

Comments about the NeXusQV document

The ESRF position is basically the one outlined by Armando Solé when the draft was sent to the nexus-developers mailing list:

<http://lists.nexusformat.org/pipermail/nexus-developers/2012/000917.html>

That position considers HDF5 as the data format and NeXus a set of conventions for structuring HDF5 files. This already provides answers to many of the questions asked in the document.

1. HDF5 is the data format

The NeXus API is therefore not needed. There are better or standard tools out there.

2. NeXus is a set of conventions

Developments, if any, should be focussed on providing tools to validate the adopted conventions.

In the NeXusQV document, there are several steps outlined in order to adopt NeXus and *“before the full benefits of NeXus can be realized:*

- 1. The community has to agree on a standard.*
- 2. The standard has to be implemented in data acquisition systems.*
- 3. The standard has to be implemented in data analysis software.”*

No wonder that with that way of thinking NeXus has been a failure. As a matter of fact, the ESRF has started the implementation of HDF5 by the data analysis software.

The NIAC has tried to answer:

- a) How to store the data and
- b) How to analyze them

NeXus as it is, answers the question of data archival while, by other hand, archival is a local issue. The ESRF does not attempt to provide any alternative to NeXus in order to fulfil that goal. Nevertheless, archival is the least of the priorities of any beamline scientists in what concern data formats. Their priorities are a) to be able to collect the data and b) to be able to analyze them.

It is our opinion that:

- a) **How to store the data is a question that belongs to the data acquisition** people of the facility and the constraints of the experiment (speed). For archival purposes, one just needs to make sure the data can be analyzed.
- b) **How to analyze the data belongs to the community employing a technique and/or to the data-analysis application developers.** Those two groups are the only ones who can define what is needed to properly carry out the analysis. The NIAC can define a set of required data to perform an analysis, but if the leading applications decide to use other parameters the NIAC has

nothing to do. As a matter of fact there are community initiatives going on for XAS and for tomography. If they arrive to a consensus, the NIAC will have **no other choice** than to accept them as they come.

The ESRF will try to conform to NeXus conventions when writing HDF5 files. We do not intend to reinvent the wheel. If things are properly designed, to add an NX_CLASS attribute to some HDF5 groups is not such a huge effort. Application definitions will be supported as defined by the communities or leading programs of the different techniques. If they just want all the information required for the analysis contained in an HDF5 group, it costs nothing to add an NX_CLASS attribute NXsubentry to it and we are done. However, do not impose the whole NeXus tree as it is the case now and leave it as optional. That “community supplied definition” is nothing else than the dictionary of the common data model (CDM). Our only difference with the CDM approach is that since HDF5 has to be supported by the analysis codes, and HDF5 can deal with external links, we consider the dictionary itself should be a group in an HDF5 file.