



Contribution ID: 171

Type: **Oral Presentation**

Scipp: Scientific data handling with labeled multi-dimensional arrays for C++ and Python

Wednesday, 16 October 2019 11:15 (25 minutes)

Scipp is heavily inspired by the Python library “xarray”. It enriches raw NumPy-like multi-dimensional arrays of data by adding named dimensions and associated coordinates. Multiple arrays are combined into datasets. On top of these basics, which xarray would also provide, scipp introduces:

- Implicit handling of physical units.
- Implicit propagation of uncertainties.
- Support for histograms, i.e., bin-edge coordinate axes, which exceed the data’s dimension extent by one.
- Support for event data, a particular form of sparse data with arrays of random-length lists, arising when collecting neutron scattering data in event-mode.

Altogether, the features enable a more natural and more concise user experience.

The combination of named dimensions, coordinates, and units helps to drastically reduce the risk for programming errors.

The core of scipp is written in C++ to open opportunities for performance improvements that a Python-based solution would not allow for. On top of the C++ core, scipp’s Python components provide functionality for plotting and other visualization, e.g., for use in Jupyter Notebooks.

Scipp is currently developed at the European Spallation Source. The released version 0.1 is for experimental use and not production-ready yet. See also <https://scipp.readthedocs.io> for detailed documentation.

Primary author: Dr HEYBROCK, Simon

Presenter: Dr HEYBROCK, Simon

Session Classification: Software

Track Classification: Software