

EOSC CZ Summer School for Data Stewards 2026

Python Toolbelt for Data Stewards in JupyterLab

Introduction

Matthias Täschner 

Data Science Center ScaDS.AI, Leipzig University, Germany



Co-funded by
the European Union



Ministry of Education,
Youth and Sports
of the Czech Republic



CC BY

Agenda

Session 1 (June 16, 13:30–17:00)

- Python projects with uv
- JupyterLab and notebooks
- AI assistant in JupyterLab
- Data steward tasks in JupyterLab

“Get to know your tools”

General handling and understanding

Session 2 (June 17, 08:30–12:00)

- Python-based toolkit and pipeline
- Data exploration and preparation
- Data validation and metadata
- Documentation, DMP
[Export via API usage, automation]

“Use your tools productively”

Detailed usage of the Python Toolbelt

Training Materials

Public GitHub Pages

- Notebooks and files for the sessions
- <https://scads.github.io/eosc-ssds-2026>

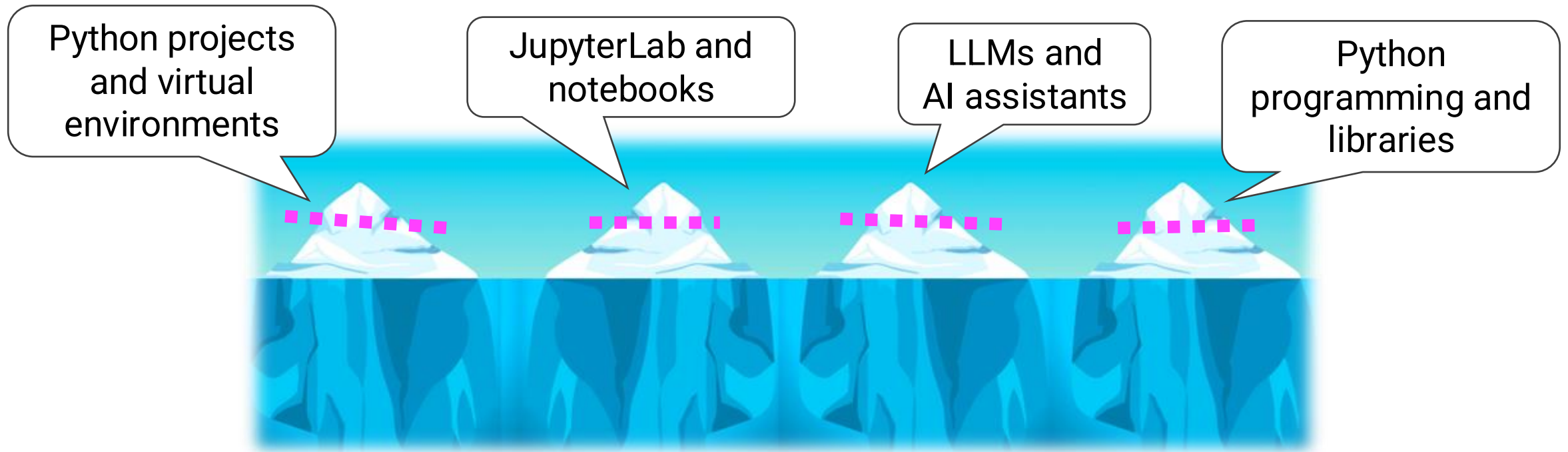
Nextcloud Share

- Additional helpful resources
- <https://urlr.me/qE92Ny>



Expectation management

Teaching the fundamentals as a foundation for further independent learning and application.



Learning Objectives Session 2

Introduction to Python tools for:

- ***bia-bob*** – AI coding assistant for JupyterLab
- ***Data validation***
- ***Data quality reports***
- ***Data cleaning***
- ***Metadata and data packaging***
- ***[optional] API use for data deposition (Zenodo sandbox)***

Agenda Session 2

- **08:30 – 08:45** **Welcome & Introduction** ✓
- **08:45 – 09:15** **Technical Setup Check, uv**
- **09:15 – 10:00** **Toolbelt overview and dataset**
- **10:00 – 10:30** ☕ **Coffee Break**
- **10:30 – 11:30** **Tools for data validation, cleaning, reporting, metadata**
- **11:30 – 11:50** **DMP and data deposit [optional]**
- **11:50 – 12:00** **Wrap-Up**

*Any
Questions?*

EOSC CZ Summer School for Data Stewards 2026

Python Toolbelt for Data Stewards in JupyterLab

Data Steward Pipeline

Matthias Täschner 

Data Science Center ScaDS.AI, Leipzig University, Germany



Co-funded by
the European Union



Ministry of Education,
Youth and Sports
of the Czech Republic



CC BY

Data Steward Pipeline



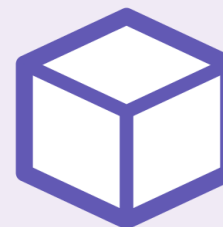
Raw data
Messy CSV
handover



Validate
Schema =
Data contract



Report
Basic cleaning &
Quality report



Package
Metadata &
FAIR data package



DMP & Deposit
Documented &
DOI



AI Assistant

Drafts the code for each step using the appropriate tools – you read, run & verify

EOSC CZ Summer School for Data Stewards 2026

Python Toolbelt for Data Stewards in JupyterLab

Wrap Up

Matthias Täschner 

Data Science Center ScaDS.AI, Leipzig University, Germany



Co-funded by
the European Union



Ministry of Education,
Youth and Sports
of the Czech Republic



CC BY