



Why do we need a(nother) new data format? A standardized data format for azimuthally integrated area detector data

Mads Ry Vogel Jørgensen,^{a,b*} Frederik Holm Gjørup,^{a,b} Meghdad Yazdi-Rizi,^b
Zdenek Matej,^b Wout De Nolf^c and Paul Bell^b

Received 26 March 2026

Accepted 21 April 2026

Edited by K. Kvashnina, ESRF – The European
Synchrotron, France

Keywords: data format; PXRD; scattering;
azimuthal integration.

^aDepartment of Chemistry & iNANO, Aarhus University, Langelandsgade 140, Aarhus C, Denmark, ^bMAX IV Laboratory, Lund University, Fotogatan 2, Lund, Sweden, and ^cESRF – The European Synchrotron, 71 Avenue des Martyrs, 38043 Grenoble, France. *Correspondence e-mail: mads@chem.au.dk

Many powder X-ray diffraction, total scattering, and small-angle X-ray scattering experiments at synchrotron facilities yield azimuthally integrated intensity data, yet current practice relies on facility-specific data formats that hinder inter-operability and reuse. We propose a standardized HDF5-based format within the NeXus convention (NXazint), designed to store 1D and 2D integrated data along with essential metadata for direct use in analysis software. This approach supports FAIR principles, simplifies handling of large datasets, and promotes compatibility across beamlines and facilities. The format is already implemented at MAX IV, with visualization and powder X-ray diffraction refinement tools available.

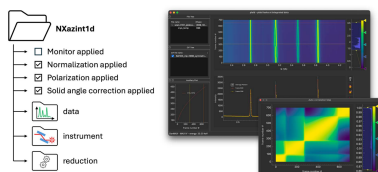
If you carry out a scattering experiment at a synchrotron beamline in the future, you may notice that the data files you bring home have changed. Data formats are seldom very exciting, so why should you appreciate this new one?

The use of synchrotron radiation for powder diffraction and scattering is ubiquitous, and today all large-scale facilities have one or more beamlines designed for these types of experiments. For powder X-ray diffraction (PXRD), total scattering (TS), and small-angle X-ray scattering (SAXS), many beamlines use 2D area detectors. The data from these detectors are azimuthally integrated to yield intensity values as a function of scattering angle and, in some cases, as a function of azimuthal angle (*e.g.* Kieffer *et al.*, 2020; Jensen *et al.*, 2022; Filik *et al.*, 2017; Toby & Von Dreele, 2013). However, while the experimental geometry and data processing are highly similar, today every facility—in many cases even each beamline—produces a distinct file format!

The HDF5 data format (HDF Group, 2026a) is widely used for storing raw and, increasingly, processed X-ray data. The benefit of the HDF5 format is that all data and metadata from a whole experiment can be stored in a single file. The caveat, however, is that the data structures in these files vary greatly from facility to facility and even from beamline to beamline. As no standard exists, analysis programs rarely work with these files. Thus, users must use custom programs or scripts to extract the data in the required format. This leaves the user to convert the ‘beamline data format’ to the format used by their chosen refinement program.

Most refinement programs for PXRD work with ASCII-based files for the intensity data. This has traditionally worked well for low data volumes but quickly becomes impractical for increasing data volumes from, for example, μ XRD mapping (Christensen *et al.*, 2023) or XRD-CT (Leemreize *et al.*, 2013) methods, where hundreds of thousands and even millions of patterns are collected in each experiment. Here, the HDF5 format offers a powerful way to aggregate data in a single file, while also including relevant processing information.

To this end, we at the MAX IV Laboratory (Robert *et al.*, 2023) and ESRF (Raimondi *et al.*, 2023) suggest an open data format (NXazint) that is now being used by all beamlines at MAX IV that perform azimuthal integration of area detector scattering data. The format will in the future also be available for use at ESRF beamlines through the EWOKS data processing workflow (De Nolf *et al.*, 2024). The format is based on



OPEN ACCESS

Published under a CC BY 4.0 licence

HDF5, utilizing the NeXus standard (Könnecke *et al.*, 2015), and has recently been approved as a set of NeXus application classes. The format is integration software agnostic, thus any integration software can be extended to produce these files. A writer for *pyFAI* (Kieffer *et al.*, 2020) has been developed and is included since version 2026.02.

