State Of The NeXus Data Format

Mark Könnecke

NeXus International Advisory Committee

October 12, 2010

NeXus Mission

- Solve the predicament of the travelling scientists
- Definition of a standard data format
  - Rules
  - Validation tools
- Promotion of NeXus
  - Documentation
  - NeXus API
  - Outreach to the scientific community

PAUL SCHERRER INSTITUT

NeXus Design

- Complete data for typical use
- Extendable, add additional data as you please
- Self describing
- Easy automatic plotting
- Platform independent, public domain, efficient
- Suitable for a wild variety of applications

PAUL SCHERRER INSTITUT

NeXus Levels

1. Physical file format and API for accessing files: HDF-4, HDF-5, XML
2. Rules for storing data in files
3. Component and application definitions
4. NeXus Utilities

PAUL SCHERRER INSTITUT

NeXus API

- NeXus-API hides complex HDF API
- Transparent access to all three supported physical file formats
- ANSI-C implementation
- Bindings: F77, Java, SWIG
- NEW: first class bindings to C++, python, IDL
- NEW: experimental python tree API added

PAUL SCHERRER INSTITUT

PSI

NeXus Objects

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

PAUL SCHERRER INSTITUT

PSI

entry:NXentry
- sample:NXsample
- instrument:NXinstrument
  - source:NXsource
  - velocity_selector:NXvelocity_selector
  - detector:NXdetector
    - data[xsize,ysize], signal=1 (1)
- control:NXmonitor
  - data
- data:NXdata
  - link to (1)

PAUL SCHERRER INSTITUT

NEW: NeXus Processed Data File Structure

entry:NXentry
- sample:NXsample
- processing_name:NXprocess
  - program
  - version
  - parameters:NXparameter
    - raw_file
- data:NXdata
  - data[nx,ny,nz], signal=1

PAUL SCHERRER INSTITUT

NEW: NXsubentry

entry:NXentry
- sample:NXsample
- instrument:NXinstrument
- .....
- sas:NXsubentry
  - sample:NXsample
  - instrument:NXinstrument
    - source:NXsource
    - velocity_selector:NXvelocity_selector
    - detector:NXdetector
      - data[xsize,ysize], signal=1 (1)
  - control:NXmonitor
    - data
  - data:NXdata
    - link to (1)

PAUL SCHERRER INSTITUT PSI

NEW: NXcollection

entry,NXentry
- measurement:NXcollection
  - positions:NXcollection
    - om
    - two_theta
  - scalars:NXcollection
    - title
    - wavelength
  - data:NXdata
    - detector1
    - mca5

PAUL SCHERRER INSTITUT

Why the Hierarchy??

- Supports self description and allows short names in components

PAUL SCHERRER INSTITUT

Why the Hierarchy??

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

PAUL SCHERRER INSTITUT

PSI

Why the Hierarchy??

- 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

PAUL SCHERRER INSTITUT

Why the Hierarchy??

- 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

PAUL SCHERRER INSTITUT

Why the Hierarchy??

- 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
- Take care once when writing, use n times

PAUL SCHERRER INSTITUT

PSI

Storing Single Data Items

- Units have to specified
- Locating axis
- Store data as physical value and in C storage order
- NEW: Taking care of scaled data and other storage orders

PAUL SCHERRER INSTITUT

PSI

- McStas Coordinate System
- Angle based polar coordinate system
- NEW: full mapping imageCIF - NeXus now possible
- NEW: General axis and transformations

PAUL SCHERRER INSTITUT

NeXus Component and Application Definitions

- 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
  - Validation by NXvalidate
- Written in NeXus Definition Language, NXDL

PAUL SCHERRER INSTITUT

NEW: Available NeXus Application Definitions

- NXarchive
- NXreftof
- NXtas
- NXtomophase
- NXxnb
- NXtomoproc
- NXindirectof
- NXsastof
- NXtofsingle
- NXmonopd
- NXsas
- NXtofraw
- NXxeuler
- NXxrot
- NXtofsingle
- NXiqproc
- NXsqom
- NXxas
- NXrefscan
- NXscan
- NXtomo
- NXxkappa
- NXiqproc
- NXdirectof
- NXlauetof
- NXtofraw
- NXxasproc

