HZB Helmholtz Zentrum Berlin

Dataset Parameters at HZB

Heike Görzig Rolf Krahl

ICAT F2F Meeting, 03 May 2023, Berlin

Problem Statement & Motivation

- Lacking standardization of parameters for search accross facilities.
- PaNOSC search portal defines only very limited set of parameters.
- Goal: work towards a common PaN Metadata Standard that could for instance also be exposed via OAI-PMH.
- ICAT schema imposes one global list of parameter names within the facility.

ICAT Schema: ParameterType

ParameterType attributes (selection):

Field	Type
pid	String [255]
name	String [255] NOT NULL
units	String [255] NOT NULL
valueType	ParameterValueType NOT NULL
description	String [255]
unitsFullName	String [255]

NeXus File

NeXus hierarchical structure:

```
entry:NXentry
    instrument:NXinstrument
        source:NXsource
            current_set:NX_FLOAT64 = Current
                @units = "mA"
            type:NX_CHAR = "Synchrotron X-ray Source"
            probe:NX_CHAR = "x-ray"
            name:NX_CHAR = "BESSY II"
            current:NX_UINT32 = Ringcurr
                @units = "mA"
        monochromator:NXmonochromator
            energy: NX_FLOAT = 850
            @units="eV"
        si_detector:NXdetector
            short_name:NX_CHAR = "si_detector"
            type:NX_CHAR = "si diode"
            data:NX_FLOAT64 = 0
                @units="V"
        apd_detector:NXdetector
            type:NX_CHAR = "APD"
            short_name:NX_CHAR = "APD detector"
            channel_1:NX_FLOAT64 = 0
                @units="V"
            channel_2:NX_FLOAT64 = 0
                @units="V"
            channel_3:NX_FLOAT64 = 0
                @units="V"
```

NeXus Ontology

- The NeXus project is working on a NeXus ontology.
- All elements of the NeXus vocabulary (appdefs, base classes, groups, fields, etc.) have their own individual concept attached.
- This ontology may (in the future) help to translate queries into some canonical form.

Use NeXus Path as Parameter Name

Considerations:

- Same parameter should use the same type accross instruments.
- Need some naming convention that works for all the instruments at the facility.
- Standardize units: don't use multiple parameter types having different units for (semantically) the same parameter.

Use NeXus Path as Parameter Name

Suggestion:

- Use the flattened NeXus path as parameter name.
- Ideally, this would require NeXus application definitions for all experiments.
- Fall back to basic NeXus classes where application definitions are not (yet) available.
- Add a numerical suffix if needed to disambiguate multiple instances having otherwise the same path.
- Use a dedicated name space for NeXus parameter types.
- Stick to SI units (with a very few exceptions), don't use SI prefixes.
- Set the NeXus ontology IRI in the pid attribute.

Example

Example dataset parameter:

type.name	value	type.units
nxs/entry/instrument/insertion_device/gap	0.029040	m
nxs/entry/instrument/insertion_device/shift	0.012834	m
nxs/entry/instrument/monochromator/energy_set	863.05	eV
nxs/entry/instrument/slitwidth/x_gap	0.00015000	m
nxs/entry/instrument/detector_1/type	si diode	N/A
nxs/entry/instrument/detector_2/type	APD	N/A
nxs/entry/instrument/detector_2/short_name	APD detector	N/A
nxs/entry/instrument/detector_2/flux	1.1699e+12	Hz

Discussion & Open Issues

- Should we aim for standardization within the ICAT project?
- Other suggestions?
- The transition from NeXus base classes to application definition may be an issue.
- Need to seek cooperation and harmonization with other projects, e.g. NOMAD.
- How to deal with parameters that do not fit into NeXus definitions?