

# Efforts toward Data Format Standardization for X-ray Absorption Spectra

Matthew Newville<sup>a</sup>, V. Armando Solé<sup>b</sup>, James R. Hester<sup>c</sup>, and Bruce Ravel<sup>d</sup>

<sup>a</sup>Center for Advanced Radiation Studies, University of Chicago, USA,

<sup>b</sup>European Synchrotron Radiation Facility, Grenoble, France,

<sup>c</sup>Bragg Institute, Australian Nuclear Science and Technology Organization, Kirrawee, Australia,

<sup>d</sup>National Institute of Standards and Technology, Gaithersburg, USA

## Abstract

In order to address the need for easier exchange and archiving of X-ray absorption data, we propose a data format for spectra of  $\mu(E)$  using a simple plaintext file, and invite comments and discussion. The XAS Data Interchange (XDI) format uses space-delimited columns for numerical absorption data, and so is similar to that currently used at many facilities, but with a well-defined structure for metadata in the file header. Several example XDI files and programming libraries for reading these files are shown, and made publicly available. Because this data format can accommodate only single spectra, we also discuss methods for combining multiple spectra into spectral libraries using either hierarchical data format (HDF5) or using relational databases based either on the CIF format or SQL-based database engine.

## X-ray Data Interchange Format

Motivation here : NEEDS WORK

Storing and Exchanging XAFS Data is a common need for everyone using XAFS. In particular, retrieving data on "Standards" or "Model Compounds" is a continuing need for both XANES and EXAFS analysis. Additionally, data taken on samples at different facilities and beamlines need to be compared and analyzed together. At this point, there is no commonly accepted data format for XAFS data. There have been a few attempts to standardize the traditional "ASCII Column File". To be sure, ASCII Column Files have some strong appeal. Most data collection software saves such files, and most processing and analysis software use some variation of this format. In addition, such files may be easily read by humans and used in a wide variety of third party applications. Still, most ASCII Column Files need some intimate knowledge of the data layout to use the data. The lack of a standard ASCII format is a serious problem. The work here is parallel to the efforts to come up with a standard. and will be able to convert data into such standard file formats.

$\mu(E)$  is the basic spectra

More text

## Example XDI File

```
#XDI/1.0 ← Version Info
#Beamline.name: APS 10ID
#Mono.name: Si 111
#Mono.d.spacing: 3.13553
#Element.symbol: Fe
#Element.edge: K
#Column.1: energy eV ← Column Labels, Units
#Column.2: mutrans
#Column.3: i0
### ← start multi-line comment
#Fe K-edge, Lepidocrocite powder
#on 4 layers of tape
#----- ← End Header, Begin Arrays Table
#
6899.9609 -1.3070486 149013.70
6900.1421 -1.3006104 144864.70
6900.5449 -1.3033816 132978.70
6900.9678 -1.3059724 125444.70
...
```

## XDI File description

Required metadata (case insensitive):

- ▶ **Element.symbol** Atomic symbol of absorbing element
- ▶ **Element.edge** K, L3, M4, etc
- ▶ **Mono.d spacing** Strongly encouraged for all data, required if mono angle is given and mono energy is not.

Supported column labels. These labels for Columns should be used, and may be required for certain data sets.

- ▶ **energy** Monochromator energy, in eV or keV.
- ▶ **angle** Monochromator angle, in degrees.
- ▶ **i0**
- ▶ **time**
- ▶ **itrans**
- ▶ **ifluor**
- ▶ **irefer**
- ▶ **mutrans**
- ▶ **mufluor**
- ▶ **murefer**
- ▶ **normtrans**
- ▶ **normfluor**
- ▶ **normrefer**
- ▶ **k**
- ▶ **chi**

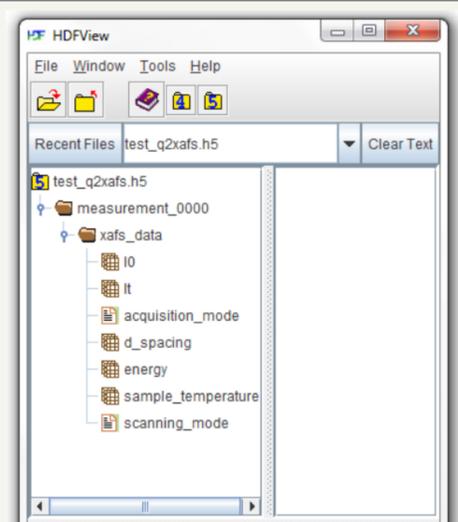
Table of metadata fields and meanings

## Options for Storing and Managing Multiple XDI Spectra

XDI describes exactly one XAFS  $\mu(E)$  or  $\chi(k)$  spectra. Obviously, having a community standard for storing and managing libraries of spectra would be useful.

Below are three options that could be used to hold multiple spectra, and be able to read / write to XDI files. each needs more explanation!

## HDF5 – Heierarchical Data Format



## CIF – Crystallographic Information File

```
data_v2o5_nanotube
_xafs_absorber.atom V
_xafs_absorber.edge K
_xafs_source.identification 'KEK-PF BL20B'
_xafs_source.location 'Tsukuba, Japan'
loop_
_xafs_detectors.label
_xafs_detectors.position
_xafs_detectors.type
monitor monitor ionisation
io-detector detector ionisation foil foil ionisation

loop_
_xafs_ionisation_detector.label
_xafs_ionisation_detector.gas_pressure
_xafs_ionisation_detector.length
_xafs_ionisation_detector.amplifier_type
_xafs_ionisation_detector.amplifier_gain
monitor 1 10 'Keithley' 10
io-detector 1 20 'Keithley' 10
foil 1 5 'Keithley' 11

loop_
_xafs_reduced.energy
_xafs_reduced.absorbance
5248.52108 0.813707373
5258.29435 0.798733337
5268.26606 0.781069442
5278.27878 0.764530778
5288.28697 0.748170706
5298.19834 0.731395959
...
```

## SQLite – relational database in a file

