BERKELEY LAB

LAWRENCE BERKELEY NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

The Wide World of NeXus

HDRMX Satellite Meeting

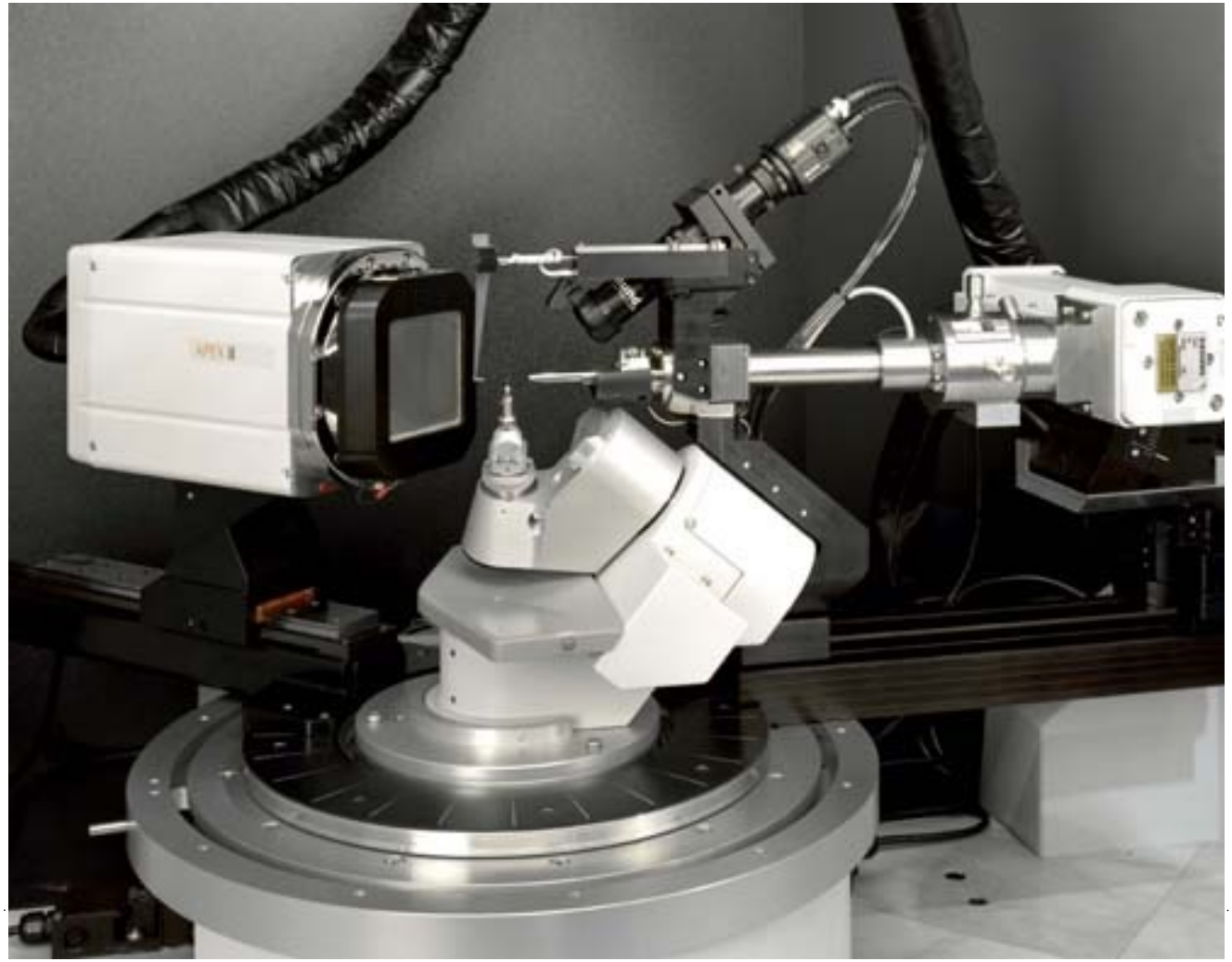
March 26th 2024

Aaron Brewster

What is NeXus ?

