



Standard Guide for Analytical Data Interchange Protocol for Mass Spectrometric Data¹

This standard is issued under the fixed designation E 2078; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide covers the implementation of the Mass Spectrometric Data Protocol in analytical software applications. Implementation of this protocol requires:

1.1.1 Specification E 2077, which contains the full set of data definitions. The mass spectrometric data protocol is not based upon any specific implementation; it is designed to be independent of any particular implementation so that implementations can change as technology evolves. The protocol is implemented in categories to speed its acceptance through actual use.

1.1.2 Specification E 2077 contains a full description of the contents of the data communications protocol, including the analytical information categories with data elements and their attributes for most aspects of mass spectrometric tests.

1.2 The analytical information categories are a practical convenience for breaking down the standardization process into smaller, more manageable pieces. It is easier for developers to build consensus and produce working systems based on smaller information sets, without the burden and complexity of the hundreds of data elements contained in all the categories. The categories also assist vendors and end users in using the guide in their computing environments.

1.3 The network common data format (NetCDF) data interchange system is the container used to communicate data between applications in a way that is independent of both computer architectures and end-user applications. In essence, it is a special type of application designed for data interchange.

1.4 The common data language (CDL) template for mass spectrometry is a language specification of the mass spectrometry dataset being interchanged. With the use of the NetCDF utilities, this human-readable template can be used to generate an equivalent binary file and the software subroutine calls needed for input and output of data in analytical applications.

2. Referenced Documents

2.1 ASTM Standards:

E 2077 Specification for Analytical Data Interchange Pro-

tol for Mass Spectrometric Data²

2.2 Other Standard:

NetCDF User's Guide³

2.3 ISO Standards:⁴

8601:1988 Data elements and interchange formats, (First edition published 1988-06-15; with Technical Corrigendum 1 published 1991-05-01)

3. List of Contents and Use

3.1 *NetCDF Toolkit*—The protocol is an application programming interface (API) layered on top of the public domain NetCDF toolkit. NetCDF is a set of tools that facilitate reading or writing platform-independent, self-describing data files. All data in a NetCDF file is written using the external data representation (XDR). XDR was developed by Sun Microsystems and is used for platform-independent file systems for all workstations and personal computers. Each NetCDF data element is self-describing - it has a name, type, and dimensionality. A NetCDF file contains three parts: a *dimensions* section, which defines the names and size of all dimensions used to describe variables; a *variables* section, which defines the names, data types, dimensionality, and attributes for all variables used in the file; and finally, a *data* section, which contains the actual values assigned to the variables. Attributes are numbers or strings which augment the description of variables or the file as a whole.

3.1.1 For example, a variable “x_axis_values” might contain an array of numbers representing the abscissa of a two-dimensional data set. It would have a dimension, possibly named “x_axis_size,” which would specify the number of abscissa points. The variable might have some descriptive attributes, such as “units” (with a value of “Seconds,” perhaps), “scale_factor” (with a value of 1000.0, specifying that all stored abscissa values should be multiplied by 1000.0 to get the actual value), or “long_name” (with value “Time”, which might be used to label the abscissa when drawing a plot).

3.1.2 The NetCDF toolkit has been placed in the public domain by the Unidata Program Center, a non-profit software

¹ This specification is under the jurisdiction of ASTM Committee E01 on Analytical Chemistry for Metals, Ores and Related Materials and is the direct responsibility of Subcommittee E01.25 on Laboratory Data Interchange and Information Management.

Current edition approved March 10, 2000. Published July 2000.

² *Annual Book of ASTM Standards*, Vol 03.06.

³ Available for Russell K. Rew, Unidata Program Center, University Corporation for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000.

⁴ Available from ISO, 1 Rue de Varembe, Case Postale 56, CH 1211, Geneva, Switzerland.