

cansas1d documentation

From canSAS

Contents

- 1 Disclaimer
- 2 Objective
 - 2.1 Status
- 3 General Layout of the XML Data
 - 3.1 Overview
 - 3.1.1 Basic elements of the cansas1d/1.0 standard
 - 3.1.2 Required XML file header
 - 3.2 Rules
 - 3.2.1 Compatibility of Geometry Definitions
- 4 Documentation and Definitions
 - 4.1 XML Schema
 - 4.2 XML Stylesheets
 - 4.3 Examples and Case Studies
 - 4.3.1 XML layout for multiple experiments
 - 4.4 Foreign Elements
 - 4.5 Support tools for Visualization & Analysis software
 - 4.5.1 Binding for IgorPro
 - 4.5.2 Binding for Java
 - 4.5.3 Support for Python
 - 4.5.4 Support for FORTRAN
 - 4.5.5 Support for Microsoft Excel
 - 4.6 Software repositories
- 5 Validation of XML against the Schema
- 6 Help for XML

Disclaimer

This description is meant to inform the community how to layout the information within the XML files. However, should the information in this document and the cansas1d/1.0 SAS XML Schema (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cansas1d.xsd>) differ, the XML Schema will be deemed to have the most correct description of the standard.

Objective

One of the first aims of the canSAS (Collective Action for Nomadic Small-Angle Scatterers) forum of users, software developers, and facility staff was to discuss better sharing of SAS data analysis software. CanSAS

(<http://www.smallangles.net/canSAS/>) identified that a significant need within the SAS community can be satisfied by a robust, self-describing, text-based, standard format to communicate reduced one-dimensional small-angle scattering data, $I(Q)$, between users of our facilities. Our goal has been to define such a format that leaves the data file instantly human-readable, editable in the simplest of editors, and importable by simple text import filters in programs that need not recognise advanced structure in the file nor require advanced programming interfaces. The file should contain both the primary data of $I(Q)$ and also any other descriptive information (metadata) about the sample, measurement, instrument, processing, or analysis steps.

The cansas1d/1.0 standard meets the objectives for a 1D standard, incorporating metadata about the measurement, parameters and results of processing or analysis steps. Even multiple measurements (related or unrelated) may be included within a single XML file.

Status

Version 1.0 was tagged from the subversion repository on 2009-05-12 as no changes were committed since January 2009. Use this command to checkout the tagged release.

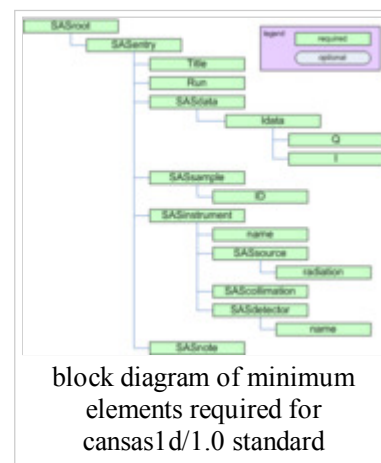
```
svn checkout http://svn.smallangles.net/svn/canSAS/ldwg/tags/v1.0 cansas1dwg-1.0
```

General Layout of the XML Data

The canSAS 1-D standard for reduced 1-D SAS data is implemented using XML files. A single file can contain SAS data from a single experiment or multiple experiments. All types of relevant data ($I(Q)$, metadata) are described for each experiment. More details are provided below.

Overview

The basic elements of the cansas1d/1.0 standard are shown in the following table. After an XML header, the root element of the file is SASroot which contains one or more SASentry elements, each of which describes a single experiment (data set, time-slice, step in a series, new sample, etc.). Details of the SASentry element are also shown in the next figure. Refer to the block diagrams for alternative depictions. See cansas1d.xml (<http://svn.smallangles.net/trac/canSAS/browser/ldwg/trunk/cansas1d.xml>) for an example XML file. Examples, Case Studies, and other background information are below. More discussion can be found on the canSAS 1D Data Formats Working Group page and its discussion page. Details about each specific field (XPath string, XML elements and attributes) are described on the cansas1d_definition_of_terms page.



