The International XAFS Society's Interchange Format Draft Specification

1 Introduction

1.1 Purpose

This document describes the International X-ray Absorption Society's Interchange Format (IXASIF), a simple file format for X-ray Absorption Spectroscopy (XAS) experiment data. We are defining this format to accomplish the following goals:

1. Establishing a common language for transferring data between beamlines, data analysis packages, and members of the international XAS community.
2. Increasing the relevance and longevity of experimental data by reducing the amount of *data archeology* future interpretations of that data will require.
3. Enhancing the user experience by promoting interoperability among data acquisition systems and data analysis packages.

In order to fulfill these goals, IXASIF files provide a flexible, consistent representation of information common to all XAS experiments. This format is simpler than a format based on XML or HDF, yields self-documenting files, and is easy for both humans and computers to read. Its structure was inspired by that of Internet electronic mail, a data format which has proven to be robust, extensible, and enduring. Due to these advantages, and because of our intention to develop free software tools and libraries that support IXASIF, we expect that this file format described in this specification will see wide adoption in the XAS community.

1.2 Scope

We do not intend this specification to dictate the file formats used by data acquisition systems during XAS experiments. An attempt to do so would be unreasonable due to the number of different data acquisition systems currently deployed at synchrotrons around the world, the variety of experiments performed at these installations, and the continuing development of new experimental techniques. Instead, this specification addresses the representation of a single scan of XAS data after an experiment has been completed.

Beamlines which adopt this specification will provide their users with tools that convert between their native file formats and IXASIF. In short, they will send their users home with their XAS data stored in this format. We intend to encourage this practice by developing tools to for reading, editing, and writing IXASIF files. Beamlines may choose to modify their data acquisition systems to write data using this format in situations where that would be appropriate. We plan to assist in this effort by developing libraries for popular programming languages which can read, manipulate, and write IXASIF files.

With their experiment data stored in IXASIF files, users may choose data analysis packages which are capable of reading this format. It is our hope that, as this specification gains wider adoption, users will be ultimately freed from the responsibility of understanding file formats. With this aim in mind, we shall assist software developers in supporting IXASIF files.

2 The Contents of the IXASIF File

IXASIF files contain two sections, a header with information about one scan of an XAS experiment and the data collected during that scan. The header consists of a series of fields that each contain a single piece of information, an area for users to store comments about the experiment, and a sequence of labels for the columns of data. The data section contains these columns, with each row corresponding to one point of the scan.

Although the header has been designed to contain arbitrary information, the meanings of several fields are explicitly defined. These fields, described below, contain the most common information about XAS experiments. We hope that users will benefit from their existence when using data analysis packages that support IXASIF files. However, none of the standard fields are required to be present. For example, some of these fields may not be appropriate for certain experiments and should be omitted in that case.

Some examples of header information follow. A complete list of standard headers along with their specifications is found in Sec. 4.1.

Beamline: location where the experiment was performed.

Crystal: information about the monochromator used in the experiment.

D Spacing: interplaner spacing of the monochromator's crystals.

Edge Energy: edge energy value defined by the data acquisition software.

Focusing: comments about the focusing optics used in the experiment.

Harmonic: undulator harmonic used in the experiment.

Mirrors: comments about the mirrors used in the experiment.

Ring Current: current of the synchrotron's storage ring.

Ring Energy: energy of the synchrotron's storage ring.

Source: type of x-ray source used in the experiment.

Step Scale: factors for converting the abscissa to photon energy

Timestamps: start and end times of this scan.

Mu-transmission: equation for calculating the experimental absorption coefficient from the data columns.

3 Definintion of the IXS Interchange Format

This section of the IXASIF specification formally describes the structure of IXASIF files.

3.1 Requirements

The key words "**must**", "**must not**", "**required**", "**shall**", "**shall not**", "**should**", "**should not**", "**recommended**", "**may**", and "**optional**" in this document are to be interpreted as described in RFC 2119.