The data format is designed for scattering and diffraction data where an azimuthal integration is performed, *e.g.* PXRD, TS, and SAXS. It is designed to include sufficient information about the data processing to enable the azimuthal integration to be reproduced, provided the raw data are available. ‘Conventional’ 1D intensity data as a function of 2θ (or magnitude of the momentum transfer vector, Q) and azimuthally binned (2D) data for, for example, texture analysis are supported. Either type of data, or both, can be stored in the file. Errors and/or pixel normalization information can optionally be stored for each dataset above. Auxiliary information and measurements, such as beam intensity and/or transmission data for normalization, can be optionally included if desired for the following analysis.

The data format is intended for direct use in analysis programs, such as Rietveld refinement. Thus, the format, for example, also requires that the wavelength be included as a required entry. Support is already available in *GSAS-II* (Toby & Von Dreele, 2013), and support is being implemented in upcoming versions of *TOPAS-Academic* (Coelho, 2018) and *Jana2020* (Petricek *et al.*, 2023). NXazint files can be explored and visualized by many existing HDF5 tools, including web-based tools, in particular *myHDF5* (HDF Group, 2026b). In order to simplify NXazint data visualization even further, we have developed a Python-based viewer, *plaid* (Gjørup, 2026). This desktop application enables rapid inspection of PXRD data, for example, during beam time. It also offers backwards compatibility by exporting to ASCII files.

Currently, the NXazint format is used at multiple beamlines at MAX IV, and we encourage other beamlines/facilities to adopt it so that it can become the standard data format for this type of experiment. In this way, users can directly repurpose scripts and other resources across different beamlines/facilities. A well documented, standardized format will hopefully also encourage more developers to implement support for it in their analysis/refinement programs.

Funders and large facilities alike are all pushing to implement the FAIR data principles, *i.e.* Findable, Accessible, Interoperable and Reusable data (Wilkinson *et al.*, 2016). A standardized universal data format for azimuthally integrated data proposed here will help ensure that the data are interoperable and reusable, thus a step towards full FAIR compliance for processed X-ray scattering data. As such, we hope this (standardized) format will help encourage researchers to make the azimuthally integrated data available when publishing. The specification requires that beamline and facility information be included, thereby increasing the traceability of the data.

The NXazint files are not intended to replace the pdCIF format (Toby, 2005), which is used to report PXRD data and refined structural models. Instead, they are designed as a

```
loop_
  _array_data_external_data.id
  _array_data_external_data.format
  _array_data_external_data.uri
  _array_data_external_data.path
  _array_data_external_data.frame
ext1 HDF5 https://zenodo.org/records/18685805/files/LaB6.h5/entry/data 1
```

Figure 1

Example concept, inspired by imgCIF, of how to link to the first pattern in an NXazint file archived in the Zenodo repository.

convenient format for azimuthally integrated experimental data, which often contain many individual patterns that cannot be practically handled as separate text files. We do, however, recommend that authors link to NXazint files when publishing pdCIF files, as shown in Fig. 1.

The full definition of the format can be found in the official NeXus documentation (NeXus International Advisory Committee, 2026a, 2026b). Example files are available on the Zenodo repository. Python code examples for reading the NXazint files are available on the *plaid* website (Gjørup, 2026).

Example file links:

LaB6.h5: Minimal NXazint1d file example. <https://doi.org/10.5281/zenodo.18685805>.

BTO.h5: NXazint1d file example, including I0 monitor, sample temperature (not calibrated), and time. <https://doi.org/10.5281/zenodo.18685853>.

Magnet.h5: NXazint2d file example, including I0 monitor, and sample rotation around the vertical axis (omega). <https://doi.org/10.5281/zenodo.18685869>.

Magnet_multi.h5: Combined NXazint1d and NXazint2d file example, including I0 monitor and sample rotation around the vertical axis (omega). <https://doi.org/10.5281/zenodo.18685879>.

Magnet_low_res.h5: Small NXazint2d file example, including I0 monitor, and sample rotation around the vertical axis (omega). <https://doi.org/10.5281/zenodo.18685892>.

Pineapple.h5: NXazint1d mapping file example, including I0 monitor, and sample position. <https://doi.org/10.5281/zenodo.18685898>.

Acknowledgements

We would like to thank the NeXus International Advisory Committee, especially Peter Chang (Diamond Light Source), for their constructive feedback during the review of the NXazint application definitions.

Funding information

Research conducted at MAX IV, a Swedish national user facility, is supported by Vetenskapsrådet (Swedish Research

Council, VR) under contract 2018-07152, Vinnova (Swedish Governmental Agency for Innovation Systems) under contract 2018-04969 and Formas under contract 2019-02496. Part of this project was carried out within the DanMAX project. DanMAX is funded by the Nationalt Udvalg for Forskningsinfrastruktur (NUFI) grant No. 4059-00009B.

