

NIAC 2008 Meeting – Rydges Cronulla Hotel, Sydney, Australia, 30th and 31st October 2008

The 7th meeting of the NeXus International Advisory Committee (NIAC) was held prior to, and as a satellite meeting of, the NOBUGS 2008 conference (<http://www.nobugsconference.org/>)

Contents

Agenda and Attendees.....	2
Thanks	2
Voting and Committee Membership	2
Renewal of Officers.....	3
Future Meeting Format.....	3
General Information	3
NeXus Mission statement	4
Review of Outstanding Actions.....	4
The CIF Format.....	4
Reduced Data.....	5
NXprocess	5
Definitions.....	6
Python API.....	6
NAPI Utility API (NXU functions).....	6
API 4.2 status	7
API 4.3 possible new features.....	7
Documentation	7
XML Schema and the NeXus Definition Language (NXDL).....	8
O-O NeXus.....	8
AOB	9
Next Meeting	9
Appendix 1: NXDL example.....	10

Agenda and Attendees

The meeting agenda is at <http://www.nexusformat.org/NIAC2008>

The following people attended the meeting – a (V) after their name indicates that they are, or represent, a current voting member of the NIAC.

Name	Company/Institute
Peter Peterson (Chair, V)	Spallation Neutron Source, Oak Ridge National Laboratory, USA
Freddie Akeroyd (Co-Secretary, V)	ISIS Facility, Rutherford Appleton Laboratory, UK
Nick Hauser (V)	Australian Nuclear Science and Technology Organisation (ANSTO), Australia
Jens-Uwe Hoffmann (V)	Helmholtz Zentrum Berlin (HMI before), Germany
Mark Koennecke (V)	Paul Scherrer Institut, Switzerland
Wayne Lewis	Australian Synchrotron , Melbourne, Australia
Jiro Suzuki (V)	KEK, Japan
Pete Jemian (V)	Argonne National Laboratory, USA
Paul Lewis (V)	Los Alamos National Laboratory, USA
Matthias Drochner (V)	Juelich, Germany

Thanks

To Nick Hauser (ANSTO) for hosting a productive and enjoyable meeting.

Voting and Committee Membership

The constitution (<http://www.nexusformat.org/NIAC>) does not explicitly mention quorum for meetings or a mechanism for appointing a proxy. There were 9 voting members of the committee at the meeting, which with the current committee size of 16 represented over half the members. It was decided that unanimous votes would be binding as their outcome would be unchanged in going to the larger group, but any close votes would be deferred to a future email vote by the whole committee.

ACTION: clarity quorum and proxy situation for future meetings. Any changes to the constitution require a 2/3 majority of current voting members.

Wayne Lewis from the Australian Synchrotron, Melbourne, attended the meeting as an observer. This facility is currently not represented on the full NIAC.

ACTION: propose addition of Wayne Lewis as the Australian Synchrotron member once a letter of support is received from his facility director.

There was also a discussion of how to keep representatives from facilities that have closed down on the NIAC as they are a source of expertise. This may require a constitution change.

ACTION: investigate above point and make proposal to NIAC

Stuart Campbell has recently moved from Diamond to the SNS and so is no longer the Diamond representative – a new Diamond representative will be nominated soon.

Renewal of Officers

Committee officers are elected every NIAC meeting.

- Peter Peterson was re-elected as chair for a 3rd term (note: this is the last year he can stand for chair as per the constitution)
- Freddie Akeroyd was re-elected as secretary for a 2nd term
- Mark Koennecke was re-elected as technical committee chair for a 2nd term.
- Nick Hauser was re-elected as definition release manager for a 2nd term

VOTE: all were in favour, none against.

Future Meeting Format

The NeXus committee has now become very large with new neutron and x-ray sources coming online. In the future holding annual meetings of the entire committee is likely to become more difficult with both budget and scheduling constraints, so after some discussion it was proposed that:

- The full committee should only meet every other year as a satellite meeting to the NOBUGS conference series
- A smaller meeting of committee officers and relevant additional people would occur in non-NOBUGS years to cover specific matters. These meetings would not make binding decisions, but instead make proposals for future email votes by rest of committee.

ACTION: Propose this change for smaller, shorter meetings to the full committee for voting (2/3 majority required as this is a constitution change)

General Information

Regular monthly NeXus officers meetings, which then became wider technical committee meetings, have been occurring via Video conference since early summer 2008. The minutes of these meetings can be found at http://www.nexusformat.org/Technical_Subcommittee and the discussions were mainly related to:

- Work on XML schema for NeXus, which was presented at NOBUGS 2008
- Documentation of definitions and of the “NeXus manual”
- 4.2 API release and 4.3 new features

These small, regular meetings have been very useful and will continue into 2009.

Mark Koennecke reported on a new EU framework initiative called EDMI which is looking at:

- AAA authentication across European facilities (neutron, synchrotron)
- Distributed file database and file access to runs – searching titles
- Analysis

Mark is heading up the data format group and they have asked for 18 months of a person's time for NeXus related work to fit in with the data archiving and retrieval. They expect an answer in early 2009.