An IXASIF implementation is not compliant if it fails to satisfy one or more of the **must** or **required** level requirements presented in this specification.

3.2 Notational Conventions

All of the representations defined in this document are described both in prose and using an augmented Backus-Naur Form (BNF). The syntax used in these grammars is defined in section 2.1 of the Internet Engineering Task Force (IETF) Request for Comments (RFC) 2616, "The Hypertext Transfer Protocol". Software developers who wish to implement support for IXASIF files themselves will need to familiarize themselves with this notation to understand this specification.

The basic rules used throughout this section to define parsing constructs are presented in appendix B.1 as part of the complete grammar. All parsing rules that consist of a sequence of multi-character tokens **must** be delimited by whitespace unless the tokens of the sequence may be unambiguously identified.

3.3 Text Encoding

The header and data sections of an IXASIF file are comprised of structured US-ASCII text. Header field values that are "free-form" or "text" **may** contain UTF-8 encoded Unicode text, although Unicode support in applications that use IXASIF files is OPTIONAL. The US-ASCII coded character set is formally by ANSI X3.4-186. The Universal Character Set (Unicode) is defined by ISO/IEC 10646. The UTF-8 translation format is defined by IETF RFC 3629.

3.4 Header Section

The header section of an IXASIF file appears at the beginning of the file and is comprised of structured text. Every line of the header **must** begin with a "#" character and **must** end with the end-of-line sequence defined below. Support for the Posix, Apple, or Microsoft end-of-line conventions is intended to increase cross-platform portability.

```
EOL             = CR  |  LF |  CRLF
```

Open Issue: *Do we need a maximum line length? BR: Preferably not – files with MED data can be very wide*

Header lines are subdivided into four subsections: version information, header fields, user comments, and column labels. With the exception of the version information line and the header fields, all of these subsections are separated by a dividing line of one or more hyphens.

```
DIVIDING-LINE  = "#"  1*"-"  EOL
```

3.4.1 Version Information

The first line of the IXASIF header contains the IXASIF version that the file conforms to. IXASIF uses a `<major>.<minor>` numbering scheme to represent versions of the file format. The `<minor>` version is incremented when changes are made to the format that do not affect compatibility with previous versions, as when new standard header fields are defined. The `<major>` version is incremented when other changes are made to the format, as when the definition of the contents of a standard header field is altered.

A series of optional version entries, separated by whitespace, may follow the IXASIF version. These version entries exist to allow various programs to annotate the file as it proceeds through the collection and analysis process. Such annotation is **optional** although version information **must** be included in this sequence by software that create IXASIF files containing extension fields (see section 4.2). The order of the optional version entries is undefined but **should** be preserved to accurately represent the sequence of programs which have manipulated the file.

```
IXASIF-VERSION  = "IXASIF"  "/"  1*DIGIT  ". " 1*DIGIT
OTHER-VERSION   = 1*ALPHA  "/"   *DIGIT  1*("." 1*DIGIT)
VERSION         = "#"  IXASIF-VERSION  *OTHER-VERSION  EOL
```

Note that the major and minor version numbers **must** be treated as separate integers that **may** contain more than a single digit. As a result, "IXASIF/1.12" is a lower version than "IXASIF/1.2".

3.4.2 Header Fields

The lines immediately following the version line of the header contain the fields of the header. These fields are arranged in a manner similar to the the header of an Internet electronic mail message, although IXASIF fields may not span multiple lines. Each field consists of a case-insensitive name, a separating colon, and an associated value. When multiple occurrences of the same field are present the value of the last occurrence **must** be used as the value for the field.

