PAUL SCHERRER INSTITUT



Benjamin Watts :: PolLux Beamline Scientist :: Paul Scherrer Institut Chair of NeXus International Advisory Committee

HDF5 and the NeXus Data Format





- HDF5 provides tools for recording data
- Many arbitrary choices





- HDF5 provides tools for recording data
- Many arbitrary choices
 - What to include?





- HDF5 provides tools for recording data
- Many arbitrary choices
 - What to include?
 - Which data type?





- HDF5 provides tools for recording data
- Many arbitrary choices
 - What to include?
 - Which data type?
 - How to name it?





- HDF5 provides tools for recording data
- Many arbitrary choices
 - What to include?
 - Which data type?
 - How to name it?
 - Where to put it?



Data Format vs File Format





Data Format vs File Format







- Upper Ontology
 - Base Classes



- Domain-Specific Ontologies
 - Application Definitions
- Organisational Framework
 - Formal Processes
 - Proposing additions and changes
 - Democratic representation
 - Community Support



NeXus base class definitions define the set of terms that *might* be used in an instance of that class.

Consider the base classes as a set of components that are used to construct a data file.

NXaperture

A beamline aperture.

NXattenuator

A device that reduces the intensity of a beam by attenuation.

NXbeam

Properties of the neutron or X-ray beam at a given location.

NXbeam_stop

A device that blocks the beam completely, usually to protect a detector.

NXbending_magnet

A bending magnet

NXcapillary

A capillary lens to focus the X-ray beam.

NXcite

A literature reference

NXcollection

An unvalidated set of terms, such as the description of a beam line.

NXcollimator

A beamline collimator.

NXcrystal

A crystal monochromator or analyzer.

NXcylindrical_geometry

Geometry description for cylindrical shapes.

NXdata

NXdata describes the plottable data and related dimension scales.

NXdetector

A detector, detector bank, or multidetector.



Application Definitions

NeXus application definitions define the minimum set of terms that must be used in an instance of that class.

Application definitions also may define terms that are optional in the NeXus data file.

NXarchive

This is a definition for data to be archived by ICAT.

NXarpes

This is an application definition for angular resolved photo electron spectroscopy.

NXcanSAS

Implementation of the canSAS standard to store reduced small-angle scattering data of any dimension.

NXdirecttof

This is a application definition for raw data from a direct geometry TOF spectrometer

NXfluo

This is an application definition for raw data from an X-ray fluorescence experiment

NXmonopd

Monochromatic Neutron and X-Ray Powder diffractometer

NXmx

Functional application definition for macromolecular crystallography

NXsas

Raw, monochromatic 2-D SAS data with an area detector

NXstxm

Application definition for a Scanning Transmission X-ray Microscope.

NXtomo

This is the application definition for x-ray or neutron tomography raw data.

NXtomophase

This is the application definition for x-ray or neutron tomography raw data with phase contrast variation at each point.



An Example NeXus Structure

- Object classes are added via attributes
- Group names are more flexible to enable human readability
- Application definition provides required features, but describing the entire experimental apparatus is always encouraged

Root I	evel of a NeXus data file				
NXent	try required				
	All data belonging to one scan or run.				
	A given NeXus file can contain multiple related scans or runs				
	NXinstrument				
	The data needed to describe an instrument.				
	Contains groups for each relevant instrument component.				
	NXsource				
	NXcollimator				
	NXattenuator				
	NXdetector				
	NXsample				
	All the information about the sample				
	NXmonitor				
	Incident intensity monitor				
	NXuser				
	User information				
	NXdata required				
	Links to plottable data in the NXdetector group				
	one instance for each detector bank. Provides support				
	for generating a view of the data automatically.				

PAUL SCHERRER INSTITUT



An Example NeXus Structure

File Help			root of HDF5 file	
 82k_SC_as_Stack.hdf5 	entry1/counter0		Root level	of a NeXus data file
 entry1 collection control 	GroupID:144115188075855 addr:1359768 fileno:1 refCn Attributes:6 NX_class NXdata	9993 it:1	NXentry All A g	data belonging to on iven NeXus file can c
 counter0 definition end time 	energy_indices [0] sample_y_indices [1] sample_x_indices [2] signal data		NX	i nstrument The data needeo Contains groups
energy	axes ['energy' 'sample	_y' 'sample_x']		NXsource
 instrument 				NXcollimator
aperture_1				NXattenuator
aperture_2				NXdetector
bendmagnet				indeceetor
control				
 energy 				All the informati
mirror			NX	monitor
Image: Content of the selecting of th				Incident intensit
sample_x			NX	user
sample_y				User information
source			NX	data
time_detector				Links to plottabl
 zone_plate sample 				one instance for
start time				ior generating a