Nick Hauser reported that ANSTO are looking at using the ICAT package to catalogue their data.

Darren Kelly (Webel) gave a presentation on the current status of NeXML – this presentation can be viewed by following the links on <http://lists.nexusformat.org/pipermail/nexus/2008/000311.html>

NeXus Mission statement

It was agreed that NeXus should have a mission statement to clarify the role and future direction of the format. Initial work was started on the wiki and will continue post-meeting – see <http://www.nexusformat.org/Mission> and the related discussion page <http://www.nexusformat.org/Talk:Mission>

ACTION: complete discussions with committee and wider NeXus community

Review of Outstanding Actions

- The anl.gov domain has now gone and all content moved to the new nexusformat.org server at ISIS, UK
- It was unclear whether the neutron news article was ever written – need to check with Andy Götz

The CIF Format

James Hester presented an overview of the history of the CIF format, concentrating on what had made the format a success and what NeXus might be able to learn from this. CIF had initial focus on reduced data for publication and succeeded as it was accessible portable ASCII and had a flexible dictionary mechanism. It was also helped by the scientific community agreeing on terminology and political support from the IUCR, who publish leading journals. Other factors were:

- The COMCIFS committee was small in size and able to get on with its job
- There was a checkcif program, which checked molecule parameters in a CIF file prior to publication
- The shelx software ecosystem came on board

The CIF jewel, however, is its comprehensive dictionary.

The “checkcif” program could be considered as a “killer app” – would the equivalent for NeXus be schema validation of files against definitions?

Reduced Data

Pete Jemian presented an overview of the CANSAS 1D XML format. This XML format is extremely similar to the NeXus XML representation (in fact they can be converted via a simple XSLT transformation), but it differs in a few areas:

- It is focused on reduced rather than raw data, so a full instrument description was not needed
- An XML schema, rather than a meta-DTD, is used for validation of data files
- Data is stored in a table like style within the XML file so it can be easily processed by programs such as Microsoft Excel.

Though NeXus has had an “NXprocessed” basic reduced data definition for some time, no work had been done on applying this to any particular domain. Recent API developments have now added an alternative table-like style for XML output and work has also been done on using XML schema for validating NeXus files. It is likely the CANSAS group will use NeXus for 2D data as it provides the option for HDF5 storage of large data sets as well as less verbose vector XML storage format.

One of the main conclusions to come out of the discussion is that NeXus should promote its use for reduced data more and provide appropriate “application definitions”. This still leaves the usual issue of who should do a definition – see later section.

NXprocess

NXprocess describes the set of operations performed on data and is currently just a list of NXnote items, which means details are rather hidden. Analysis programs should be able to record the operation they performed in an NXprocess in a way that would allow it to be reproduced.

It was decided that each data process should result in 1 NXentry, but there may be more than one NXdata present.

The suggestion for NXprocess itself was to add details of the program and an NXparameters class to store program arguments/input data :

```
<NXprocess>
  <program type="NX_CHAR"> program name </program>
  <version type="NX_CHAR"> version </version>
  <date type="ISO8601"> 09-MAR-2008 10:07:30 </date>
  <NXparameters name="">+
    <any_thing>1.0<any_thing>
  </NXparameters>
  <NXnote name=" ">+</NXnote>
</NXprocess>
```

VOTE: new NXprocess: all in favour, but defer to full committee

Definitions

Currently the committee has been responsible for doing most of the instrument definitions work, but the committee does not have either the time or expertise to do all instruments. Definitions from the wider community have not always been forthcoming, so a suggestion from Mark was that:

- first person dealing with a situation defines a format for it
- there is a one year curation period of their definition during which changes can be made; if no problems occur, it is adopted

Mark agreed to produce a tutorial on writing definitions.

At previous NIAC meetings it has been suggested that application based definitions should also be provided and this approach will now be accelerated. Such a definition is based on a “technique” or particular “analysis program” rather than a specific instrument implementations and so is much more generally applicable.

VOTE: for application based definitions – all for

ACTION: NXevent_data be renamed to NXtofevent_data As far as we know, nobody is using the current definition and so it should cause no problems.

ACTION: define a reduced data definition for S(Q,W) – this was mentioned at NOBUGS 2008 and would provide an interface between reduction and analysis for the DANSE project.

ACTION: Remove trailing letter S from NXcharacterisations template

Python API

Paul Kienzle has produced some excellent new python bindings for the NeXus API, which will be included in the 4.2 release.

Paul has also produced a “high level” tree interface to a NeXus file, which creates objects and defines plotting methods for them. Though such high level functionality will be vital in the future, it was decided that it should not be included with the 4.2 release python bindings as committee members had not been able to familiarize themselves with it sufficiently. The ultimate goal would be to use this higher level interface to provide an “Open GENIE” like NeXus aware program.

ACTION: If Paul wishes, the high level tree interface can be included as part of 4.2 contrib area. Otherwise it will be evaluated for possible inclusion in the 4.3 release.