Although the values of some fields have a required structure, all values are assumed to be free-form text in the following rules. Rules for each of the standard fields are defined in section 4.1 and the complete definition of the `FIELDS` rule may be found in section 4.3.

```
FIELD-NAME     = ALPHA  *(ALPHA  |  DIGIT  |  "-")
FIELD-VALUE    = *TEXT
FIELD-LINE     = "#"  FIELD-NAME  ":"  FIELD-VALUE EOL
FIELDS         = *FIELD-LINE  DIVIDING-LINE
```

The header fields subsection is ended with a dividing line. Note that because no fields are required to be present, this subsection may contain no lines. Even so, the dividing line **must** be present. Also note that because `FIELD-VALUE` matches zero or more `TEXT` characters, it is not required to contain any text.

3.4.3 User Comments

Following the dividing line at the end of the header fields subsection is the area of the header that contains user comments. Please note that this area is reserved for comments supplied by the experimenters and **must**

notbe used by software as a place to store other information. Refer to section 4.2 for information about using extension fields for this purpose.

```
COMMENT-LINE   = "#"  *TEXT  EOL
COMMENTS       = *COMMENT-LINE  DIVIDING-LINE
```

As with the header fields, this section may contain no lines of commentary or lines that contain no comment text but **must** end with a dividing line. When extracting the comment subsection from an IXASIF file, software **should** remove a single leading space and any trailing white space from each comment line but **must not** further alter the line's contents.

3.4.4 Column Labels

The final line of the IXASIF header contains the labels for each column of data in the data section of the file, separated by whitespace. There **must** be one label present for each column of data present in the data section.

```
LABEL          = *ALPHA  *(ALPHA  |  DIGIT  |  "_"  |  "-")
LABELS         = "#"  1*LABEL  EOL
```

3.5 Data Section

The data section of the file contains whitespace delimited columns of floating-point numbers. The first column of this section **must** contain the abscissa and the remaining columns **must** correspond to experimental values at that abscissa. If the abscissa is not the photon energy, then either or both of the `Step-scale` and `Step-offset` fields **must** be present in the header. See section 4.1 for more information about how these fields are used to calculate the photon energy from the abscissa.

```
DATA-LINE      = *FLOAT EOL
DATA           = *DATA-LINE
```

Open Issue: *single-precision v.s. double-precision representations (e.g. floats vs. doubles)*

Note that blank lines in this section **must** be discarded, the number of columns **must** be the same for all lines that contain data, and any measurements of times present **must** be represented as floating point numbers.

4 IXASIF Fields

4.1 Standard Fields

When present, the following header fields **must** comply with their associated parsing rules. Any fields which fail to do so **must** be ignored by preprocessing and analysis software. The text in brackets to the right of

the token provides a quick overview of the expected format, and any text following the line of dots is an example of a valid value.

♦ **Beamline:** `[<text>]` ..... *APS 10ID*
The location where the experiment was performed.

```
BEAMLINE       = "#"  "Beamline"  ":"  *TEXT EOL
```

♦ **Crystal:** `[<material> <cut>]` ..... *Si 111*
Information about the crystals of the monochromator used in the experiment.

```
MATERIAL       = "Si"
CUT            = 3*DIGIT
CRYSTAL        = "#" "Crystal"  ":" MATERIAL CUT EOL
```

Open Issue: *what other materials should we support? should it be free-form instead? BR: Probably, although a nominal list (Ge, Diamond, YB66, and so on) could be implemented*

♦ **D-spacing:** `[<float>]` *3.13555*
The interplaner spacing of the monochromator's crystals, in angstroms.

```
DSPACING       = "#"  "D-spacing"  ":"  FLOAT  EOL
```

♦ **Edge-energy:** `[<float>]` *5465*
The absorption edge of the absorbing, as defined in the data acquisition software, in electron volts.

```
EDGEENERGY     = "#"  "Edge-energy"  ":"  FLOAT  EOL
```

♦ **End-time:** `[<timestamp>]` *2005-05-03 13:42:23*
The date and time that this scan of the experiment completed, in ISO format.

```
DATE           = 4*DIGIT  "-"  2*DIGIT  "-"  2*DIGIT
TIME           = 2*DIGIT  ":"  2*DIGIT  ":"  2*DIGIT
ENDTIME        = "#"  "End-time"  ":"  DATE  TIME  EOL
```

