

## Olis Dataset File Specification version 1.0

### Syntax

Structure of the Olis dataset file (later in the document referred as “file”) resembles structure of XML documents. Data elements are enclosed in a pair of opening and closing tags. Each tag and data will start on a new line as follows:

```
<Name>
Absorbance
</Name>
```

Because parsing of the file relies on angle bracket characters (<>), use of these characters is reserved for tags only.

Note that element indentation shown in examples in the rest of this document is used for illustration purposes only.

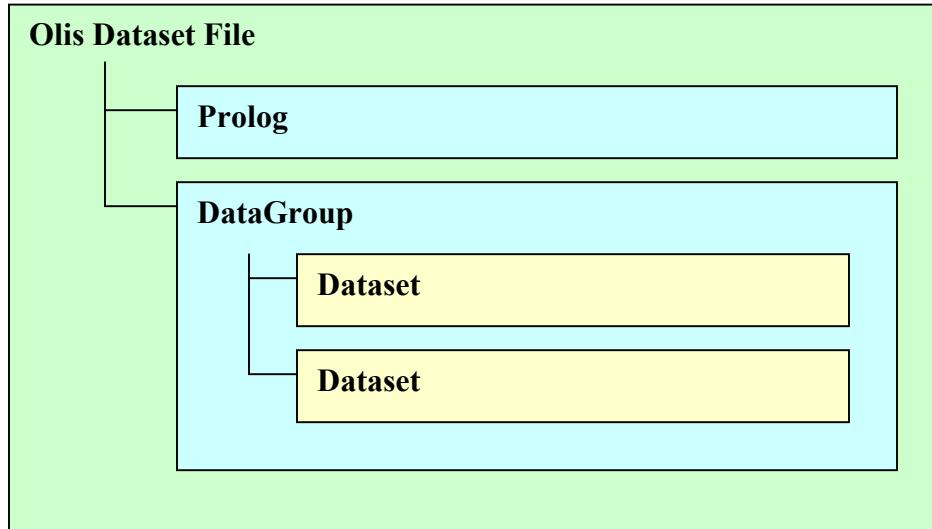
Data elements can be nested (i.e. embedded one within another). Element nesting sets up parent /child relationship. Every child element resides completely within its parent element.

```
<Parent>
  <Child1>
    data here
  </Child1>
  <Child2>
    data here
  </Child2>
</Parent>
```

Data of all elements other than BinData, are written in a file as strings of ASCII characters. Data of BinData elements are written as a binary array of bytes. Each 8 bytes of the binary array represent a one double-precision floating-point value. Therefore the size of the binary array is NumberOfPoints \* 8 bytes. For two-dimensional arrays, data layout follows row-major form used by SafeArrays. This means that moving to next array element increments the last index. For example, if a two-dimensional array (e.g. Data(i,j)) is written in such one-dimensional binary byte array form, moving to the next 8 byte element of the binary array increments last index of the original two-dimensional array (i.e. Data(i,j+1)). After the last element of the binary array the combination of carriage return and linefeed characters (ANSI characters 13 and 10) is written.

## Logical Data Structure

Logical structure of the data file is presented on the figure below:



The first structural element in a file is a mandatory prolog that identifies the type of file as Olis dataset and its specification version:

<Olis dataset version 1.0>

Next structural element is DataGroup. It contains one or more Datasets. Datasets are usually grouped when there is a relationship among them, for example they compose a result of SVD analysis or Global fit.

**DataGroup** required elements:

<b>Name</b>	-arbitrary name
<b>Dataset</b>	-at least one Dataset must be present

Dataset element describes a single set of Z data as a function of X (2d data) or both X and Y (3d data).

**Dataset** required elements:

<b>Name</b>	-arbitrary name
<b>Type</b>	-designates origin of data (collected by an instrument; created as a result of SVD analysis, fitting, etc.); Data collected by instruments is designated by constant 3301.
<b>XAxis</b>	-element describing X axis
<b>YAxis</b>	-element describing Y axis
<b>ZAxis</b>	-element describing Z axis; Note: XAxis and YAxis elements must precede the ZAxis element

**XAxis, YAxis** required elements:

**Name** -name of axis data  
**Units** -units of axis data  
**IsLinear** -True if data is distributed in a linear manner along this axis;  
False if not linear  
**Number of Points** -number of data points  
**Start** -value of the first axis point; required only if IsLinear = True  
**Step** -step between two consecutive axis points; required only if  
IsLinear = True; For linear point distribution one can compute  
values along the axis using Number of Points, Start and Step  
values.  
**BinData** -contains values of all axis points; data is written in a binary form;  
required only if IsLinear = False;

**ZAxis** required elements:

**Name** -name of axis data  
**Units** -units of axis data  
**BinData** -contains values of all points; data is written in a binary form;

**Important:** Some files may contain additional elements. Specification of these elements has not been finalized. Therefore, these elements should not be used for data import. Using such elements may cause errors.

Following example shows a file, which contains a single 3d dataset (Absorbance as a function of Wavelength and Time). The example shows all required elements. Note that indentation in the example is shown for illustration purposes only. Note that X Axis has a linear data distribution and Y Axis has a non-linear distribution.

```
<Olis dataset version 1.0>
<DataGroup>
  <Name>
  GroupName
  </Name>
  <Dataset>
    <Name>
    DatasetName
    </Name>
    <Type>
    3301
    </Type>
```

```
<XAxis>
  <Name>
  Wavelength
  </Name>
  <Units>
  nm
  </Units>
  <IsLinear>
  True
  </IsLinear>
  <Number of Points>
  201
  </Number of Points>
  <Start>
  400
  </Start>
  <Step>
  2
  </Step>
</XAxis>
<YAxis>
  <Name>
  Time
  </Name>
  <Units>
  sec
  </Units>
  <IsLinear>
  False
  </IsLinear>
  <Number of Points>
  150
  </Number of Points>
  <BinData>
  ....binary data here
  </BinData>
</YAxis>
<ZAxis>
  <Name>
  Absorbance
  </Name>
  <Units>
  </Units>
  <BinData>
  .... binary data here
  </BinData>
</ZAxis>
</Dataset>
</DataGroup>
```