real workflow allows. Spreadsheet applications can reinterpret dates, remove leading zeroes from identifiers, or alter text. A readable export and its dictionary should agree on encoding, separators, and units.

4. Record effort and quality decisions

Missing observations, unexamined observations, and observations with no detection are not interchangeable. Document when an instrument was functioning, when collection was scheduled, which intervals were recovered, and which intervals were actually reviewed. Otherwise a gap can quietly become evidence that nothing happened.

Consider a fictional camera scheduled to capture ten images but returning six. If two returned images were unreadable and four were checked, the reviewed effort is four usable images. Reporting “no animals in ten images” would conflate the plan, recovery, and review. The same structure applies to water samples, listening windows, and repeated vegetation photographs.

Effort ledger template

Interval / batch	Planned	Recovered	Usable and reviewed	Gap reason
Write ID	State unit	State unit	State criteria	Known / unknown

Keep quality flags separate from the measured value. State who assigned each flag, when, using which procedure, and whether the procedure was automated or manual. “Accepted” means it passed the stated checks, not that the record is perfect. Preserve conflicting judgements until the agreed reviewer resolves them.

Build a transformation ledger alongside the data. For each derived release, identify its source files, the operation performed, software or calculation version, responsible person, and known effect on interpretation. Examples include a unit conversion, a timestamp correction, duplicate removal, and a revised exclusion rule.

Retain the original input unchanged in controlled storage. A corrected table is a new version, with a note explaining what changed and why. Do not issue a fresh file with the old version label and expect readers to discover the difference themselves.

Quality question to resolve: What evidence would change this flag, and who has authority to make that change?

5. Package understanding, not just files

For a small permitted release, gather a concise README, the data dictionary, the released data, and the method and change records. Name the release date and version, link the files by stable identifiers, and provide a citation that resolves to the intended release. Do not invent a DOI; a repository assigns one through its own process.

Readable does not necessarily mean public. Metadata can reveal sensitive locations, activities, names, or relationships even when measurements are withheld. Agree what may be disclosed, to whom, under whose authority, and for which purposes before publishing a catalogue entry or uploading a file.

The [National Inuit Strategy on Research](#) identifies Inuit access, ownership, and control over data and information as a priority in Inuit Nunangat research. An open-data preference does not override that authority. This workbook grants no permission and cannot substitute for local agreements, ethical review, or repository rules.

The handover conversation

1. Ask a permitted colleague to explain one row using only the package.
2. Ask them to distinguish a missing value from a measured zero.
3. Ask which intervals were reviewed and which were not.
4. Ask where a correction would be recorded and who would approve release.

Keep their unanswered questions as revision notes. Successful documentation lets another person use the record within its limits, and lets them decline an unsuitable use without having to reconstruct the project.

Demonstration disclosure. Still North and its staff are fictional. Generated images and all sample records in this workbook are illustrative. No dataset, deployment, calibration, certification, or permission is asserted.