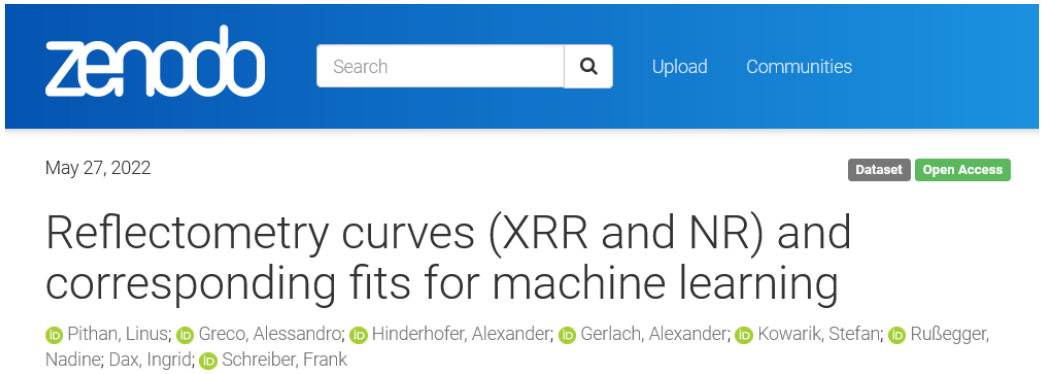
How to extend the existing experimental dataset?

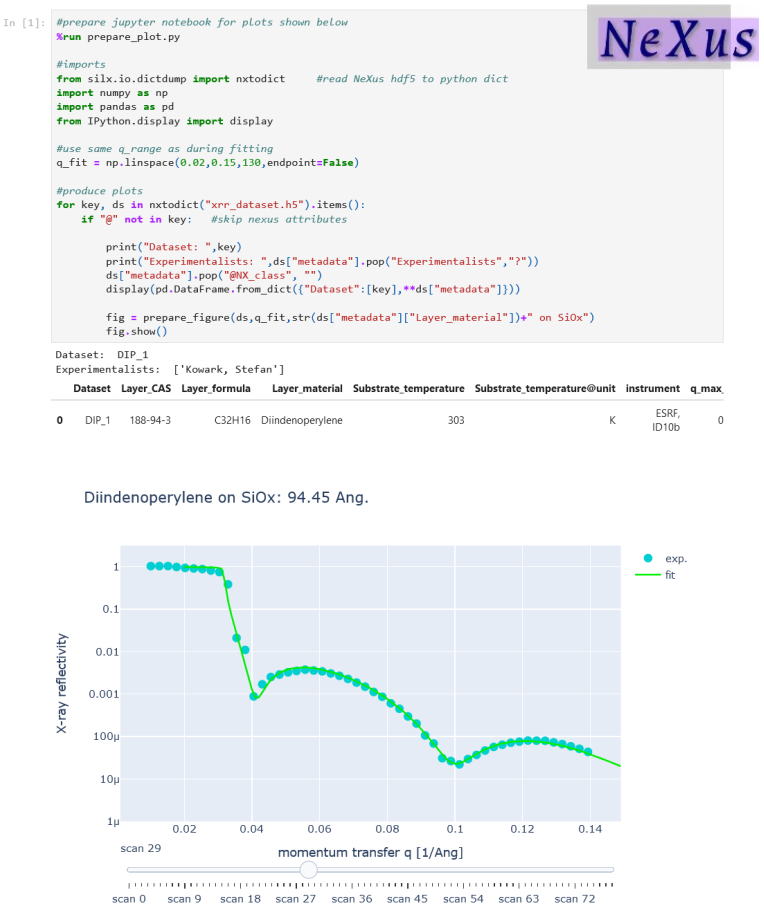
Current status:



