Sharing is encouraged

NFDI4Earth Pilot

SoilPulse

-

Towards Fair soil process data

→ 2min Madness

→ More on aims

→ More on data and issues

→ Metadataschema

→ (Meta-)Data queries

→ See the team

→ Even more on SoilPulse (Links)

1 Jonas Lenz, Jan Devátý, Conrad Jackisch

Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation

NFDI4Earth

TU Bergakademie Freiberg TUBAF Die Ressourcenuniversität. Seit 1765.

Faculty of Civil Engineering CTU in Prague

IPROconsult

Sharing is encouraged

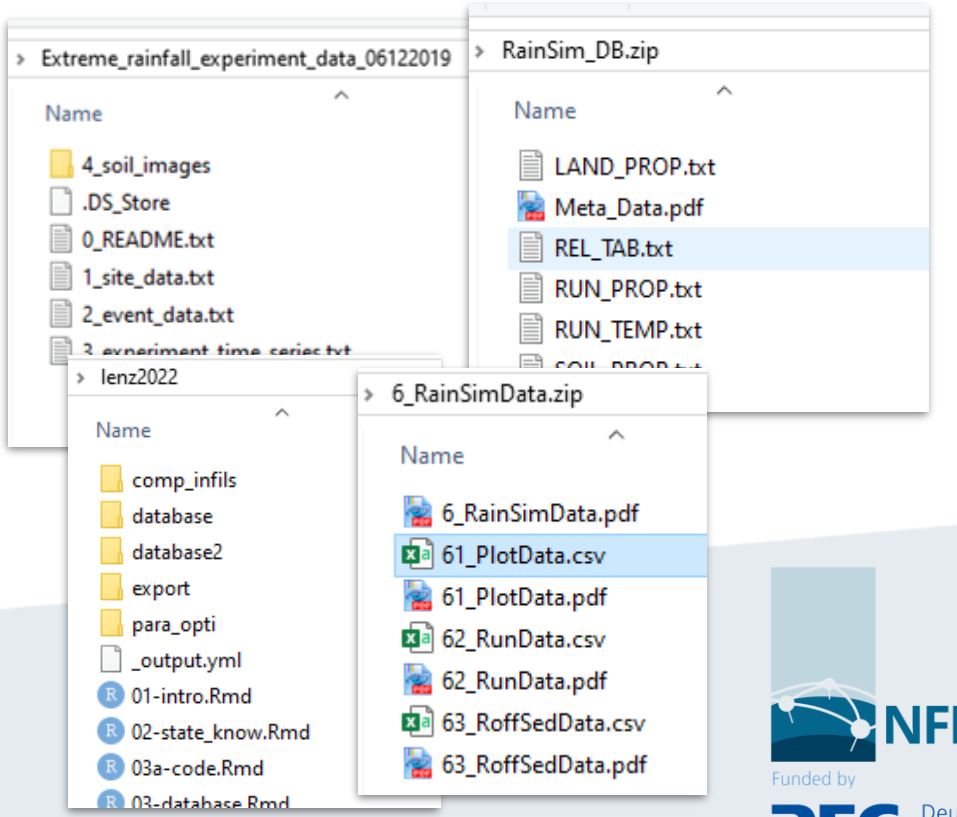
SoilPulse - Motivation

Topic: Observation data for model calibration

- e.g. rainfall-runoff simulations in field or lab

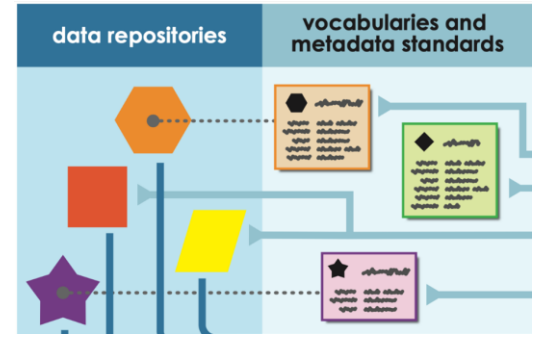
Issue: Missing standards

- In data management
- (In experimental methods)