♦ **Focusing:** `[<text>]` ..... *none*
A brief description of the focusing optics used in the experiment.

```
FOCUSING       = "#"  "Focusing"  ":"  *TEXT  EOL
```

♦ **Harmonic:** `[<int>]` *2*
The undulator harmonic used in the experiment.

```
HARMONIC-VALUE = <any US-ASCII digit "1".."7">
HARMONIC       = "#"  "Harmonic"  ":"  HARMONIC-VALUE  EOL
```

Open Issue: *are seven harmonics enough? BR: probably*

♦ **Mirrors:** `[<text>]` ..... *single harmonic rejection*
A brief description of the mirrors used in the experiment.

```
MIRRORS        = "#" "Mirrors"  ":"  *TEXT  EOL
```

♦ **Mu-fluorescence:** `[<expression>]` ..... *$4/$2*
The expression for calculating the $\mu(E)$ of fluorescence from this file's data section.

```
MUFLUOR        = "#"  "Mu-fluorescence"  ":"  EXPRESSION  EOL
```

Open Issue: *we need to define the* ***EXPRESSION*** *language and the language needs to allow for addition of columns, as for an MED. Does the language allow for dead-time correction calculations?*

♦ **Mu-reference:** `[<expression>]` ..... *ln($3/$5)*
The expression for calculating the $\mu(E)$ of the reference channel from this file's data section.

```
MUREF          = "#"  "Mu-reference"  ":"  EXPRESSION  EOL
```

♦ **Mu-transmission:** `[<expression>]` ..... *ln($2/$3)*
The expression for calculating the $\mu(E)$ of transmission from this file's data section.

```
MUTRANS        = "#"  "Mu-transmission"  ":"  EXPRESSION  EOL
```

♦ **Ring-current:** `[<float>]` *101.2*
The current of the synchrotron's storage ring, in milliamperes.

```
RINGCURRENT    = "#"  "Ring-current"  ":"  FLOAT  EOL
```

♦ **Ring-energy:** `[<float>]` *7.01*
The energy of the synchrotron's storage ring, in giga-electron volts (GeV).

```
RINGENERGY     = "#"  "Ring-energy"  ":"  FLOAT  EOL
```

♦ **Start-time:** `[<timestamp>]` *2005-05-03 11:17:52*
The date and time that this scan of the experiment began, in ISO format.

```
STARTTIME      = "#"  "Start-time"  ":"  DATE  TIME  EOL
```

♦ **Source:** `[<text>]` ..... *undulator A*
The type of x-ray source used in the experiment.

```
SOURCE         = "#"  "Source"  ":"  *TEXT  EOL
```

♦ **Step-offset:** `[<float>]` *0.0*
The photon energy is calculated from the abscissa using the the value of this field in the equation $E(a) = a * \text{scale} + \text{offset}$. If the `Step-scale` field is not present, the default value of 1 is used.

```
STEPOFFSET     = "#"  "Step-offset"  ":"  FLOAT  EOL
```

♦ **Step-scale:** `[<float>]` *1.0*
The photon energy is calculated from the abscissa using the the value of this field in the equation $E(a) = a * \text{scale} + \text{offset}$. If the `Step-offset` field is not present, the default value of 0 is used.

```
STEPSCALE      = "#"  "Step-scale"  ":"  FLOAT  EOL
```

4.2 Extension Fields

Extension fields are fields present in the header of an IXASIF file that are not defined in that file's version of IXASIF. Such fields are interpreted as having values of free-form text. Any field not defined in section 4.1 **must** be considered an extension field, providing backwards compatibility between different minor versions of this specification.

