Fluorescence Spectroscopy Data Sharing Guide

This document defines recommended conventions for organizing and sharing fluorescence spectroscopy datasets, with a focus on measurements commonly used in the characterization of SWCNT-based optical nanosensors. For storage, there seems to be no field standard – consider Zenodo, Figshare, Dryad (if tied to a publication) or the Chemotion Repository based on your needs.

1. Scope

Fluorescence spectroscopy experiments often generate large collections of raw measurement files with heterogeneous structures. Different measurement modes may produce substantially different data formats, and instrument-generated files frequently lack sufficient context for interpretation.

This guide defines a set of recommended conventions for sharing fluorescence spectroscopy data in a structured and reusable manner. The document provides guidelines for:

- organizing datasets into experiment-level directories
- defining file naming conventions
- standardizing sample identifiers
- documenting experimental metadata
- reporting acquisition parameters
- describing the internal structure of raw data files

The conventions described here are intended to ensure that fluorescence spectroscopy datasets can be correctly interpreted and reused by researchers who were not involved in the original experiment.

Although the guide was developed with SWCNT-based nanosensor experiments in mind, the proposed structure can be applied broadly to fluorescence spectroscopy datasets.

2. Dataset Organization

Fluorescence spectroscopy datasets should be organized into **experiment-level directories**, where each directory corresponds to a single independent measurement.

An experiment is defined as a single acquisition performed using one of the following measurement modes:

- **EXEM** – excitation–emission fluorescence scan
- **TIMETRACE** – time-resolved fluorescence measurement
- **ABS** – absorbance spectroscopy measurement
- **SINGLE** – single-excitation fluorescence emission spectrum

Each experiment directory must contain the following components:

- **raw/** - Raw measurement files exported directly from the instrument. These files must be preserved exactly as produced by the instrument.
- **instrument_metadata/** - Instrument-generated metadata files describing acquisition parameters such as exposure time, wavelength ranges, detector settings, and scan timing information.
- **README.md** - Experiment-level documentation describing the purpose of the experiment, sample composition, file naming conventions, and the internal structure of the raw data files.

A recommended experiment directory structure is shown below:

```
experimentID_measurementType/
    raw/
    instrument_metadata/
    README.md
```

If the instrument does not generate metadata files automatically, the required acquisition parameters must be documented manually in the experiment README.

3. File Naming Conventions

Raw data files must follow consistent naming conventions that allow each measurement file to be uniquely identified. File naming rules may vary depending on the spectroscopy measurement type.

The following conventions are recommended:

3.1. EXEM Measurements

Excitation–emission fluorescence scans are typically performed using multi-well plates. Each well produces a separate measurement file.

File naming format: WELL_SAMPLE_raw.csv
Example: A02_sensor1_raw.csv
Where:

- **WELL** - well position in the plate (e.g., A02, B03)
- **SAMPLE** - sample identifier corresponding to the well
- **raw** - indicates that the file contains unprocessed instrument output

3.2. TIMETRACE Measurements

Time-trace fluorescence experiments produce a sequence of measurements recorded at regular time intervals.

File naming format: EXPERIMENTID_TIME_raw.csv
Example: 245650911_120_raw.csv
Where:

- **EXPERIMENTID** - unique identifier of the experiment
- **TIME** - time in seconds from the start of the measurement
- **raw** - indicates unprocessed data

3.3. ABS Measurements

Absorbance measurements are often recorded sequentially for different samples.

File naming format: EXPERIMENTID_SAMPLE_TIMESTAMP.txt
Example: 245650912_sensor1_115000.txt
Where:

- **EXPERIMENTID** - experiment identifier
- **SAMPLE** - sample identifier
- **TIMESTAMP** - measurement time (HHMMSS)

3.4. SINGLE Measurements

Single excitation fluorescence measurements produce a single emission spectrum.

File naming format: EXPERIMENTID_raw.csv
Example: 245650913_raw.csv
Where:

- **EXPERIMENTID** - experiment identifier
- **raw** – indicates the file contains the original measurement data

4. Sample Identifier Convention

4.1. Main Samples

Sample identifiers should follow a consistent naming convention that encodes the key experimental variables.

The recommended format is: sensorID-analyteConcentration-replicateNumber
Example: sensor1-10-2
Meaning:

- **sensor1** – identifier of the SWCNT sensor
- **10** – analyte concentration
- **2** – replicate number

The identity of the analyte and the units of concentration must be documented in the experiment metadata.

Replicates are used to represent repeated measurements of the same experimental condition.

4.2. Control Samples

Control samples may use simplified identifiers.