State of data resources

- incompatible
- unFAIR
- unpublished



Jonas Lenz, Jan Devátý, Conrad Jackisch

NFDI4Earth
Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation
TU BERGAKADEMIE FREIBERG TUBAF Die Ressourcenuniversität. Seit 1765.
FACULTY OF CIVIL ENGINEERING CTU IN PRAGUE
IPROconsult

Navi-gation

Sharing is encouraged

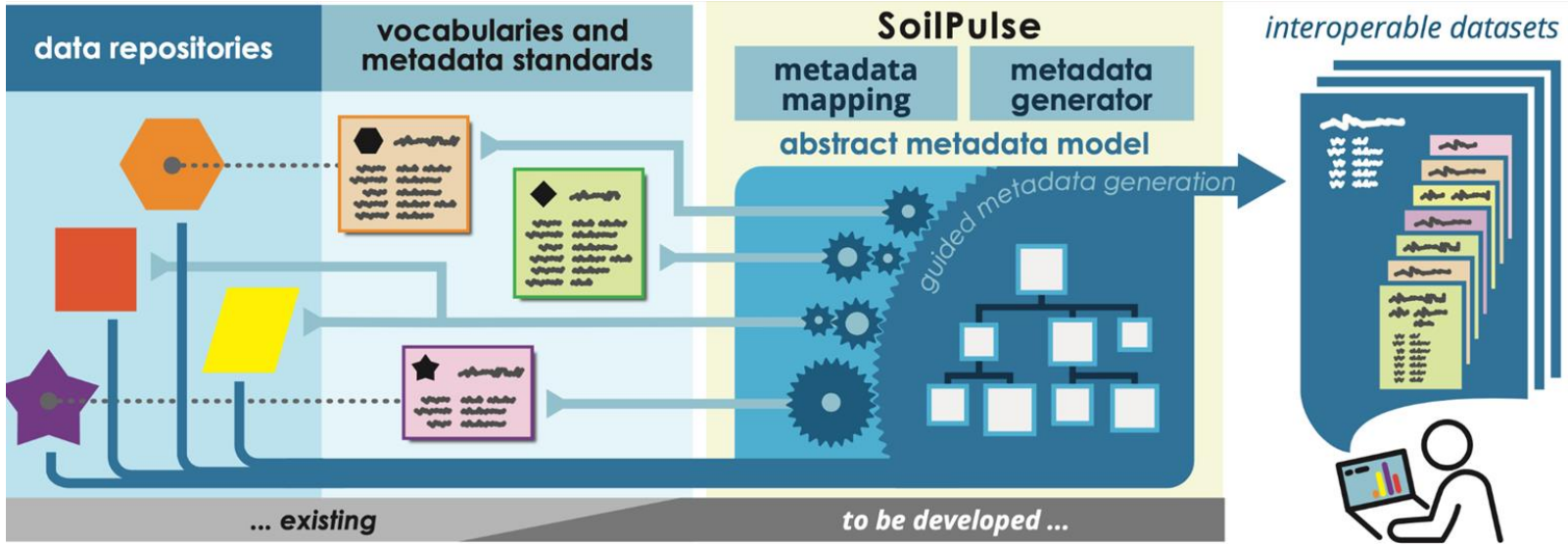
SoilPulse - Aims

1 - Data harmonization assisted by a tool for metadata generation.

2 - Allow for semi-automatic analysis through (meta-)data querying.

3 - Learn from others (meta-)data needs

Make (existing) data sets reusable!



You are an experimenting / data wrangling / modelling / interested person?

–> See you on the screen!

NFDI4Earth
Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation
TU BERGAKADEMIE FREIBERG TUBAF Die Ressourcenuniversität. Seit 1765.
FACULTY OF CIVIL ENGINEERING CTU IN PRAGUE
IPROconsult

TUBAF
Die Ressourcenuniversität.
Seit 1765.

