LNADA Capacity Building Training

Session 1: Introduction & Metadata Fundamentals

23 February 2026 | Lao Statistics Bureau

Welcome to the Training

- **Training:** LNADA Capacity Building
- **Dates:** 23–24 February 2026
- **Location:** Lao Statistics Bureau, Vientiane Capital
- **Trainer:** Frank Lieber, International Database Specialist
- **Project:** LAOSTAT II — Lao PDR Strengthening the National Statistical System

Training Goals

1. **Understand** what metadata is and why it matters for data management
2. **Learn** key metadata standards (DDI, Dublin Core, ISO 19139, SDMX)
3. **Master** the World Bank Metadata Editor to document datasets
4. **Publish** documented data to LNADA (Laos National Data Archive)
5. **Manage** the LNADA platform (admin essentials)

Day 1 Agenda

Session	Time	Topic
Session 1	09:00–10:20	Introduction & Metadata Fundamentals
Session 2	10:40–12:00	Data Preparation & Editor Basics
Lunch	12:00–13:30	
Session 3	13:30–14:45	Documenting Data — Document, Microdata & DDI
Session 4	15:00–16:00	Publishing to LNADA

Day 2 Agenda

Session	Time	Topic
Session 5	09:00–10:20	Day 1 Recap & Additional Data Types
Session 6	10:40–12:00	Cataloging Workflows & LSB Strategy
Lunch	12:00–13:30	
Session 7	13:30–14:45	System Administration
Session 8	15:00–16:00	Configuration, Maintenance & Wrap-up

Meet Your Trainer & Participants

Let's go around the room:
- Name
- Role at your institution
- Experience with data or metadata (if any)

This is an interactive training — please ask questions anytime.

Training Website

Visit: **lnada-training.lsb.lao-stat.de**

On this website you will find:
- Training agenda and schedule
- Training manual (PDF download)
- **`complete-set-of-training-datasets.zip`** — download this now!
- Links to documentation and resources

Important: Download Your Materials Now

The training website will be **offline in 15 days**.

Download today:
- **`complete-set-of-training-datasets.zip`** — all training datasets in one file
- Training manual (PDF)
- Trainer contact information

Don't wait — get these files now while the site is active.

Why We Are Here: LNADA Context

The Challenge:
- The previous LNADA catalog was **lost** during office migration + ransomware attack
- Approximately **35 World Bank datasets** are documented for Laos but need to be re-catalogued
- We must **rebuild from scratch**

The Solution:
- This training enables **LSB staff** to independently manage the LNADA platform
- You will have the skills to document, publish, and maintain data

Key Insight: Data Outlives Tools

"Data has a much longer lifespan than the software tools used to create it."

Consider:

- A survey conducted in 2010 is still valuable in 2026 — 16 years later
- But the original software (SPSS version 2010? Access database?) may no longer work
- **Without documentation, that data becomes unusable**
- With documentation, that data remains valuable forever

Real-World Example: Undocumented Data

Scenario:
You find a folder with survey data from 10 years ago:
- File name: `data_final_v2_REAL_USE_THIS.csv`
- Columns: `age`, `sex`, `inc`, `educ`, `q1`, `q2`, ... `q47`
- Values: `1`, `2`, `3`, `999`, `-1`
- No legend. No questionnaire. No dates.

Question: What does `999` mean? Is it missing data? An error code?

Answer: You don't know. The data is unusable without someone who remembers.

The FAIR Principles

FAIR data is:

- **F**indable — can be discovered through data catalogs
- **A**ccessible — can be retrieved via standard protocols
- **I**nteroperable — can be combined with other datasets
- **R**eusable — has sufficient documentation for reuse

Metadata makes data FAIR. go-fair.org/fair-principles

Metadata Makes Data AI-Ready

Modern data analysis relies on automated tools:
- Machine learning algorithms
- Data integration systems
- Search and discovery engines

These tools require **structured, standardized metadata** to work effectively.

Without metadata: Data is invisible to automated systems
With metadata: Data becomes discoverable and usable at scale

What Is Metadata?

Metadata = Data about data

Examples:
- What is this dataset? (title, abstract)
- Who created it? (author, institution)
- When was it collected? (dates)
- What variables does it contain? (data dictionary)
- How was it created? (methodology)
- How can it be accessed? (data access terms)

Four Types of Metadata

1. **Cataloguing** — what datasets exist (discovery)

2. **Descriptive** — what the data is about (content)

3. **Structural** — how data is organized (relationships)

4. **Administrative** — rights, dates, who manages it (management)

All four types work together. A complete metadata record includes all four.

Metadata Standards: An Overview

The world has agreed on **standard ways to document data**. Different standards for different data types.