Data acquisition systems and data analysis packages may embed additional information in an IXASIF file by adding extension fields to the header. Extension fields created by applications **must** begin with the same application name used in the version line, followed by a hyphen (see appendix A). This requirement prevents field name collisions between different applications and between applications and future versions of this specification.

Applications that read IXASIF files **may** attempt to parse the values of extension fields to extract the additional information about the scan. They **may** propagate these fields into output files they create, but **must** propagate the associated version information if they do so.

```
EXT-FIELD-NAME = ALPHA  *(ALPHA | DIGIT)  "-"  FIELD-NAME
EXT-FIELD      = "#"  EXT-FIELD-NAME  ":"  *TEXT EOL
```

4.3 Grammar for the Header Fields

Having defined the rules of the standard header fields, it is now possible to create a complete version of the `FIELDS` rule that was provisionally defined in section 3.4.2. This definition may be found in appendix B.3.

A Example XISIF File

Here is an example of a file conforming to this specification. This was edited by hand from a real data file measured at beamline ID10 at the APS.

```
# IXASIF/1.0 MX/2.0
# Crystal: Si 111
# Harmonic: 3
# Beamline: APS 10ID
# Source: undulator a
# Mirrors: single harmonic rejection mirror
# Focussing: none
# Start-time 2005-03-08 20:08:57
# Ring-energy: 7.00
# Edge-energy: 7112.00
# Mu-transmission: ln($2/$3)
# Mu-reference: ln($3/$5)
# MX-Num-regions: 1
# MX-SRB: 6900
# MX-SRSS: 0.5
# MX-SPP: 0.1
# MX-Settling-time: 0
# MX-Offsets: 11408.00 11328.00 13200.00 10774.00
# MX-Gains: 8.00 7.00 7.00 9.00
#---
# Fe K-edge, Lepidocrocite powder on kapton tape, RT
# 4 layers of tape
# exafs, 20 invang
#---
# energy     mcs3     mcs4     mcs6     mcs5
 6899.9609  48120 19430 2250 54540
 6900.1421  48390 19540 2260 54860
 6900.5449  48520 19610 2250 55110
 6900.9678  48930 19780 2280 55650
 6901.3806  48460 19590 2250 55110
     (....etc....)
```

B Grammar of the IXASIF

```
IXASIF          = VERSION FIELDS COMMENTS LABELS DATA
```

B.1 Basic Constructs

```
OCTET           = <any 8-bit sequence of data>
UPALPHA         = <any US-ASCII uppercase letter "A".."Z">
LOALPHA         = <any US-ASCII lowercase letter "a".."z">
ALPHA           = UPALPHA | LOALPHA
DIGIT           = <any US-ASCII digit "0".."9">
CTL             = <any US-ASCII control character (octets 0 - 31)
                   and DEL (127)>
CR              = <US-ASCII CR, carriage return (13)>
LF              = <US-ASCII LF, linefeed (10)>
SP              = <US-ASCII SP, space (32)>
HT              = <US-ASCII HT, horizontal-tab (9)>
WS              =  SP  |  HT
EOL             = CR  |  LF  |  CRLF
TEXT            = <any OCTET except CTLs, but including WS>

SIGN            = "+"  |  "-"
EXPONENT        = ("e"  |  "E"  |  "d"  |  "D")  [SIGN]  1*DIGIT
NUMBER          = 1*DIGIT  ["."  *DIGIT]  [EXPONENT]
INF             = ("i"  |  "I")  ("n"  |  "N")  ("f"  |  "F")
NAN             = ("n"  |  "N")  ("a"  |  "A")  ("n"  |  "N")
FLOAT           = [SIGN]  (NUMBER  |  INF  | NAN )

DIVIDING-LINE   = "#"  1*"-"  EOL
```

B.2 Version Information

```
IXASIF-VERSION  = "IXASIF"  "/"  1*DIGIT  ". " 1*DIGIT
OTHER-VERSION   = 1*ALPHA  "/"   *DIGIT  1*("." 1*DIGIT)
VERSION         = "#"  IXASIF-VERSION  *OTHER-VERSION  EOL
```