FACULTY OF CIVIL ENGINEERING
CTU IN PRAGUE

IPROconsult

NFDI4Earth
Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

NFDI4Earth Pilot:

SoilPulse Team

Navi-
gation

SoilPulse @ EGU 2024

April 16th 2024

jonas.lenz@iproconsult.com

conrad.jackisch@tbt.tu-freiberg.de

jan.devaty@fsv.cvut.cz

SoilPulse - Aims 1/2

Metadata generation in a web interface (Demo: https://soilpulse-egu.streamlit.app/Metadata_retriever)

- User/Data creator provides files (**non standardized structure**), metadata shall be generated as automatically as possible, so the user “only” needs to approve and complete it.
  → Data structure needs to be mapped within metadata.
  → User gets feedback how the machine understands his data, while preparing the metadata.
  → User gets feedback if his data complies with data of other resources.
  → Semi-automatic generation of submission ready metadata to data files.

- Also applicable to **already published** resources (e.g. Datasets on Zenodo) -> Reference to the resource is then included in metadata to avoid republication.

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Navi-
gation

Sharing is encouraged

SoilPulse - Aims 2/2

(Meta-)Data querying (Demo: https://soilpulse-egu.streamlit.app/Explorer)

- Metadata becomes access point for data
- Making data points queryable: “Get runoff values from all rainfall simulation experiments with total organic carbon content > 3% of a soil sample.”
- Feed data aggregates to models by defining model requirement templates.
- Combination with data from other resources.

- Requires (self-hosted) live system/ server holding (temporarily) all data.

Which Metadata do we need to generate in addition to existing metadata schemas to increase data reusability?

Funded by
DFG Deutsche Forschungsgemeinschaft German Research Foundation

SoilPulse - Metadata schema 1/2

Adaptation of bonares metadata schema (Gärtner et al. 2017)

- Extensive soil specific schema, building upon INSPIRE and DataCite (https://doi.org/10.1016/j.cageo.2019.07.005)
- Bonares has metadata down to table structure and relation of tables

Extension:

- Assignment of controlled vocabulary concepts to single data points/columns:
  - e.g. “SOC”/”TOC”/”Corg” of original datasets becomes “total organic carbon” of AGROVOC (http://aims.fao.org/aos/agrovoc/c_c35fdd26).
- Make metadata within files readable for machines (e.g. table structure, timesteps, experiment ID). → Map down to single values.

Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation

SoilPulse - Metadata schema 2/2

Implementation (in progress):

- Devátý, J., Lenz, J., and Jackisch, C.: SoilPulse – A software package for semi-automated metadata management and publication, EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-18775, https://doi.org/10.5194/egusphere-egu24-18775, 2024.

- will be available as python package

SoilPulse - Data and issues 1/2

Soil, Erosion/Infiltration Experiments

- at the boundary between hydrology, agriculture and soil properties
- various process' observation
- state dependent (intial water content, plant development, …)
- functional characteristics of soils/ experimental sites