- Neutron/X-ray unified standard data format, used in neutron, X-ray, muon, and electron data science
  - https://www.nexusformat.org/
- In use and being adopted by many facilities/communities/detector manufactures
  - SOLEIL and ESRF (France), Diamond and ISIS (UK), PSI, SINQ and SLS (Switzerland), NSLS-II, SNS, APS, Oak Ridge National Laboratory (SNS/HFIR), and Lujan/LANL (USA), KEK and Spring 8 (Japan), DESY and European XFEL (Germany), MAX IV (Sweden), μSR (muon spin rotation/relaxation/resonance) community, Extreme Light Infrastructure (Czech Republic, Hungary and Romania), CLS (Canada), and ALBA (Spain)
  - Dectris: Eiger
- Governed by the NeXus International Advisor Committee (NIAC)

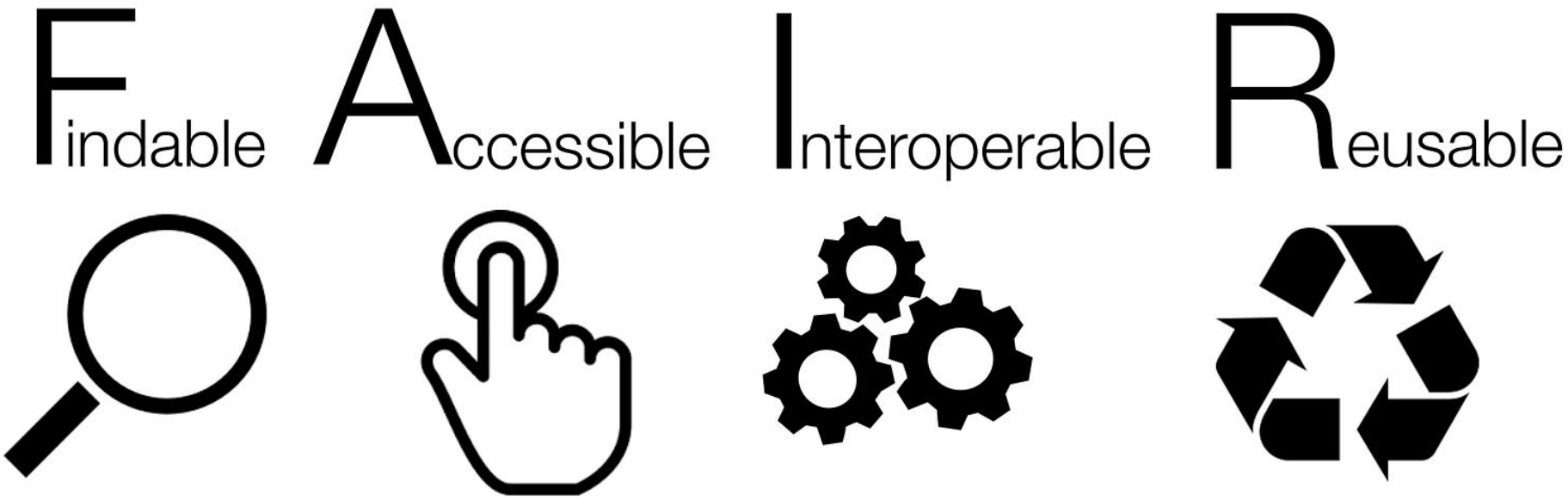
What is NeXus ?



A file format designed to capture the entire experiment

https://serc.carleton.edu/msu_nanotech/methods/SXD.html

What is NeXus ?



Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

What is NeXus ?

- A series of design guidelines
- A set of data objects
  - Base classes: detector, beam, goniometer
  - Application definitions: data recipe
- A community of like-minded developers
  - Docs
  - Software
  - Support

Our NeXus: NXmx

research papers

IUCrJ
ISSN: 2052-2525
BIOLOGY | MEDICINE

Volume 7 | **Part 5** | September 2020 | Pages 784-792
https://doi.org/10.1107/S2052252520008672
OPEN ACCESS
Cited by 15

Gold Standard for macromolecular crystallography diffraction data

Herbert J. Bernstein,[a*] Andreas Förster,[b] Asmit Bhowmick,[c] Aaron S. Brewster,[c] Sandor Brockhauser,[d,e,f] Luca Gelisio,[g] David R. Hall,[h] Filip Leonarski,[i] Valerio Mariani,[g] Gianluca Santoni,[j] Clemens Vonrhein[k] and Graeme Winter[h]