- **SASroot**: the root element of the file (after the XML header)
- **SASentry**: describes a single experiment (data set, time-slice, step in a series, new sample, etc.)
- block diagrams
- cansas1d.xml (<http://svn.smallangles.net/trac/canSAS/browser/ldwg/trunk/cansas1d.xml>) example XML file
- discussion of this format: basic more
- Seek outside help for XML

- **Definition of terms:** Details about each specific field (XPath string, XML elements and attributes)

Basic elements of the cansas1d/1.0 standard

element	description
XML header	descriptive info required at the start of every XML file
SASroot	
SASentry	data set, time-slice, step in a series, new sample, etc.
Title	for this particular SASentry
Run	run number or ID number of experiment
{any}	any non-cansas1d/1.0 element can be used at this point
SASdata	this is where the reduced 1-D SAS data is stored
Idata	a single data point in the dataset
{any}	any non-cansas1d/1.0 element can be used at this point
SASsample	description of the sample
SASinstrument	description of the instrument
SASsource	description of the source
SAScollimation	description of the collimation
SASdetector	description of the detector
SASprocess	for each processing or analysis step

SASnote

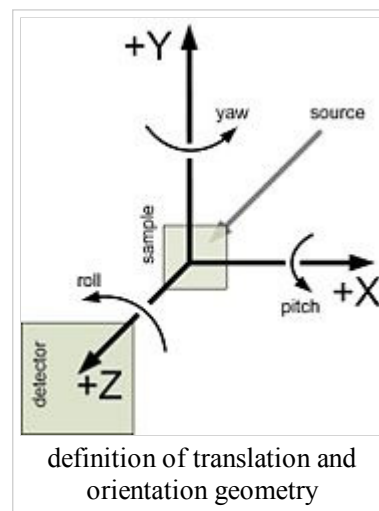
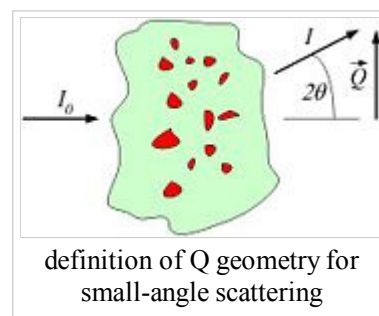
anything at all

Required XML file header

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="cansasxml-html.xsl" ?>
<SASroot version="1.0"
  xmlns="cansas1d/1.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="cansas1d/1.0 http://svn.smallangles.net/svn/canSAS/1dwg/trunk/cansas1d.xsd"
>
```

Rules

1. canSAS1d/1.0 XML data files will adhere to the standard if they can successfully validate against the established XML Schema (cansas1d.xsd (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cansas1d.xsd>))
2. $Q = (4\pi / \lambda) \sin(\theta)$
where λ is the wavelength of the radiation and 2θ is the angle through which the detected radiation has been scattered.
3. units to be given in standard SI abbreviations (eg, m, cm, mm, nm, K) with the following exceptions:
 1. um=micrometres
 2. C=celsius
 3. A=Angstroms
 4. percent=.
 5. fraction
 6. a.u.=arbitrary units
 7. none=no units are relevant (such as dimensionless)
4. where reciprocal units need to be quoted the format shall be "1/abbreviation"
5. when raised to a power, use similar to "A^3" or "1/m^4" (and not "A3" or "m-4")
6. axes:
 1. z is along the flight path (positive value in the direction of the detector)
 2. x is orthogonal to z in the horizontal plane (positive values increase to the right when viewed towards the incoming radiation)
 3. y is orthogonal to z and x in the vertical plane (positive values increase upwards)
7. orientation (angles):
 1. roll is about z
 2. pitch is about x
 3. yaw is about y
8. Unicode characters MUST NOT be used
9. Binary data is not supported



Compatibility of Geometry Definitions

Note: translation and orientation geometry used by canSAS are consistent with:

1. http://en.wikipedia.org/wiki/Cartesian_coordinate_system
2. http://en.wikipedia.org/wiki/Right-hand_rule
3. http://www.nexusformat.org/Coordinate_Systems
4. <http://mcstas.risoe.dk/documentation/tutorial/node6.html>
5. <http://webhost5.nts.jhu.edu/reza/book/kinematics/kinematics.htm>

The translation and orientation geometry definitions used here are different than those used by **SHADOW** (<http://www.nanotech.wisc.edu/shadow/>) where the *y* and *z* axes are swapped and the direction of *x* is changed.

Documentation and Definitions

- **Documentation:** [cansas1d_documentation](#) (this page)
- **Definitions:** [cansas1d_definition_of_terms](#)
- **Block diagrams:** [cansas1d_block_diagrams](#)

XML Schema

XML Schema: The [cansas1d.xsd](http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cansas1d.xsd) (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cansas1d.xsd>) XML Schema (<http://www.w3schools.com/xsd>) defines the rules for the XML file format (TRAC (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cansas1d.xsd>) , SVN (<http://svn.smallangles.net/svn/canSAS/1dwg/trunk/cansas1d.xsd>)) and is used to validate any XML file for adherence to the format.

XML Stylesheets

- **cansasxml-html.xsl:** XSLT stylesheets (<http://www.w3schools.com/xsl/>) can be used to extract metadata or to convert into another file format. The default canSAS stylesheet [[cansasxml-html.xsl](http://svn.smallangles.net/svn/canSAS/1dwg/trunk/cansasxml-html.xsl) (<http://svn.smallangles.net/svn/canSAS/1dwg/trunk/cansasxml-html.xsl>)] should be copied into each folder with canSAS XML data file(s). It can be used to display the data in a supporting WWW browser (such as Firefox or Internet Explorer) or to import into Microsoft Excel (with the added XML support in Excel). (See the excellent write-up by Steve King, ISIS, at http://www.isis.rl.ac.uk/LargeScale/LOQ/xml/cansas_xml_format.pdf for an example.) By default, MS Windows binds *.xml files to start Internet Explorer. Double-clicking on a canSAS XML data file with the **cansasxml-html.xsl** (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cansasxml-html.xsl>) stylesheet in the same directory will produce a WWW page with the SAS data and selected metadata.
- Suggestion for support software that writes canSAS1d/1.0 XML data files:
 - be sure to update to the latest SVN repository revision (command: `svn update`)
 - check the output directory to see if it contains the default XSLT file.
 - copy the latest XSLT file to the output directory if either:
 - the output directory contains an older revision
 - the output directory does not have the default XSLT file
 - The most recent XSLT file can be identified by examining the file for the **\$ Revision:** string. For example

