1st Conference on Research Data Infrastructure

Poster presentations II

https://doi.org/10.52825/CoRDI.v1i.258

Published: 07 Sept. 2023

Metadata Fields and Quality Criteria - XAS Reference Database under DAPHNE4NFDI

Abhijeet Gaur[1]*, Sebastian Paripsa[3], Frank Förste[4], Dmitry Doronkin[1,2], Wolfgang Malzer[4], Christopher Schlesiger[4], Birgit Kanngießer[4], Dirk Lützenkirchen-Hecht[3], Edmund Welter[5], and Jan-Dierk Grunwaldt[1,2]

[1] Institute for Chemical Technology and Polymer Chemistry, Karlsruhe Institute of Technology (KIT), Engesserstr. 20, Karlsruhe, D-76131 (Germany)

[2] Institute of Catalysis Research and Technology, Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, D-76344 (Germany)

[3] Fk. 4, Physik, Bergische Universität Wuppertal, Gauß-Str. 20, Wuppertal, D-42097 (Germany)

[4] Technische Universität Berlin, Hardenbergstr. 36, Berlin, D-10623

[5] Deutsches Elektronen-Synchrotron (DESY), Notkestraße 85, Hamburg, D-22607

Abstract. X-ray absorption spectroscopy is important to analyse solid materials, in particular amorphous materials, disordered or multicomponent materials. Due to its vast application in diverse scientific fields XAS has become an essential tool for studying, e.g., catalytic reactions or battery materials to mention just a few. In the field of XAS, data are often evaluated by comparing them to previously measured or calculated reference spectra. This sets the high requirements concerning both spectral quality and documentation of the measurements. Under DAPHNE4NFDI, we have been working on to set up a XAS reference database including raw and processed data with an interface developed for uploading and evaluating the data. In this context, defining metadata fields about the performed XAS experiments and documenting this information along with data is essential to make the measured data reusable by any researcher in a similar field and beyond. Another important aspect of a curated database is that users should be able to easily judge the quality and the usability of each data set by looking at the mentioned quality criteria. In the present work, we have discussed and highlighted the importance of metadata fields and quality criteria for the data to be uploaded at the XAS database.

Keywords: X-ray absorption spectroscopy, Database, Metadata, Quality criteria

1. Extended Abstract

X-ray absorption spectroscopy is important to analyse solid materials, in particular amorphous materials, disordered or multicomponent materials. Due to its vast application in diverse fields XAS has become an essential tool for studying, e.g., catalytic reactions or battery materials to mention just a few. In the field of XAS, data are often evaluated by comparing them to previously measured or calculated reference spectra [1]. This sets the high requirements concerning both spectral quality and documentation of the measurements. Previous databases for XAS suffer from problems such as not providing detailed information about the sample and the data acquisition process itself, unknown and inconsistent data formats, difficulties in adding new data to the database, non-standardized organisation of the database, and a review process to ensure and assess the quality of submitted data [2]. Recently, XAFS databases in Japan have

been integrated to create a new database called MDR XAFS database [3], which hosts approximately spectra of 2000 samples and 700 unique materials with machine-readable metadata. However, data quality is not included in the inclusion criteria for this collection and hence use of data is at the discretion of the user.

Under DAPHNE4NFDI [4], we have been working on to set up a XAS reference database including raw and processed data with an interface developed for uploading and evaluating the data. In this context, defining metadata fields about the performed XAS experiment and documenting this information along with data is essential to make the measured data reusable by any researcher in a similar field and beyond. For the present database, we have categorized meta data fields under “Sample”, “Spectra”, “Instrument” and “Bibliography”, subfields under these categories are shown in Scheme 1. Hence, the metadata fields include contributions from users as well as experimental facilities [5]. Another important aspect of a curated database is that users should be able to easily judge the quality and the usability of each data set by looking at the mentioned quality criteria. Quality criteria, i.e., edge step, energy resolution, signal to noise ratio, have been considered for the automated check of any uploaded data. The edge step is directly related to the elemental composition of the sample, the concentration of elements, as well as sample preparation. The energy resolution is generally dependent on the limits of the used beamline/instrument. The assigned values of quality criteria shown in Scheme 2 corresponds to references, e.g., metal foils. During the upload of the data, an automatic check of these defined quality criteria and noise estimation is performed. Details about data upload and quality check procedures are available at our database public webpage [6].

Scheme 1. Defined metadata fields for the XAS database

Sample

- Collection code
- CAS no. (optional)
- Physical state (crystalline, powder, thin film, liquid, gas)
- Structural parameters for crystalline sample
- Crystal orientation
- X-ray or neutron diffractogram (if available)
- Temperature
- Pressure
- Remarks about sample preparation – Foil, Pellet, Capillary, Powder on tape, etc.
- Sample environment - Cell/Microreactor/Batch, gases, solvents, potential, etc.
- General remarks e.g., sample properties (hygroscopic metastable, etc).