[a]Ronin Institute for Independent Scholarship, c/o NSLS II, Brookhaven National Laboratory, Upton, New York, USA, [b]DECTRIS Ltd, Täfernweg 1, 5405 Baden-Dättwil, Switzerland, [c]Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA, [d]European XFEL GmbH, Holzkoppel 4, 22869 Schenefeld, Germany, [e]Biological Research Centre Szeged (BRC), Temesvári krt. 62, 6726 Szeged, Hungary, [f] University of Szeged, Arpad ter 2, 6720 Szeged, Hungary, [g]Center for Free-Electron Laser Science, Notkestrasse 85, 22607 Hamburg, Germany, [h] Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot OX11 0DE, United Kingdom, [i]Swiss Light Source, Paul Scherrer Institut, Forschungsstrasse 111, 5232 Villigen PSI, Switzerland, [j]Structural Biology Group, European Synchrotron Radiation Facility, 71 Avenue des Martyrs, 38000 Grenoble, France, and [k]Global Phasing Ltd, Sheraton House, Castle Park, Cambridge CB3 0AX, United Kingdom
[*]Correspondence e-mail: yayahjb@gmail.com

Edited by M. Takata, SPring-8, Japan (Received 6 February 2020; accepted 26 June 2020; online 10 July 2020)

Macromolecular crystallography (MX) is the dominant means of determining the three-dimensional structures of biological macromolecules. Over the last few decades, most MX data have been collected at synchrotron beamlines using a large number of different detectors produced by various

BERKELEY LAB

The Gold Standard

Small example Eiger 2X 16M data set from Diamond Light Source I04 revised for HDRMX Gold Standard Discussion

Winter, Graeme[1]; Brewster, Aaron[2]; Bernstein, Herbert J.[3]

Show affiliations

Revised useful small (488 frame) Eiger data set recorded during routine testing, useful for software testing as it is small. Data recorded is from a thaumatin crystal by Graeme Winter, The original dataset is https://zenodo.org/record/3385862 which contains two Eiger

https://doi.org/10.5281/zenodo.3484187

68 image lysozyme dataset recorded on the Jungfrau 16M detector at SwissFEL and formatted as a NeXus file

Brewster, Aaron[1]; Wang, Meitian[2]; Bernstein, Herbert J[3]

Show affiliations

Data provided by Meitian Wang at PSI and master file revised May 2020 for full NXmx compliance.

https://doi.org/10.5281/zenodo.3352357

BERKELEY LAB

Aiming for Gold Standard compliance!

I'd like someone to tell me how to make Nexus images for my geometry

(1) beamline staff being aware that metadata exists

Accurate beamline geometry from a transmission electron microscope

BERKELEY LAB

It's a bit easier now!*

- dxtbx.any2nexus
- Inputs
  - Any old image file
  - DIALS output (.expt)
- Creates a NeXus file!
  - Supports multipanel detectors!
  - Supports goniometers!
  - Supports beam spectra!

**for macro-molecular crystallographers*

BERKELEY LAB

NeXus 101

3.3.1. Base Class Definitions

A description of each NeXus base class definition is given. 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.

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

...

https://manual.nexusformat.org/classes/base_classes/

BERKELEY LAB

NeXus 101

3.3.2. Application Definitions

A description of each NeXus application definition is given. 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. The definition, in this case, reserves the exact term by declaring its spelling and description. Consider an application definition as a *contract* between a data provider (such as the beam line control system) and a data consumer (such as a data analysis program for a scientific technique) that describes the information is certain to be available in a data file.

https://manual.nexusformat.org/classes/applications

Application definitions

NXarchive: This is a definition for data to be archived by ICAT (http://www.icatproject.org/).

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

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

NXiqproc: Application definition for any data.

NXlauetof: This is the application definition for a TOF laue diffractometer

NXmonopd: Monochromatic Neutron and X-Ray Powder diffractometer

NXmx: functional application definition for macromolecular crystallography

NXrefscan: This is an application definition for a monochromatic scanning reflectometer.

NXreftof: This is an application definition for raw data from a TOF reflectometer.

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

NXsastof: raw, 2-D SAS data with an area detector with a time-of-flight source

NXscan: Application definition for a generic scan instrument.

