

The NeXus Data Format for muon Spectroscopy and Neutron or X-ray Scattering

Mark Könnecke

Paul Scherrer Institute
Switzerland

January 11, 2010

Why a Common Dataformat?

- Today:
 - Lots of different data formats
 - Time wasted converting data
 - Old formats no longer capable of delivering for new high throughput detectors
 - Difficult to add additional data
 - Often, for DA multiple different files needed
 - Badly documented formats
- Tomorrow, with NeXus:
 - Single, efficient, platform independent data format
 - All information in one file
 - Self describing
 - Extendable

- Devised from three independent proposals by Jonathan Tischler, APS, Przemek Klosowski, NIST and Mark Koennecke, ISIS, PSI in 94-96
- Improved during various NOBUGS conferences
- NeXus International Advisory Committee, NIAC, since 2003
- Since 2003 yearly meetings of the NIAC
- We already considered many issues!

- 1 Physical file format and API for accessing files
- 2 Rules for storing data in files
- 3 Component and application definitions
- 4 NeXus Utilities

- Portable, self describing, extendable, public domain
- Hierarchical data format, NCSA, HDF-4, later HDF-5
- HDF-5:
 - grouping support
 - on the fly compression
 - reading/writing subsets
 - first dimension appendable
 - Public domain C, F77 access library
 - Used by: NASA, Boing, the weathermen,
- XML

- HDF libraries: complex API
- NeXus-API hides HDF complexity
- Transparent access to all supported physical file formats
- ANSI-C implementation
- Bindings: C++, F77, Java, python, IDL, SWIG
- January, 4, 2010: 1311217 files processed at PSI alone

```
nxfile = nxs.open('hrpt2008n152088.hdf','r')
nxfile.openpath('/entry1/data1/two_theta')
x = nxfile.getdata()
nxfile.openpath('/entry1/data1/counts')
y = nxfile.getdata()
nxfile.openpath('/entry1/title')
txt = nxfile.getdata()
nxfile.close()

plot(x,y)
xlabel('two theta')
ylabel('counts')
title(txt)
show()
```

- Files
- Groups identified by name and a classname beginning with NX
- Scientific data sets
- Attributes
- Links

Rules for Storing Data in NeXus Files

- NeXus files have a hierarchy
- NXentry
 - NXuser
 - NXsample
 - NXmonitor
 - NXdata
 - NXinstrument
 - NXmonochromator
 - NXdetector
 - ...
- Units have to be given
- Rules to locate axis data for multi dimensional data sets

- Supports self description and allows short names in components

- Supports self description and allows short names in components
- Name, classname pair allows for multiple components of the same type

- Supports self description and allows short names in components
- Name, classname pair allows for multiple components of the same type
- NXentry allows for multiple datasets in the same file

- Supports self description and allows short names in components
- Name, classname pair allows for multiple components of the same type
- NXentry allows for multiple datasets in the same file
- NXdata supports automatic plotting

CIF uses Hierarchies too, but hides them:

```
_exptl_crystal_description plate  
_exptl_crystal_colour colourless  
_exptl_crystal_size_max 0.30
```

CIF uses Hierarchies too, but hides them:

```
_exptl_crystal_description plate  
_exptl_crystal_colour colourless  
_exptl_crystal_size_max 0.30
```

```
/exptl/crystal/description plate  
/exptl/crystal/colour colourless  
/exptl/crystal/size/max 0.30
```

NeXus Component and Application Definitions

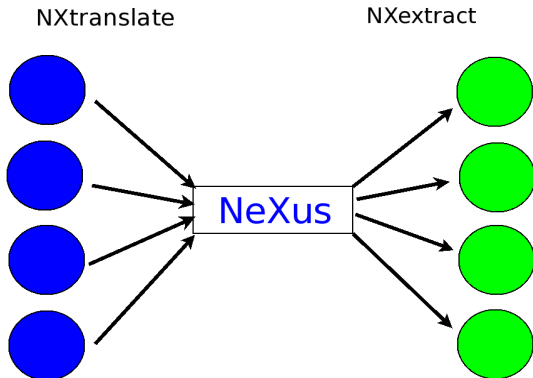
- Written in NeXus Definon Language, NXDL
- Component definitions: dictionaries of allowed field names for the various NeXus groups
- **APPLICATION DEFINITIONS**
 - **DEFINE WHAT HAS TO BE IN A NEXUS FILE FOR A CERTAIN APPLICATION**
 - **DEFINES STANDARDS**
 - **ANOTHER VIEW: CONTRACT BETWEEN FILE PRODUCERS AND USERS ABOUT WHAT HAS TO BE IN A NEXUS FILE FOR A WELL DEFINED PURPOSE**
- NXvalidate

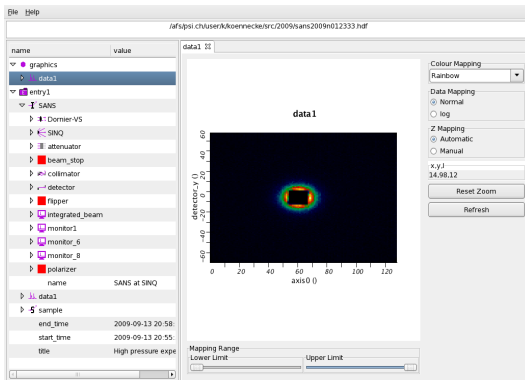
Available NeXus Application Definitions

NXARCHIVE	NXMONOPD	NXREFSCAN
NXREFTOF	NXSAS	NXSCAN
NXTAS	NXTOFRAW	NXTOMO
NXTOMOPHASE	NXXEULER	NXXKAPPA
NXXNB	NXXROT	NXIQPROC
NXTOMOPROC	NXGISAS	

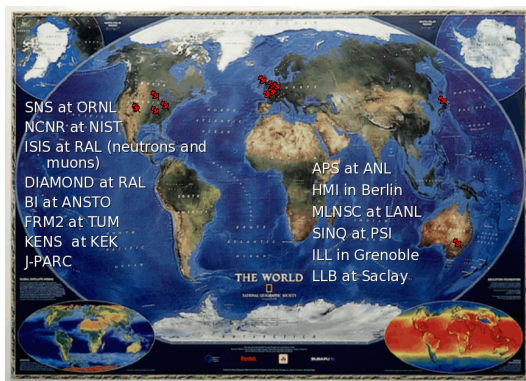
- 1 Construct an application definition with advice from the NIAC
 - 2 Cure for a year; data should be produced in the new format in this time
 - 3 After curation and review: this is the standard for this application type.
- No promises, but the NIAC may do it for you
 - Description of experiment
 - Minimum set of data items necessary form common use
 - Example data

- nxbrowse, nxtree, nxtranslate, nxvalid, nxextract, nxplot
- DANSE
- DAVE
- FABLE (ESRF)
- ISAW
- LAMP
- openGenie
- ICAT
- Mantid
- numerous smaller utilities





Who commits to NeXus?



- The NeXus data format is capable to store data for a wide variety of instruments efficiently
- Any compliant NeXus file is discoverable
- Strict standards can be expressed and validated through NeXus and its tools
- More uses: archiving, workflow, without breaking standard compliance
- You are invited to join the NIAC and contribute to NeXus
- More information: <http://www.nexusformat.org>