Bibliography

- DOI
- Title
- Author
- Reference
- Funding agency /Grant details

Spectra

- Raw data file (optional)
- Absorbance spectrum (3000 data points)
- Reference spectrum
- Data/File format
- Header information

Instrument

- Facility (synchrotron/ lab instrument)
- Beamline
- Acquisition mode
- Crystals
- Mirrors
- Harmonic rejection
- Detectors
- Element
- Edge
- Maximum k range (EXAFS)

The metadata fields and quality criteria are still an open point of discussion to cover the different types of experiments (ex situ, in situ, operando), acquisition modes, instruments (synchrotron beamline/laboratory facility), detection modes, data formats etc. The database will be filled with real spectra covering a wide range of functional materials, with criteria for meta-data and quality assessment. It will allow researchers to easily access and analyze XAS data. The interface for data submission will make it easy for users to contribute their own data

to the database, and the automated assessment of data quality will ensure that the data in the database is of high quality. This will not only benefit the researchers who contribute data, but also other researchers who use the database for their own work.

Scheme 2. Formulated quality criteria for uploaded data (values corresponds to metal foils)

Quality criteria

- edge step [0.5 - 2.0]
- maximum k-range [15 - 20 $Å^{-1}$]
- energy resolution [0.5 - 2.0 eV]
- signal to noise ratio [comparison procedure]
- amplitude reduction factor [0.7 - 1.0]*

*analysed data

One of the application of such curated database is that it would be possible to compare the data for identical samples from different facilities and hence the effect of different parameters of an instrument on the data quality can be studied. As an example Fig. 1 shows the comparison of Pt foils measured at XAS beamlines from different synchrotron facilities. However, these Pt foils may not be identical, e.g., thickness, purity etc., hence their data is also affected by these factors other than beamline parameters. Thus, identical samples need to be measured at different XAS facilties (synchrotron/laboratory) by using either their own or standardized analytical protocols. This is basic idea behind a recently initiated Round Robin test [7] which could further help to standardize the meta data fields across different laboratories.

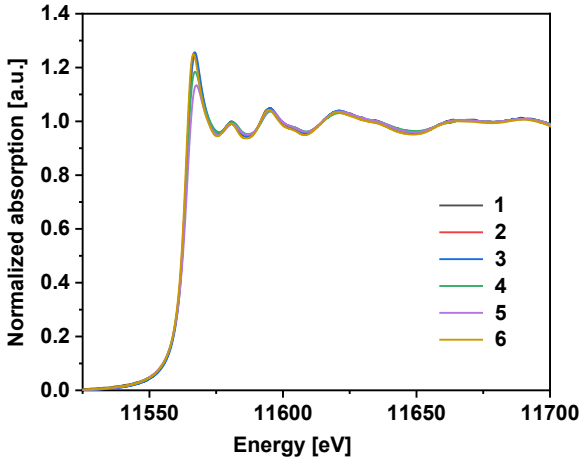

Figure 1. Comparison of Pt foils (distinct) data measured at different beamlines.

As such, our objective is to establish a database for X-ray Absorption Spectroscopy (XAS) data in a timely manner to create a self-accelerating effect and extend the knowledge to other fields. Our overarching goal under DAPHNE4NFDI is to develop and endorse efficient and systematic data and metadata capturing tools during experiments, establish appropriate metadata schemata for respective communities, create federated data catalogues and repositories, and provide tools for data processing, visualisation, and analysis.

Underlying and related material

--

Author contributions

Please include a statement on authors' contributions according to the CreDIT guidelines here. CRediT (Contributor Roles Taxonomy)'s intention is to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Competing interests

The authors declare that they have no competing interests.

Funding

Deutsche Forschungsgemeinschaft http://dx.doi.org/10.13039/501100001659 460248799

Acknowledgement

This work was supported by the consortium DAPHNE4NFDI (https://www.daphne4nfdi.de/) in the context of the work of the NFDI e.V.

References

1. A. Gaur and B. D. Shrivastava, "Speciation using X-ray absorption fine structure (XAFS)" Rev. J. Chem., 5, pp. 361-398, November, 2015, doi: https://doi.org/10.1134/S2079978015040032
2. K. Asakura, H. Abe., M. Kimura, "The challenge of constructing an international XAFS database" *J. Synchrotron Rad.,* 25, pp. 967-971, July, 2018, doi: https://doi.org/10.1107/S1600577518006963
3. M. Ishii, K. Tanabe, A. Matsuda, H. Ofuchi, T. Matsumoto, T. Yaji, Y. Inada, H. Nitani, M. Kimura and K. Asakura, "Integration of X-ray absorption fine structure databases for data-driven materials science" *Sci. Technol. Adv. Mater.: Methods*, 3, pp. 2197518, doi: https://doi.org/10.1080/27660400.2023.2197518
4. DAPHNE4NFDI - Consortium Proposal https://doi.org/10.5281/zenodo.8040606
5. R. Dimper, A. Götz, A. De Maria, M. Solé V.A., Chaillet, B. Lebayle, "ESRF Data Policy, Storage, and Services", Synchrotron Radiation News, 32, pp.7-12, May, 2019, doi: https://doi.org/10.1080/08940886.2019.1608119
6. https://san-wierpa.github.io/xafsdb4daphne/
7. C. T. Chantler, B. A. Bunker, H. Abe, M. Kimura, M. Newville, E. Welter, "A call for a round robin study of XAFS stability and platform dependence at synchrotron beamlines on well defined samples", *J. Synchrotron Rad.*, 25, pp. 935-943, July, 2018, doi: https://doi.org/10.1107/S1600577518003752