NXspe: NXSPE Inelastic Format. Application definition for NXSPE file format.

NXsqom: This is the application definition for S(Q,OM) processed data.

Application definitions

NXstxm: Application definition for a STXM instrument.

NXtas: This is an application definition for a triple axis spectrometer.

NXtofnpd: This is a application definition for raw data from a TOF neutron powder diffractometer

NXtofraw: This is an application definition for raw data from a generic TOF instrument

NXtofsingle: This is a application definition for raw data from a generic TOF instrument

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.

NXtomoproc: This is an application definition for the final result of a tomography experiment: a 3D construction of some volume of physical properties.

NXxas: This is an application definition for raw data from an X-ray absorption spectroscopy experiment.

NXxasproc: Processed data from XAS. This is energy versus I(incoming)/I(absorbed).

NXxbase: This definition covers the common parts of all monochromatic single crystal raw data application definitions.

NXxeuler: raw data from a four-circle diffractometer with an eulerian cradle, extends NXxbase

NXxkappa: raw data from a kappa geometry (CAD4) single crystal diffractometer, extends NXxbase

NXxlaue: raw data from a single crystal laue camera, extends NXxrot

NXxlaueplate: raw data from a single crystal Laue camera, extends NXxlaue

NXxnb: raw data from a single crystal diffractometer, extends NXxbase

NXxrot: raw data from a rotation camera, extends NXxbase

3.3.3. Contributed Definitions

A description of each NeXus contributed definition is given. NXDL files in the NeXus contributed definitions include propositions from the community for NeXus base classes or application definitions, as well as other NXDL files for long-term archival by NeXus. Consider the contributed definitions as either in *incubation* or a special case not for general use. The NIAC: The NeXus International Advisory Committee is charged to review any new contributed definitions and provide feedback to the authors before ratification and acceptance as either a base class or application definition.

Some contributions are grouped together:

- Optical Spectroscopy
- Multi-dimensional Photoemission Spectroscopy
- Atom Probe Microscopy
- Electron Microscopy
- Transport Measurements
- Geometry and Microstructures

and others are simply listed here:

NXaberration
: Quantified aberration coefficient in an aberration_model.

NXaberration_model
: Models for aberrations of electro-magnetic lenses in electron microscopy.

NXaberration_model_ceos
: CEOS definitions/model for aberrations of electro-magnetic lenses.

...

Sandor Brockhauser, Center for Materials Science Data (HU Berlin), Germany (NIAC Executive Secretary)

The NeXus international advisory committee (NIAC)

Services:

- Manage and develop the NeXus definitions (base classes, application definitions, contributed definitions)
- Document and provide examples
- Provide automated validators
  - cnxvalidate: https://github.com/nexusformat/cnxvalidate.git
  - punx: https://punx.readthedocs.io/

Acknowledgments: the NIAC!

- Sandor Brockhauser, Center for Materials Science Data (HU Berlin), Germany (Executive Secretary)
- Peter Chang, Diamond Light Source, UK (Documentation Release Manager)
- Benjamin Watts, Swiss Light Source, Switzerland
- Mark Könnecke, PSI, Switzerland
- Herbert Bernstein, CIF
- Thomas Caswell, Brookhaven National Laboratory/NSLS-II, USA
- Bjørn Clausen, Los Alamos National Laboratory, USA
- Heike Görzig, Helmholtz Zentrum Berlin, Germany
- Pete Jemian, Advanced Photon Source, USA
- Raymond Osborn, Argonne National Laboratory, USA
- Wout de Nolf, European Synchrotron Radiation Facility, France
- Takahiro Matsumoto, Spring8, Japan
- Balázs Bagó, the Extreme Light Infrastructure, Czech Republic, Hungary and Romania
- Russ Berg, Canadian Light Source, Canada
- Majid Ounsy, Synchrotron Soleil, France
- Chen Zhang, Oak Ridge National Laboratory (SNS/HFIR), USA
- Luca Geliso, European XFEL, Germany
- Paul Millar, DESY, Germany
- Zdenek Matej, MAX IV, Sweden
- Fernan Saiz, ALBA, Spain

NIH/NIGMS grants R24GM154040 and R35-GM151988
DOE/Office of Science contract DE-AC02-05CH11231

BERKELEY LAB