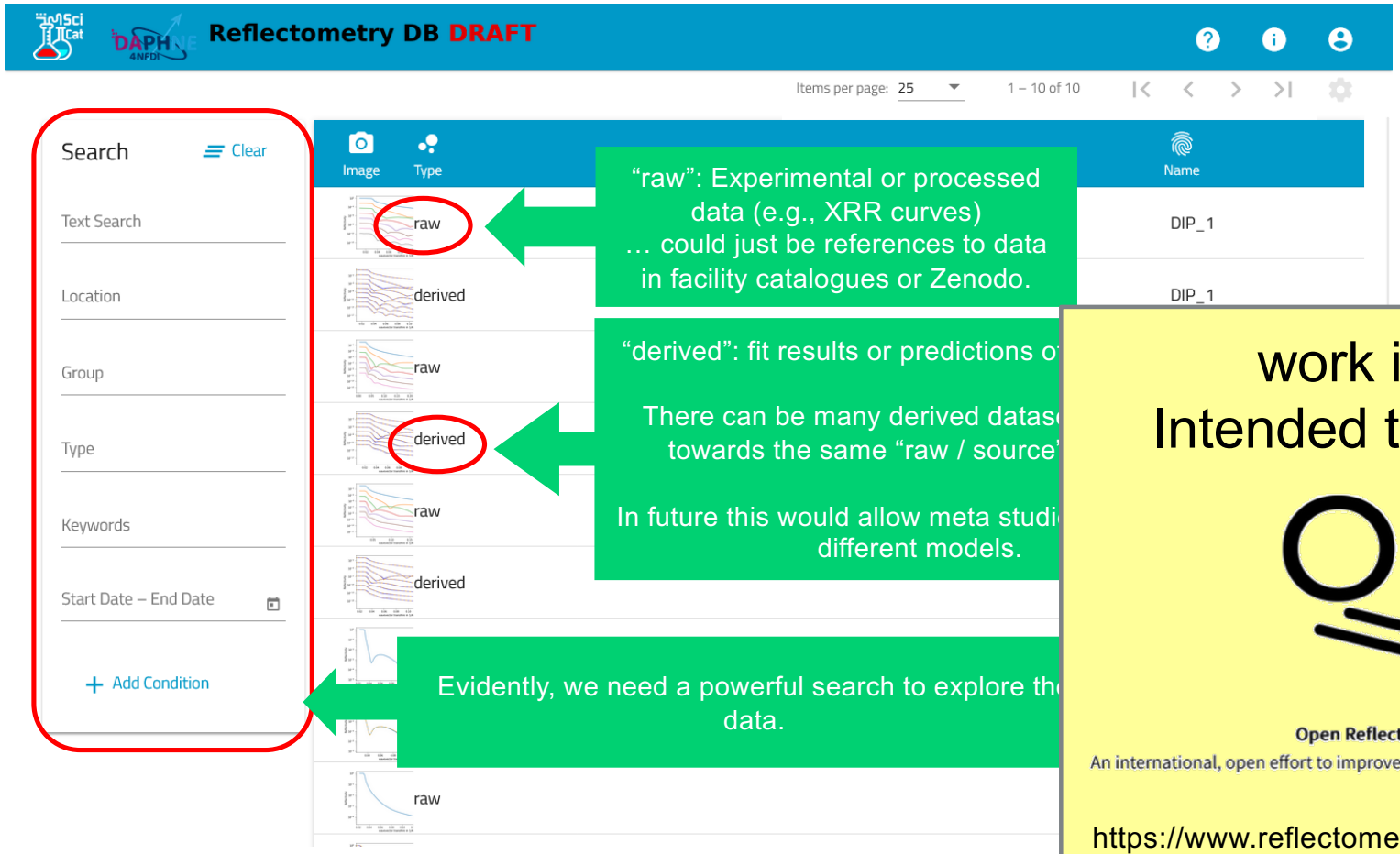
Public reflectometry data collection

- about 250 experimental XRR profiles
- sample parameters & classical fits are published together with measurements (in ML-context: “labelled data”)

L.Pithan et al. 2022 https://doi.org/10.5281/zenodo.6497438
https://schreiber-lab.github.io/reflectometry-dataset/



Idea: Use a tool like SciCat to collect reduced reflectometry data



work in progress –
Intended to be proposed to

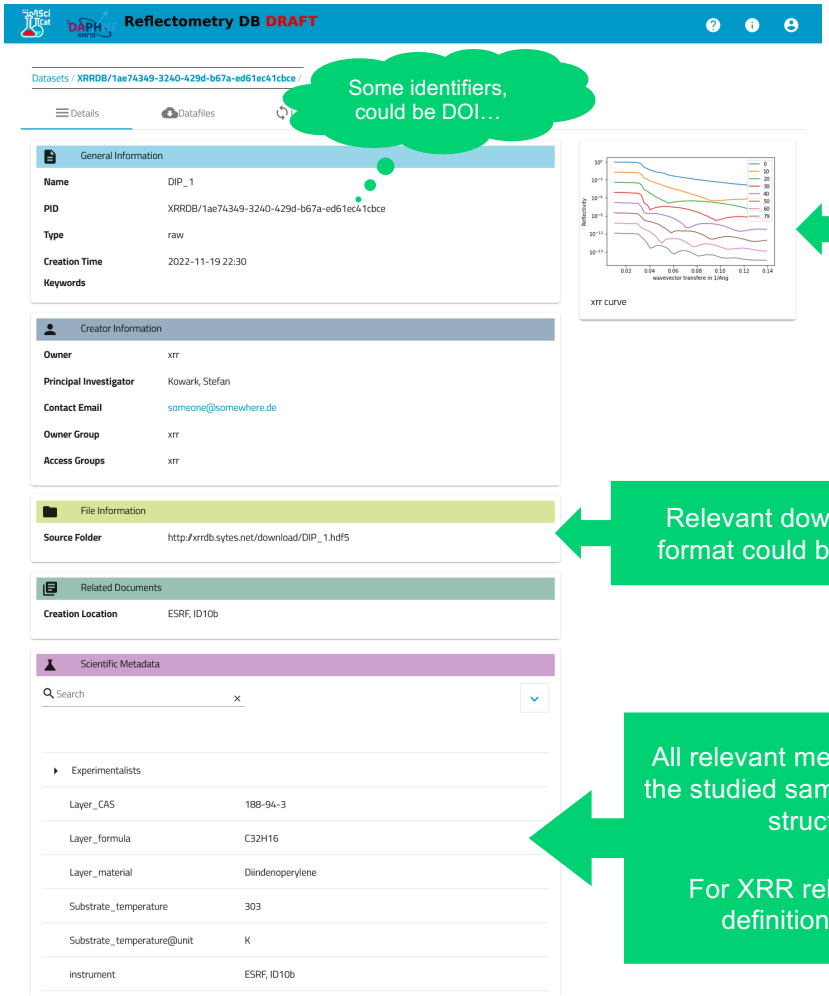
ORSO

Open Reflectometry Standards Organisation

An international, open effort to improve the scientific techniques of neutron and X-ray reflectometry