```
# $Revision: 66 $
```

is version 66 (updated 2009-01-12).

Examples and Case Studies

- `cansas1d.xml` (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cansas1d.xml>) basic example: Note that, for clarity, only one row of data is shown. This is probably a very good example to use as a starting point for creating XML files with a text editor.
- `bimodal-test1.xml` (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/bimodal-test1.xml>) : Simulated SAS data to test size distribution calculation routines.
- `dry chick collagen`: illustrates the minimum information necessary to meet the requirements of the standard format
- `AF1410 steel`: SANS study using magnetic contrast variation (with multiple samples and multiple data sets for each sample), the files can be viewed from TRAC (no description yet): <http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/examples/af1410/>
- `cansas1d-template.xml` (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cansas1d-template.xml>) : This is used to test all the rules in the XML Schema. This is probably not a very good example to use as a starting point for creating XML files with a text editor since it tests many of the special-case rules.

XML layout for multiple experiments

Each experiment is described with a single **SASentry** element. The fragment below shows how multiple experiments can be included in a single XML file. Full examples of canSAS XML files with multiple experiments include:

- `ISIS LOQ SANS instrument` (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/W1W2.XML>)
- `NIST SANS data` (http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/cs_af1410.xml)

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="cansasxml-html.xsl" ?>
<SASroot version="1.0"
  xmlns="cansas1d/1.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="cansas1d/1.0 http://svn.smallangles.net/svn/canSAS/1dwg/trunk/cansas1d.xsd"
>
  <SASentry name="071121.dat#S22">
    <!-- contents of the first experiment in the file go here -->
  </SASentry>
  <SASentry name="example temperature series">
    <!-- example with two SAS data sets related to the same sample -->
    <Title>title of this series</Title>
    <Run name="run1">42-001</Run>
    <Run name="run2">42-002</Run>
    <SASdata name="run1">
      <!-- data from 42-001 run comes here -->
    </SASdata>
    <SASdata name="run2">
      <!-- data from 42-002 run comes here -->
    </SASdata>
    <!-- other elements come here for this entry -->
  </SASentry>
  <SASentry name="other sample">
    <!-- any number of additional experiments can be included, as desired -->
    <!-- SASentry elements in the same XML file do not have to be related -->
  </SASentry>
</SASroot>
```

Foreign Elements

To allow for inclusion of elements that are not defined by the cansas1d.xsd XML Schema, XML *foreign elements* are permitted at select locations in the cansas1d/1.0 format. Please refer to the references (and others) below for deeper discussions on foreign elements.