Data types

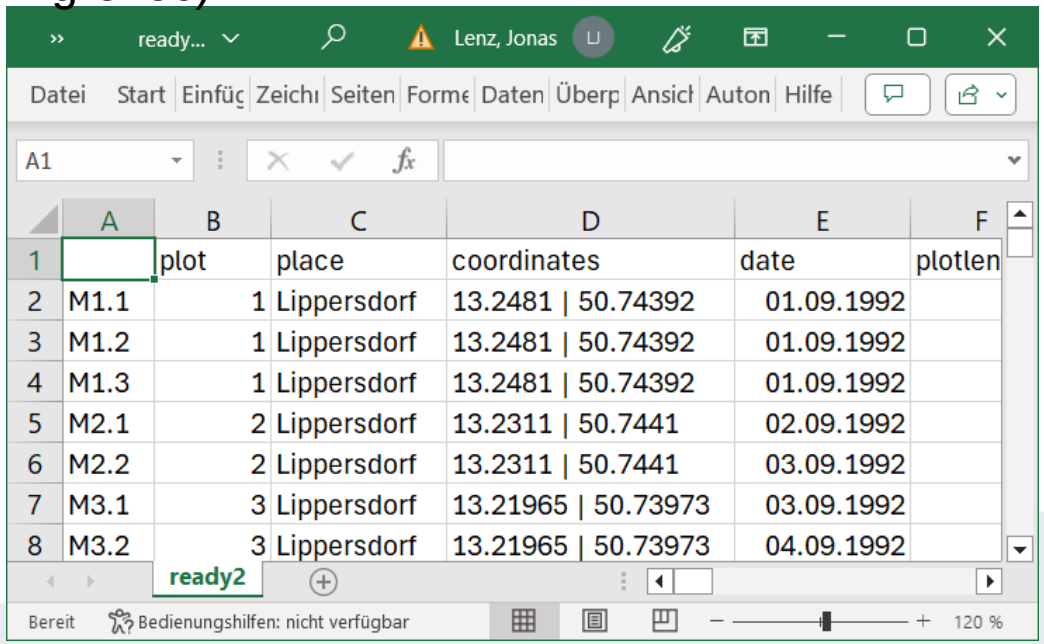
- single measurements of soil properties
- descriptions of treatment (last or history), plant development
- time series of processes (runoff, irrigation intensity)
- images
- …

Funded by
Deutsche Forschungsgemeinschaft
German Research Foundation

SoilPulse - Data and issues 2/2

Missing standards in datamanagement

- Differing table structures → capture table reading instructions in metadata
- Differing namings → Use controlled Vocabulary (e.g. Agrovoc)



ready...

Lenz, Jonas

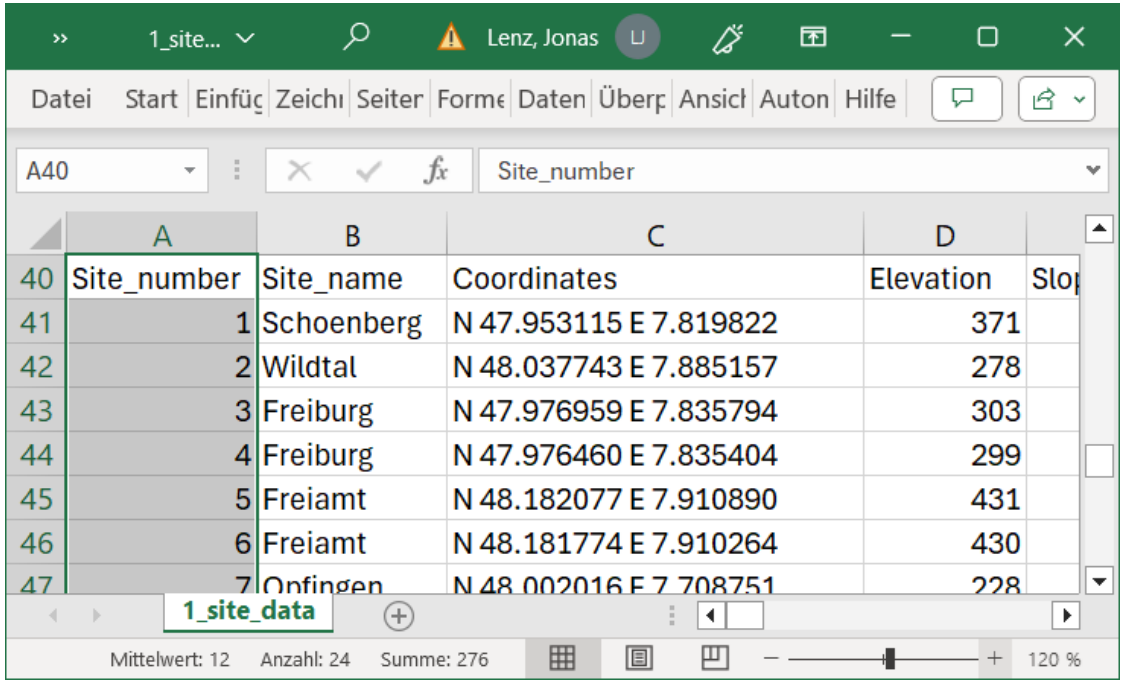
Datei Start Einfüg Zeichn Seiten Forme Daten Überp Ansicl Auton Hilfe