PAUL SCHERRER INSTITUT

An Example NeXus Structure

ile Help		root of HDF5 file
 82k_SC_as_Stack.hdf5 entry1 collection control counter0 	entry1/counter0 GroupID:144115188075855993 addr:1359768 fileno:1 refCnt:1 Attributes:6 NX_class NXdata energy_indices [0] sample_y_indices [1]	Root level of a NeXus data file NXentry All data belonging to or A given NeXus file can NXinstrument
 definition end_time energy 	sample_x_indices [2] signal data axes ['energy' 'sample_y' 'sample_x']	Contains groups
 instrument aperture_1 aperture_2 bendmagnet control counter0 energy mirror order_selecting_aperture sample_x sample_y source time_detector zone_plate sample start_time 	File Help File Help S2k_SC_as_Stack.hdf5 File Help S2k_SC_as_Stack.hdf5 File Help Collection Collection Counter0 Counter0 Counter0 Count_time Count_time Count_time Count_time Sample_x Sample_y Stxm_scan_type Collection Counter0 Counter0 Count_time	entry1/counter0/data DatasetID:360287970189640240 addr:1360552 fileno:1 refCnt:1 Attributes:1 units shape: (3, 200, 200) type: float64 chunk:(1, 1, 200) fill time:2 alloc_time:3 layout:2 nfilters:2







Why should I follow your rules?

Why give up my freedom?

Why reinvent the wheel?

What's in it for me?

Efficiency means less work in the long term

Why tie myself down?

You only commit to your own community through your AD

Write an AD and make them *your* own rules!

Why fix something that's not broken?

Because you'll have little time when it breaks

Working with others is a hassle!

Interoperability shares the work!



- NeXus International Advisory Committe (NIAC)
 - Members recommended by stakeholder institutions.
 - Mostly from X-ray and neutron facilities.
 - Other types of stakeholders are accepted.
- NIAC activities:
 - Proposed changes and additions discussed and voted on.
 - Full meeting every second year (following NOBUGS conference).
 - Smaller "code camp" meeting in between years.
 - Monthly teleconferences to discuss issues as they arise.
- Mailing lists for general support.



NeXus International Advisory Committee

Benjamin Watts, Paul Scherrer Institute, Swiss Light Source, Switzerland (Chair) **Mark Basham**, Diamond Light Source, UK (Executive Secretary) **Pete Jemian**, Advanced Photon Source, USA (Documentation Release Manager) **Stuart Campbell**, Brookhaven National Laboratory/NSLS-II, USA (Technical Manager) Herbert Bernstein, CIF (non-facility member) Aaron Brewster, Lawrence Berkeley Laboratory, USA Bjorn Clausen, Los Alamos National Laboratory, USA **Stephen Cottrell**, Rutherford Appleton Laboratory, UK (Muon Representative) Ricardo Ferraz-Leal, SNS and HFIR at ORNL, USA Jens-Uwe Hoffmann, Helmholtz Zentrum Berlin, Germany Mark Könnecke, Paul Scherrer Institute, Swiss Spallation Source, Switzerland **Raymond Osborn**, Argonne National Laboratory, USA (non-facility member) **Tobias Richter**, European Spallation Source, Sweden Armando Sole, European Synchrotron Radiation Facility, France Jiro Suzuki, KEK, Japan Takahiro Matsumoto, Spring8, Japan



Appendix: Multi-Technique Experiments

- An extra level of hierarchy with Nxsubentry
- Multiple application definitions can be included in parallel

root of HDF5 file					
	NXent	NXentry			
		All data			
		NXinst			
			NXsource		
		NXsample NXuser			
		NXdata	required		
	Links to data representing a default view				
	NXsubentry				
	e.g. small-angle scattering application definition				
	NXsubentry				
	wide-angle scattering application definition				



Appendix: Processed Data

Both raw and processed data can be stored, even together in the same file

root of HDF5 file						
N		required				
		All dat	a belonging to this processed data entry			
	NXprocess					
			Data needed to describe this processing step			
			input:NXparameter			
ou			output:NXparameter			
	NXsample					
	NXdata					
		Links to data representing a default view				