References

- Christensen, T. E. K., Chua, J. Q. I., Wittig, N. K., Jørgensen, M. R. V., Kantor, I., Thomsen, J. S., Miserez, A. & Birkedal, H. (2023). *IUCrJ* **10**, 288–296.
- Coelho, A. A. (2018). *J. Appl. Cryst.* **51**, 210–218.
- De Nolf, W., Huder, L., Svensson, O., Payno, H., Autran, P.-O., Broche, L., Koumoutsos, G. & Götz, A. (2024). *Synchrotron Radiat. News* **37**(6), 9–15.
- Filik, J., Ashton, A. W., Chang, P. C. Y., Chater, P. A., Day, S. J., Drakopoulos, M., Gerring, M. W., Hart, M. L., Magdysyuk, O. V., Michalik, S., Smith, A., Tang, C. C., Terrill, N. J., Wharmby, M. T. & Wilhelm, H. (2017). *J. Appl. Cryst.* **50**, 959–966.
- Gjørup, F. H. (2026). *plaid – plaid looks at integrated data*, Version 1.0.0, <https://github.com/fgjurup/plaid>.
- HDF Group (2026a). *Hierarchical Data Format*, Version 5, <https://www.hdfgroup.org/solutions/hdf5/>.
- HDF Group (2026b). *myHDF5 service*, <https://myhdf5.hdfgroup.org/>.
- Jensen, A. B., Christensen, T. E. K., Weninger, C. & Birkedal, H. (2022). *J. Synchrotron Rad.* **29**, 1420–1428.
- Kieffer, J., Valls, V., Blanc, N. & Hennig, C. (2020). *J. Synchrotron Rad.* **27**, 558–566.
- Könnecke, M., Akeroyd, F. A., Bernstein, H. J., Brewster, A. S., Campbell, S. I., Clausen, B., Cottrell, S., Hoffmann, J. U., Jemian, P. R., Männicke, D., Osborn, R., Peterson, P. F., Richter, T., Suzuki, J., Watts, B., Wintersberger, E. & Wuttke, J. (2015). *J. Appl. Cryst.* **48**, 301–305.
- Leemreize, H., Almer, J. D., Stock, S. R. & Birkedal, H. (2013). *J. R. Soc. Interface.* **10**, 20130319.
- NeXus International Advisory Committee (2026a). *NeXus documentation v. 2026.01, NXazint2d*, <https://manual.nexusformat.org/classes/applications/NXazint2d.html>.
- NeXus International Advisory Committee (2026b). *NeXus documentation v. 2026.01, NXazint1d*, <https://manual.nexusformat.org/classes/applications/NXazint1d.html>.
- Petricek, V., Palatinus, L., Plasil, J. & Dusek, M. (2023). *Z. Krist.* **238**, 271–282.
- Raimondi, P., Benabderrahmane, C., Berkvens, P., Biasci, J. C., Borowiec, P., Bouteille, J. F., Brochard, T., Brookes, N. B., Carmignani, N., Carver, L. R., Chaize, J. M., Chavanne, J., Checchia, S., Chushkin, Y., Cianciosi, F., Di Michiel, M., Dimper, R., D’Elia, A., Einfeld, D., Ewald, F., Farvacque, L., Goirand, L., Hardy, L., Jacob, J., Jolly, L., Krisch, M., Le Bec, G., Leconte, I., Liuzzo, S. M., Maccarrone, C., Marchial, T., Martin, D., Mezouar, M., Nevo, C., Perron, T., Plouviez, E., Reichert, H., Renaud, P., Revol, J. L., Roche, B., Scheidt, K. B., Serriere, V., Sette, F., Susini, J., Torino, L., Versteegen, R., White, S. & Zontone, F. (2023). *Commun. Phys.* **6**, 82.
- Robert, A., Cerenius, Y., Tavares, P. F., Hultin Stigenberg, A., Karis, O., Lloyd Whelan, A., Runéus, C. & Thunnissen, M. (2023). *Eur. Phys. J. Plus* **138**, 495.
- Toby, B. H. (2005). *International Tables for Crystallography Volume G: Definition and exchange of crystallographic data* edited by S. R. Hall & B. McMahon, pp. 258–269. Dordrecht: Springer Netherlands.
- Toby, B. H. & Von Dreele, R. B. (2013). *J. Appl. Cryst.* **46**, 544–549.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., ’t Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. (2016). *Sci Data* **3**, 160018.