Examples:
PBS-1
PBS-2
PBS-3
Where:

- PBS indicates a buffer-only control measurement
- the number indicates the replicate index

5. Required Metadata

Fluorescence spectroscopy datasets must include sufficient metadata to allow other researchers to correctly interpret the measurements.

Metadata may be provided through:

- instrument-generated metadata files (when available)
- experiment-level README documentation

When instrument metadata files are available, they should be preserved and included in the dataset. If the instrument does not export metadata automatically, the required parameters must be documented manually.

The metadata requirements can be grouped into three categories:

- experimental metadata
- acquisition metadata
- data structure metadata

5.1. Experimental metadata

Experimental metadata describes the biological or chemical context of the experiment and the identity of the samples.

Metadata Field	Description
experiment identifier	Unique identifier assigned to the experiment
measurement type	Type of spectroscopy measurement (EXEM, TIMETRACE, ABS, SINGLE)
sensor identity	Identifier of the SWCNT sensor used in the experiment
sensor concentration	Concentration of the sensor in the measurement solution
analyte identity	Name of the analyte used in the experiment (if applicable)
analyte concentration units	Units used to express analyte concentration
incubation time	Incubation time between sensor and analyte (if applicable)
buffer composition	Composition of the buffer solution (e.g. PBS)

5.2. Data structure metadata

Data structure metadata describes how the raw data files are organized and how the measurements are encoded within the files.

Metadata Field	Description
dataset folder structure	Organization of the dataset into experiment directories
file naming convention	Naming rules used for raw measurement files
data file format	Format of the raw data files (e.g. CSV, TXT)
column naming convention	Names used for columns within raw data files
description of column meaning	Explanation of what each column represents
measurement units	Units associated with measured values

5.3. Acquisition metadata

Acquisition metadata describes the measurement conditions used by the instrument.

These parameters are often automatically generated by spectroscopy instruments and stored in metadata files.

Metadata Field	Description
excitation wavelength	Excitation wavelength used for fluorescence measurements
emission wavelength range	Range of emission wavelengths recorded
detector exposure time	Exposure time used during acquisition
detector gain	Detector amplification setting
detector (camera) temperature	Temperature of the detector during acquisition
plate layout	Mapping between well positions and sample identifiers
scan timing information	Order and timing of individual measurements

6. Metadata Requirements by Measurement Type

Different fluorescence spectroscopy measurement modes require different types of metadata in order to be correctly interpreted.

The following table summarizes the minimum metadata that must be documented for each measurement type defined in this guide.

Measurement types include:

- **EXEM** – excitation–emission fluorescence scans
- **TIMETRACE** – time-resolved fluorescence measurements
- **ABS** – absorbance spectroscopy measurements
- **SINGLE** – single-excitation fluorescence emission spectra

The table below indicates which metadata fields are required for each measurement type.

Metadata Field	EXEM	TIMETRACE	ABS	SINGLE
experiment identifier	✓	✓	✓	✓
measurement type	✓	✓	✓	✓
sensor identity	✓	✓	✓	✓
sensor concentration	✓	✓	✓	✓
analyte identity	—	✓	—	✓
analyte concentration units	—	✓	—	✓
incubation time	—	—	—	✓
buffer composition	✓	✓	✓	✓
dataset folder structure	✓	✓	✓	✓
file naming convention	✓	✓	✓	✓
data file format	✓	✓	✓	✓
column naming convention	✓	✓	✓	✓
description of column meaning	✓	✓	✓	✓
measurement units	✓	✓	✓	✓
excitation wavelength	✓	✓	—	✓
emission wavelength range	✓	✓	—	✓
detector exposure time	✓	✓	—	✓
detector gain	✓	✓	—	✓
detector (camera) temperature	✓	✓	—	✓
plate layout / well mapping	✓	✓	—	✓
scan timing information	✓	✓	—	✓
time interval between scans	—	✓	—	—

7. Raw Data Policy

Datasets shared under this guide should follow the following principles:

1. **Raw data files must be preserved exactly as exported from the instrument.**
   File contents should not be modified or overwritten.
2. **Processed data must not replace raw data.**
   Any processed or derived datasets should be provided separately and must not modify the original measurement files.
3. **Instrument metadata files must be preserved.**
   Metadata generated automatically by the instrument should be included whenever available.
4. **Experiment-level documentation must accompany the raw data.**
   Each experiment directory should contain a README file describing the experimental conditions, sample identifiers, and the structure of the raw data files.

Following these practices ensures that fluorescence spectroscopy datasets remain interpretable, reproducible, and reusable by other researchers.