B.3 Header Fields

```
FIELDS          = *(  BEAMLINE    |  CRYSTAL    |  DSPACING
                    | EDGEENERGY  |  ENDTIME    | FOCUSSING
                    | HARMONIC    |  MIRRORS    | MUFLUOR
                    | MUREF       |  MUTRANS    | RINGCURRENT
                    | RINGENERGY  |  STARTTIME  | SOURCE
                    | STEPOFFSET  |  STEPSCALE  | EXT-FIELD
                    | FIELD-LINE
                   ) DIVIDING-LINE
```

B.3.1 Standard Fields

```
CUT            = 3*DIGIT
ELEMENT        = "Si"
DATE           = 4*DIGIT  "-"  2*DIGIT  "-"  2*DIGIT
TIME           = 2*DIGIT  ":"  2*DIGIT  ":"  2*DIGIT
HARMONIC-VALUE = <any US-ASCII digit "1".."7">

BEAMLINE       = "#"  "Beamline"      ":"  *TEXT           EOL
CRYSTAL        = "#"  "Crystal"       ":"  ELEMENT  CUT    EOL
DSPACING       = "#"  "D-spacing"     ":"  FLOAT           EOL
EDGEENERGY     = "#"  "Edge-energy"   ":"  FLOAT           EOL
ENDTIME        = "#"  "End-time"      ":"  DATE  TIME      EOL
FOCUSING       = "#"  "Focusing"      ":"  *TEXT           EOL
HARMONIC       = "#"  "Harmonic"      ":"  HARMONIC-VALUE  EOL
MIRRORS        = "#"  "Mirrors"       ":"  *TEXT           EOL
MUFLUOR        = "#"  "Mu-fluor"      ":"  EXPRESSION      EOL
MUREF          = "#"  "Mu-ref"        ":"  EXPRESSION      EOL
MUTRANS        = "#"  "Mu-trans"      ":"  EXPRESSION      EOL
RINGCURRENT    = "#"  "Ring-current"  ":"  FLOAT           EOL
RINGENERGY     = "#"  "Ring-energy"   ":"  FLOAT           EOL
STARTTIME      = "#"  "Start-time"    ":"  DATE  TIME      EOL
SOURCE         = "#"  "Source"        ":"  *TEXT           EOL
STEPOFFSET     = "#"  "Step-offset"   ":"  FLOAT           EOL
STEPSCALE      = "#"  "Step-scale"    ":"  FLOAT           EOL
```

B.3.2 Extension Fields

```
FIELD-NAME     = ALPHA  *(ALPHA  |  DIGIT  |  "-")
EXT-FIELD-NAME = ALPHA  *(ALPHA  |  DIGIT)  "-"  FIELD-NAME

EXT-FIELD      = "#"  EXT-FIELD-NAME  ":"  *TEXT  EOL
FIELD-LINE     = "#"  FIELD-NAME      ":"  *TEXT  EOL
```

B.4 User Comments

```
COMMENT-LINE   = "#"  *TEXT  EOL
COMMENTS       = *COMMENT-LINE  DIVIDING-LINE
```

B.5 Column Labels

```
LABEL          = *ALPHA  *(ALPHA  |  DIGIT  |  "_"  |  "-")
LABELS         = "#"  1*LABEL  EOL
```

B.6 Data Section

```
DATA-LINE      = *FLOAT EOL
DATA           = *DATA-LINE
```

C References On the Web

To do: *add more references*

ASCII tables
`http://www.lookuptables.com/`

Timezone information
`http://wwp.greenwichmeantime.com/info/timezone.htm`

RFC822: Standard for ARPA Internet Text Messages
`http://www.w3.org/Protocols/rfc822/`

Farrel Lytle database at IIT
`http://ixs.csrri.iit.edu/database/data/index.html`