We'll cover the five most relevant to your work at LSB:

1. **DDI Codebook 2.5** — microdata and surveys
2. **Dublin Core** — documents and general resources
3. **ISO 19139** — geographic data
4. **SDMX** — statistical data exchange
5. **Others** — MARC21, IPTC, Schema.org, DCAT, PROV-O

Key Standard 1: DDI Codebook 2.5

DDI = Data Documentation Initiative

Use for: Microdata datasets, household surveys, census data, individual-level records

Covers:
- Study description (who, when, where, why)
- Methodology (sampling, data collection)
- Data dictionary (variable names, labels, value labels)
- External resources (questionnaires, reports)

Example: A household survey with 5,000 records documenting income, education, health

Key Standard 2: Dublin Core

Use for: Documents, reports, publications, images, any general resource

Covers:
- Title, creator, date
- Subject, description
- Language, rights
- Minimal but flexible

Example: A research report on poverty trends in PDF format

The beauty of Dublin Core: it works for almost anything.

Key Standard 3: ISO 19139 (Geographic Data)

Use for: Maps, shapefiles, geographic datasets, spatial boundaries

Covers:
- Spatial extent and reference systems
- Data quality and lineage
- Temporal resolution

Example: Provincial boundary maps of Laos in GIS format

Key Standard 4: SDMX (Statistical Data Exchange)

SDMX = Statistical Data and Metadata Exchange

Use for: Statistical time series, aggregate data, indicators

Covers:
- Definitions of dimensions (country, time, indicator)
- Concepts and code lists
- Data structure definitions

Example: National GDP series by quarter, 2010–2026

Other Standards Worth Knowing

- **MARC21** — library metadata (books, theses)
- **IPTC** — news, multimedia, press photos
- **Schema.org** — web-friendly semantic markup
- **DCAT** — data catalog standards (EU, governments)
- **PROV-O** — data provenance (lineage and workflow tracking)

You won't need to master these, but it's good to know they exist.

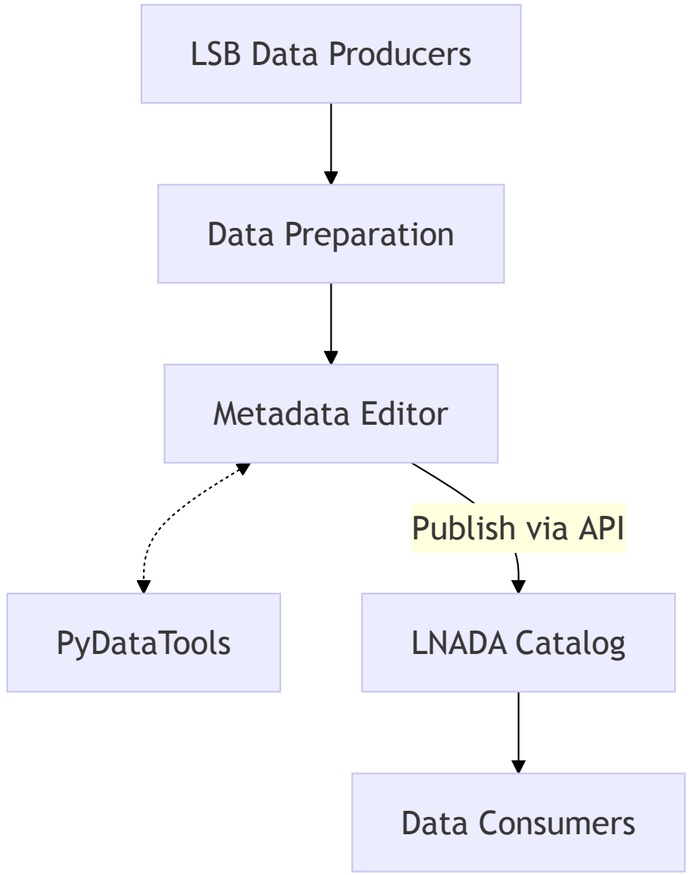
The NADA & Metadata Editor Ecosystem

NADA = National Data Archive
- A **catalog and dissemination platform** for data
- Hosted on government servers
- Provides public access to datasets

LNADA = Laos NADA instance (at LSB)

Metadata Editor = Tool to create structured metadata
- Open-source software by the World Bank
- Web-based interface for entering metadata
- Automatically extracts metadata from data files
- Publishes to NADA via API

LNADA & Metadata Editor Overview Diagram



Metadata Editor Supported Data Types (8 Total)

Data Type	Format	Standard	Example
Microdata	.dta, .sav, .csv	DDI Codebook 2.5	Household survey
Indicators	.csv, time series	WB schema	GDP, poverty rates
Geographic	Shapefile, GeoJSON	ISO 19139	Provincial boundaries
Statistical Tables	.xlsx, .csv	SDMX	Aggregate statistics
Documents	.pdf, .docx	Dublin Core / MARC21 / BibTex	Research reports
Images	.jpg, .png	IPTC	Photos, infographics
Videos	.mp4, .mov	Schema.org	Training videos
Scripts	.r, .py, .ipynb	Dedicated schema	Data cleaning code

Questions So Far?

- Anything about metadata?
- Anything about standards?
- Anything about the training plan?

Speak up — these are foundational concepts.

Thank You & Let's Begin!

Session 1 Complete (80 min)

Next break: 10:20–10:40

Then Session 2: Data Preparation & Metadata Editor Basics