A1

	A	B	C	D	E	F
1		plot	place	coordinates	date	plotlen
2	M1.1	1	Lippersdorf	13.2481 | 50.74392	01.09.1992	
3	M1.2	1	Lippersdorf	13.2481 | 50.74392	01.09.1992	
4	M1.3	1	Lippersdorf	13.2481 | 50.74392	01.09.1992	
5	M2.1	2	Lippersdorf	13.2311 | 50.7441	02.09.1992	
6	M2.2	2	Lippersdorf	13.2311 | 50.7441	03.09.1992	
7	M3.1	3	Lippersdorf	13.21965 | 50.73973	03.09.1992	
8	M3.2	3	Lippersdorf	13.21965 | 50.73973	04.09.1992	

ready2

Bereit Bedienungshilfen: nicht verfügbar 120 %



1_site...

Lenz, Jonas

Datei Start Einfüg Zeichn Seiten Forme Daten Überp Ansicl Auton Hilfe

A40 Site_number

	A	B	C	D
40	Site_number	Site_name	Coordinates	Elevation
41	1	Schoenberg	N 47.953115 E 7.819822	371
42	2	Wildtal	N 48.037743 E 7.885157	278
43	3	Freiburg	N 47.976959 E 7.835794	303
44	4	Freiburg	N 47.976460 E 7.835404	299
45	5	Freiamt	N 48.182077 E 7.910890	431
46	6	Freiamt	N 48.181774 E 7.910264	430
47	7	Opfingen	N 48.002016 E 7.708751	228

1_site_data

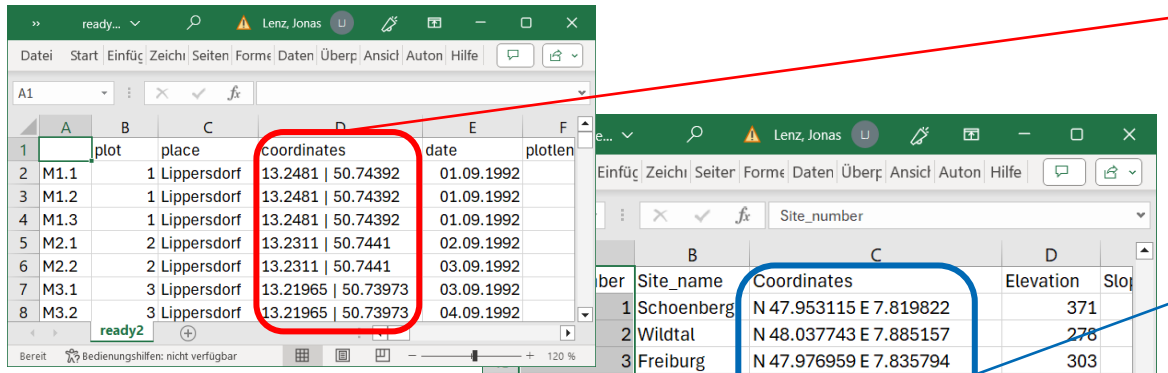
Mittelwert: 12 Anzahl: 24 Summe: 276 120 %

Jonas Lenz, Jan Devátý, Conrad Jackisch

TU Bergakademie Freiberg
TUBAF Die Ressourcenuniversität. Seit 1765.
NFDI4Earth
Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation
Faculty of Civil Engineering CTU in Prague
IPROconsult