NAPI Utility API (NXU functions)

The Fortran 90 interface had various helper functions that grouped together common operations on the API. These functions were planned to be implemented in C via the napiu.h header, but a few questions on the implementation were outstanding (see <http://trac.nexusformat.org/code/ticket/1>)

The group reviewed napiu.h and concluded the following:

- The globals setting functions are no longer needed
- Readdata/writedata : name histogram->data, data->slab
- Remove set compress as deprecated
- All functions should allocate memory for the caller

API 4.2 status

This is about to go into release candidate status – details are at http://www.nexusformat.org/NeXus_42_Testing

New features in 4.2 include:

- C++ API
- IDL bindings
- Table style XML output support

API 4.3 possible new features

The following were suggested for debate

- A table data (tuple) type
- A NXUgetplotdata() function
- Providing the API as web services – babel, cobra etc.
- Base64 encoding for binary data in XML – possible byte order issues though
- An NXUsetdefinition function
- Nxdiff command

It was agreed to ask via the nobugs/nexus lists for feedback

Documentation

We need two types: online and offline (paper). The NeXus API currently uses doxygen, which can provide both formats. A conversion of wiki to DocBook was tried over summer using a MediaWiki collections extension, but the output still requires substantial editing. After some discussion it was agreed that:

- Docbook should be used as the primary format for the NeXus manual - HTML pages can be generated from this. The DocBook source would be version controlled via subversion.
- The WIKI should be used for discussion/prototyping and items from this moved into DocBook where appropriate

We also need documentation for definitions – this will become easier with XML schema.

XML Schema and the NeXus Definition Language (NXDL)

Having definitions that allow easy machine validation of data files is very important if NeXus files are to be the interface between e.g. reduction and analysis. They aid developers in checking that their programs are working and help find problem when things go wrong. Machine validation removes ambiguity and errors, which is where the old meta-DTDs failed as much validation was done manually – it can also help find ambiguities in the standard and mistakes in definitions as well as in data files.

Freddie Akeroyd started work on using XML schema for validating NeXus files, which was presented at NOBUGS 2008. After discussions at the NIAC meeting it was decided that XML schema were a little too verbose to use directly for defining instrument definitions, so a replacement for the Meta-DTD (called NXDL) would be created that would be transformed into schema via an XSLT stylesheet. Starting with a new definition language gives us the opportunity to correct some of the issues with Meta-DTDs that made them hard to validate such as:

- Dimension information will be split from data type
- Proper units constraints, min/max occurs and enumerations
- Better way to document a definition

Also with NXDL we can have a schema for a definition itself as well as a generated schema for checking a file written against that definition. The NXDL schema, when combined with a schema-aware editor, make writing a definition much easier – it also stops us producing definitions that do not follow our set of rules.

In Freddie's schema work he needed to make use of "unbounded sequences" to allow for the fact that elements can occur in any order in a NeXus XML data file. By using NXDL and a new *nxvalidate* tool we could order elements internally so that we do not need to use unbounded sequences in the schema files.

See appendix 1 for an example of an NXDL definition file.

ACTION: Freddie Akeroyd, Peter Peterson and Pete Jemian to continue work on NXDL and schema

0-0 NeXus

Object oriented features could be applied in two areas – in programs/application and in definitions. The Schema based NeXus already makes use of inheritance to avoid duplication of common items in classes – the question is how many of these "convenience classes" become proper classes that are used by programs and can be created/requested/searched by an API. Also where is knowledge of the object hierarchy stored – the NeXus file itself or externally?

The initial proposal is to use an XML schema to store the object hierarchy, so NeXus files themselves will remain unchanged. An NXU function will be provided to find a class or subclass instance and this will make use of the schema.

ACTION: formulate a proposal

AOB

We need to check that the NX_CHAR data type can hold valid XML when the underlying representation is XML. Cases of XML being stored in NX_CHAR items in HDF files are known and these should not break if nxconvert is used to turn these into an XML nexus file.

Next Meeting

It is planned to meet again in November 2009 – Pete Jemian has offered to hold the meeting at Argonne, USA. Pending a committee vote, this will be a smaller “non-NOBUGS year” meeting.

DRAFT

Appendix 1: NXDL example

NXDL is currently “work in progress”, but this gives an idea of the new layout style.

```
<definition xmlns="http://definition.nexusformat.org/schema/3.1"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://definition.nexusformat.org/schema/3.1 ../nxdl.xsd"
name="NXfermi_chopper" type="group" extends="../nxdl" >

  <doc>Description of a Fermi chopper, possibly with curved slits.</doc>

  - <field name="type" >

    <doc>Fermi chopper type</doc>

  </field>

  - <field name="rotation_speed" type="NX_FLOAT" units="NX_FREQUENCY" >

    <doc>chopper rotation speed</doc>

  </field>

  - <field name="radius" type="NX_FLOAT" units="NX_LENGTH" >

    <doc>radius of chopper</doc>

  </field>

  - <group type="NXgeometry" >

    <doc>geometry of the fermi chopper</doc>

  </group>
```