https://www.reflectometry.org/

Display of raw and processed data

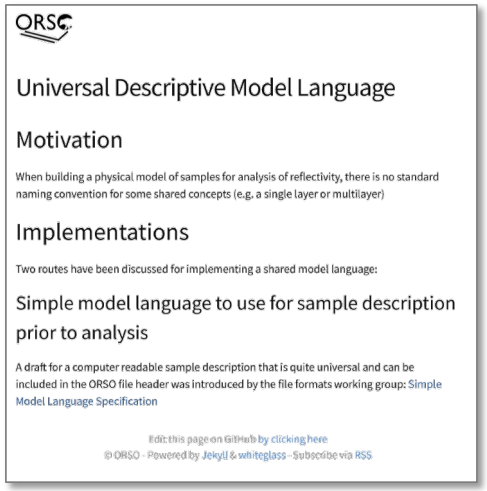


Graphical preview of the dataset

Relevant download info. For XRR the file format could be ORSO .txt or ORSO .hdf5

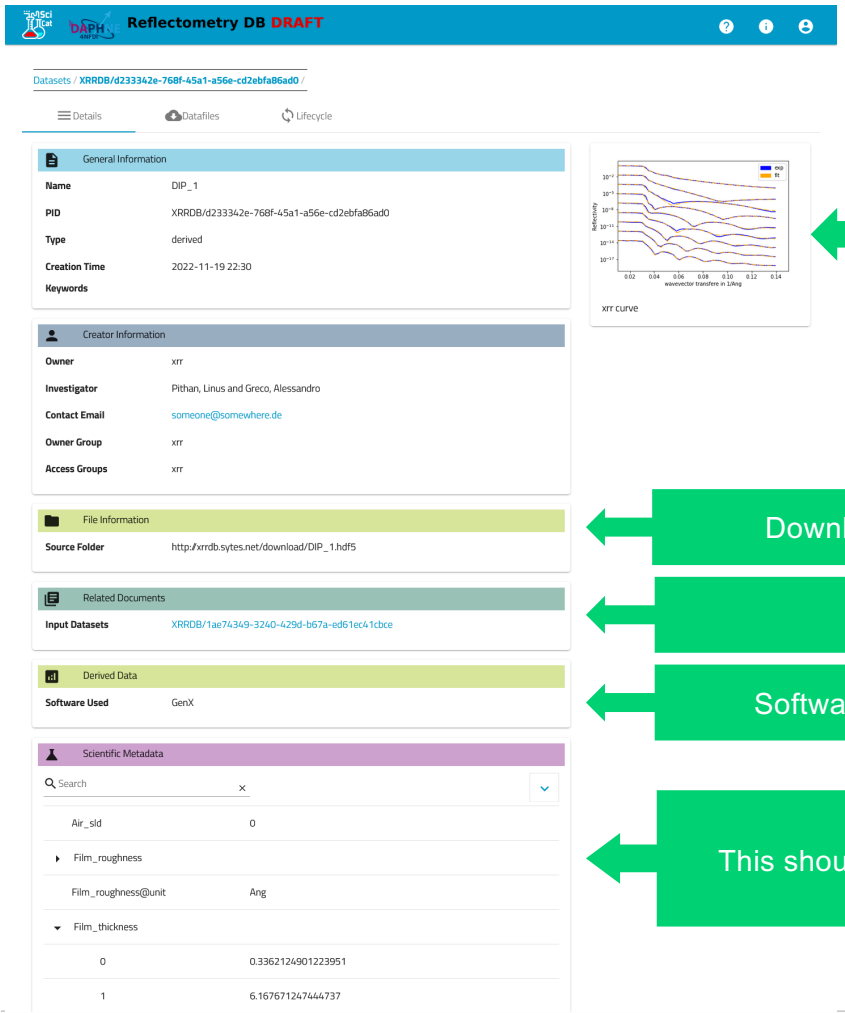
All relevant metadata to find the dataset and the studied sample (materials, expected layer structure, thicknesses…)

For XRR relatively easy: follow ORSO definitions and extend them later



https://www.reflectometry.org/projects/model_language/

Display of analysed data



Graphical preview of analysis & fit results

Download link for analysed data

Link to raw dataset

Software used for analysis / fitting

Searchable metadata describing the analysis.
This should rather contain a machine-readable descriptive summary of the analysis than the fit itself…

Why do we care about this?

- Collect open data for machine learning efforts
- Systematic comparison of different analysis tools
- Long-term: Build reproducible processing chains