PAUL SCHERRER INSTITUT

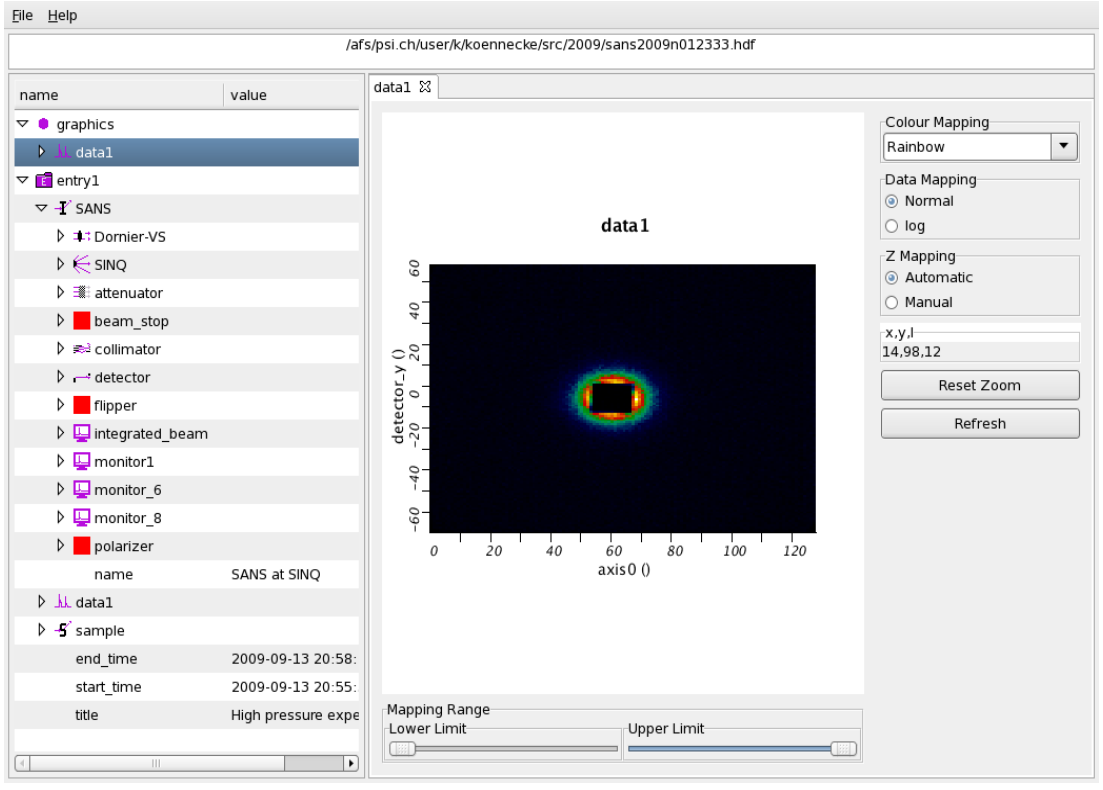
NeXus Tools

nxbrowse	CLI NeXus browser
nxtree	prints NeXus tree
NXmeta	dumps all NeXus meta data
nxtranslate	transforms into NeXus
nxextract	converts from NeXus to ASCII and binary
nxvalidate	NEW: validates NeXus files
nxplot	NEW: plots any NeXus file

PAUL SCHERRER INSTITUT

PSI

NXplot


File Help
/afs/psi.ch/user/k/koennecke/src/2009/sans2009n012333.hdf
name
value
graphics
data1
entry1
SANS
Dornier-VS
SINQ
attenuator
beam_stop
collimator
detector
flipper
integrated_beam
monitor1
monitor_6
monitor_8
polarizer
name
SANS at SINQ
data1
sample
end_time
2009-09-13 20:58:
start_time
2009-09-13 20:55:
title
High pressure expe
data1
data1
detector_y ()
axis0 ()
Colour Mapping
Rainbow
Data Mapping
Normal
log
Z Mapping
Automatic
Manual
x,y,I
14,98,12
Reset Zoom
Refresh
Mapping Range
Lower Limit
Upper Limit

NeXus Political

- PanData: european initiative for SSO, shared data catalogs and some.
  - Works with NeXus to make the shared data catalog fly
- Synchrotrons hitting the wall with current data formats
  - Two workshops in europe addressing this, at ESRF and PSI
  - Another one upcoming

PAUL SCHERRER INSTITUT

PSI

Benefit 1 By using a discoverable data format like NeXus, XML, HDF-5, people can at least figure out what is in the data file.

PAUL SCHERRER INSTITUT

NeXus Benefits

Benefit 1 By using a discoverable data format like NeXus, XML, HDF-5, people can at least figure out what is in the data file.

Benefit 2 Using predefined names from a dictionary gives meaning to the data in a file.

PAUL SCHERRER INSTITUT

NeXus Benefits

Benefit 1 By using a discoverable data format like NeXus, XML, HDF-5, people can at least figure out what is in the data file.

Benefit 2 Using predefined names from a dictionary gives meaning to the data in a file.

Benefit 3 Using a shared API reduces learning costs and increases application stability.

PAUL SCHERRER INSTITUT

PSI

NeXus Benefits

Benefit 1 By using a discoverable data format like NeXus, XML, HDF-5, people can at least figure out what is in the data file.

Benefit 2 Using predefined names from a dictionary gives meaning to the data in a file.

Benefit 3 Using a shared API reduces learning costs and increases application stability.

Benefit 4 With NeXus, HDF-5 plus professional programming techniques a DA application can read any file which contains the required data.

PAUL SCHERRER INSTITUT

PSI

NeXus Benefits

Benefit 1 By using a discoverable data format like NeXus, XML, HDF-5, people can at least figure out what is in the data file.

Benefit 2 Using predefined names from a dictionary gives meaning to the data in a file.

Benefit 3 Using a shared API reduces learning costs and increases application stability.

Benefit 4 With NeXus, HDF-5 plus professional programming techniques a DA application can read any file which contains the required data.

Benefit 5 Storing as much data as possible increases the likelihood that the needed data is actually on file, even for unforeseen uses.

PAUL SCHERRER INSTITUT

PSI

NeXus Benefits

Benefit 1 By using a discoverable data format like NeXus, XML, HDF-5, people can at least figure out what is in the data file.

Benefit 2 Using predefined names from a dictionary gives meaning to the data in a file.

Benefit 3 Using a shared API reduces learning costs and increases application stability.

Benefit 4 With NeXus, HDF-5 plus professional programming techniques a DA application can read any file which contains the required data.

Benefit 5 Storing as much data as possible increases the likelihood that the needed data is actually on file, even for unforeseen uses.

Benefit 6 Application Definitions

PAUL SCHERRER INSTITUT

PSI

NeXus Conclusion And Outlook

- NeXus is already in use at quite a number of facilities and is understood by a long list of applications
- New systems tend to use NeXus
- No competitor for a general purpose data format
- Planned:
  - Refine application definitions together with communities
  - Release application definitions and NXvalidate
  - Update manuals and the NeXus WWW-site
  - www.nexusformat.org, embarrasingly outdated, download manual

PAUL SCHERRER INSTITUT

PSI