No examples exist. At present, all examples of canSAS xml files using foreign namespaces have been converted to bring that data into either the **SASprocessnote** or **SASnote** elements. Refer to the TRAC changes (<http://svn.smallangles.net/trac/canSAS/changeset/47>) for an example of arranging the content in **SASprocessnote** to avoid the use of foreign namespace elements.

Support tools for Visualization & Analysis software

(under development as of May 2008)

Binding for IgorPro

An import tool (a.k.a. binding) for IgorPro (<http://www.wavemetrics.com/>) has been created (cansasXML.ipf (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/IgorPro/cansasXML.ipf>)). Documentation is available.

Note that this tool is not a true binding (http://en.wikipedia.org/wiki/Binding_%28computer_science%29) in that the structure of the XML file is not replicated in IgorPro data structures. This tool reads the vectors of 1-D SAS data (Q, I, ...) into IgorPro waves (Qsas, Isas, ...). The tool also reads most of the metadata into an IgorPro textWave for use by other support in IgorPro.

Status:

- Capabilities
 - test suite of XML files developed
 - import the XML files into IgorPro
- open TRAC tickets (as of 2008-05-15)
 - <http://svn.smallangles.net/trac/canSAS/ticket/3>
 - <http://svn.smallangles.net/trac/canSAS/ticket/10>
- Further
 - conversation with Andrew Nelson shows more efficiency can be gained from minor redesign.
 - Development of a GUI (to support the Irena (<http://usaxs.xor.aps.anl.gov/staff/ilavsky/irena.html>) package) has begun
 - Development to add export capabilities (from IgorPro) back to the cansas1d/1.0 format has begun

Binding for Java

A JAXB (<http://java.sun.com/developer/technicalArticles/WebServices/jaxb/>) binding (http://en.wikipedia.org/wiki/Binding_%28computer_science%29) for Java has been created. (<http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/java/cansas1d>). Documentation is in progress.

Note that this tool replicates the structure of the XML file into Java data structures. Additional Java software must be applied to convert the /SASdata/Idata/* elements into vectors of $I(Q)$ data.

Status:

- Capabilities
 - import the XML files into Java

- tested by use in collimation-correction (desmearing) software (not publicly available at this time)
- Further
 - need to test export capabilities

Support for Python

Specific support (<http://danse.us/trac/sans>) for the cansas1d/1.0 data standard in Python is being developed by NIST/NCNR (<http://www.ncnr.nist.gov/>) as part of their contribution to the DANSE (<http://wiki.cacr.caltech.edu/danse>) project. See the Python Documentation page for more details.

Support for FORTRAN

Steve King[1] (<mailto:s.m.king@rl.ac.uk>) (ISIS) has provided a F77 routine (SASXML_G77.F (http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk/fortran/SASXML_G77.F)) that will read CanSAS XML v1.0 files. See the Fortran Documentation page for more details.

Support for Microsoft Excel

Support for Microsoft Excel is provided through the default canSAS stylesheet [[cansasxml-html.xml](http://svn.smallangles.net/svn/canSAS/1dwg/trunk/cansasxml-html.xml) (<http://svn.smallangles.net/svn/canSAS/1dwg/trunk/cansasxml-html.xml>)]. An excellent description (http://www.isis.rl.ac.uk/LargeScale/LOQ/xml/cansas_xml_format.pdf) of how to import data from the cansas1d/1.0 format into Excel is available from the ISIS LOQ instrument (<http://www.isis.rl.ac.uk/LargeScale/LOQ/loq.htm>).

Software repositories

- **TRAC:** <http://svn.smallangles.net/trac/canSAS/browser/1dwg/trunk>
- **Subversion:** <http://svn.smallangles.net/svn/canSAS/1dwg>
(svn checkout <http://svn.smallangles.net/svn/canSAS/1dwg/trunk/> cansas-1dwg)

Validation of XML against the Schema

1. open browser to: <http://www.xmlvalidation.com/>
2. paste content of candidate XML file (with reference in the header to the XML Schema as shown above) into the form
3. press <validate>
4. paste content of cansas1d.xsd (<http://svn.smallangles.net/svn/canSAS/1dwg/trunk/cansas1d.xsd>) XSD file into form and press <continue validation>
5. check the results

Help for XML

The various references for help on XML have been moved to their own wiki page: XmlHelp

Retrieved from "http://www.smallangles.net/wgwiki/index.php/cansas1d_documentation"

- This page was last modified 21:12, 12 May 2009.