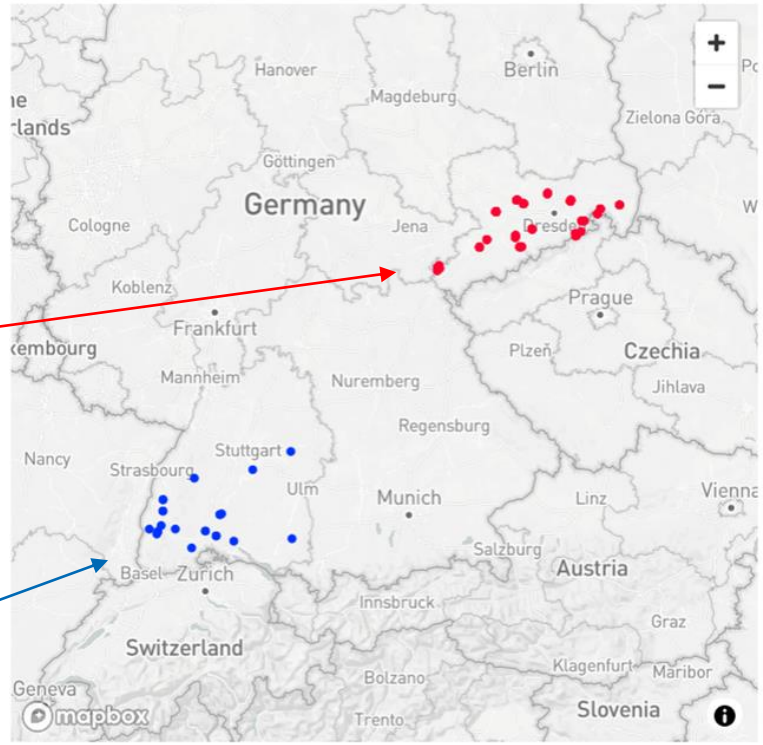
SoilPulse - (Meta-)Data queries 1/1

- Query multiple data sets at once
- Define your queries in templates
- Easier analysis



	plot	place	coordinates	date	plotlen
2	M1.1	1 Lippersdorf	13.2481 | 50.74392	01.09.1992	
3	M1.2	1 Lippersdorf	13.2481 | 50.74392	01.09.1992	
4	M1.3	1 Lippersdorf	13.2481 | 50.74392	01.09.1992	
5	M2.1	2 Lippersdorf	13.2311 | 50.7441	02.09.1992	
6	M2.2	2 Lippersdorf	13.2311 | 50.7441	03.09.1992	
7	M3.1	3 Lippersdorf	13.21965 | 50.73973	03.09.1992	
8	M3.2	3 Lippersdorf	13.21965 | 50.73973	04.09.1992	

	Site_name	Coordinates	Elevation
1	Schoenberg	N 47.953115 E 7.819822	371
2	Wildtal	N 48.037743 E 7.885157	278
3	Freiburg	N 47.976959 E 7.835794	303
44	4 Freiburg	N 47.976460 E 7.835404	299
45	5 Freiamt	N 48.182077 E 7.910890	431
46	6 Freiamt	N 48.181774 E 7.910264	430
47	7 Opfingen	N 48.002016 E 7.708751	228



SoilPulse - Links

Navi-gation

Sharing is encouraged

More on the project:

https://soilpulse.github.io/

Test it on streamlit:

https://soilpulse-egu.streamlit.app

Get in touch:

mailto:conrad.jackisch@tbt.tu-freiberg.de

Abstract for EGU 2024:

https://doi.org/10.5194/egusphere-egu24-19497

Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation

NFDI4Earth

TU Bergakademie Freiberg TUBAF Die Ressourcenuniversität. Seit 1765.

Faculty of Civil Engineering CTU in Prague

IPROconsult