IAEA
International Atomic Energy Agency

INDC(NDS)- 0565
Distr. LP,NE,SK

INDC International Nuclear Data Committee

XML Schema for Atomic and Molecular Data

Summary Report of Consultants' Meeting

Observatoire de Paris, Paris, France

6–7 December 2007

Prepared by

D. Humbert
International Atomic Energy Agency, Vienna, Austria

April 2008

Selected INDC documents may be downloaded in electronic form from
http://www-nds.iaea.org/indc_sel.html or sent as an e-mail attachment.
Requests for hardcopy or e-mail transmittal should be directed to *services@iaeand.iaea.org*
or to:
Nuclear Data Section
International Atomic Energy Agency
PO Box 100
Wagramer Strasse 5
A-1400 Vienna
Austria

Printed by the IAEA in Austria

April 2010

INDC(NDS)- 0565
Distr. LP,NE,SK

XML Schema for Atomic and Molecular Data

Summary Report of Consultants' Meeting

Observatoire de Paris, Paris, France
6–7 December 2007

Prepared by

D. Humbert

Abstract

Advanced developments in computer technologies offer exciting opportunities for new distribution tools and applications in various fields of physics. The convenient and reliable exchange of data is clearly an important component of such applications. Therefore, in 2003, the A+M Data Unit initiated within the collaborative efforts of the DCN (Data Centre Network) a new standard for atomic, molecular and particle surface interaction data exchange (AM'PSI) based on XML (eXtensible Markup Language). A working group composed of staff from the IAEA, NIST, ORNL and Observatoire Paris-Meudon meets biannually to discuss progress made on the XML schema, and to foresee new developments and actions to be taken to promote this standard for AM/PSI data exchange.

April 2008

TABLE OF CONTENTS

IAEA Consultants' Meeting: XML Schema for Atomic and Molecular Data

6–7 December 2007, Observatoire de Paris, Paris, France

Present

Mr. D. Humbert, IAEA *(Scientific Secretary)*
Ms. M-L. Dubernet (France)
Ms. E. Roueff (France)
Mr. N. Moreau (France)
Mr. Yu. Ralchenko (USA)
Mr. D.R. Schultz (USA)
Mr. R.E.H. Clark (IAEA)

Opening of the Meeting

This two-day meeting was held in Paris at the Observatoire de Paris. Ms. Evelyne Roueff welcomed the participants and opened the meeting.

1. Presentation of the Schema at Various Technical Meetings

Current status of the schema was reported at:

- IAEA DCN Meeting, 3-5 October 2007, Yuri Ralchenko
- IAEA CRP on surface erosion: 1st RCM, 17-19 October 2007, Denis Humbert
- ADAS Workshop, 12-13 October 2007, Denis Humbert
- IVOA Conference, M-L. Dubernet

2. Review of Status of XML Schema

The schema is split into different "containers" to improve maintenance and development. Single- and multi-reactant processes have been merged into a single "processes" container with three complex elements: radiative, autoionization and collisions. Current containers with responsible persons are as follows:

- general schema *Yuri Ralchenko*
- biblio *Dave Schultz*
- functions *Denis Humbert*
- methods *Denis Humbert*
- types and attributes *Yuri Ralchenko*
- statesAtoms *Yuri Ralchenko*
- statesMolecules *M-L Dubernet*
- statesSolids *Dave Schultz*

- Processes
    - Radiative *Evelyne Roueff*
    - Autoionization *Denis Humbert*
    - Collisions *Denis Humbert*

General Notes
Participants noted that several new projects could develop their own XML schema for A+M/PSI data. The current schema should be publicised, and the first version released as soon as possible.

The schema is well advanced. Some applications are already available or under development; this situation has been helpful in ensuring and improving the quality of the schema. XML libraries are available for most common programming languages such as Perl and Python, and files can be checked through these libraries. Data from the NIST ASD database for energy levels can now be as XML, and will soon be extended to cover spectral lines. Similar applications are in progress for the IAEA ALADDIN database and BASECOL at Observatoire de Meudon.

General Notes on Labelling and Structure:
A formal document for the data model structure will be developed, including a compilation of rules formulated during the present and previous meetings. This activity is important to the maintenance of the consistency of the data model.

Action: draft document (*Denis*)
Action: check the entire data model for consistency (*everyone*)

The attribute *bibMethRefs* is now split into two aspects: *bibRef* and *methodRef.*

Some data representations such as isoelectronic series, isonuclear sequences, quantum defects, average atom, band properties and solid spectroscopy do not comply with the current schema,. These topics are deferred to a future release of the schema.

Atom
The AtomState type has no major changes.

Molecule
The molecular excited state is very complex to describe. Developing a detailed model which includes all possible descriptions is impossible, as the physics is still evolving. Discussions took place on the placement of QN (to the left or to the right). The description of the excited states is linear at the present time. For the *description* element, Marie-Lise presented a "boxes schema" – *description* could also get an attribute "encoding", with values like ASCII, html or latex.

Action: Marie-Lise to decide between a linear or parallel structure for the excited states.

Surface/Solid
The SolidState type should include dust and ices and should fulfill the needs in fusion energy research, astrophysics and industrial plasmas. At present only a simple schema for surface has been proposed, including:

- temperature, density, surface charge (all optional)
- description of single- and multi-layer materials
- description of mixed and single materials, with composition and properties

Action: first draft document (*Dave*)
Review the schema

Processes
This section has been re-organized, taking into account the numerical data type provided. Three data types are identified:

- transition probabilities for radiative processes (*Data type*)
- transition probabilities for autoionization (*Data type*)
- cross sections, rate coefficients and reaction rates for collisions (*DataXY type*)

Radiative processes and autoionization are separate to reflect their own specificity.

Radiative: Evelyne will review this part, essentially for the molecular radiative processes. Emission and absorption have the same description – distinction is made from the initial and final state. Air broadening and stark broadening will be introduced into the schema later.

Collisions: Denis to review this part

DataXY
Errors can be reported: positive or negative dx and dy. The x units are given in the *xType* (former *arguments*), the y units in the yType (former *collParamType*). Type of errors are systematic, computing, experimental…, to be compared with IVOA.

Actions: review the structure for a better fit with *processes* (ex. units) *(Denis)*

Bibliography
Simple approach is sufficient for our needs.

3. Processes

The element *process* is now complex with two optional elements. A user-defined element and a list of codes (for example: dissociative recombination will have 2 values: dissociation and recombination).

A proposed classification of surface processes was described in a document for the IAEA ALADDIN database. From this document Dave will propose a list of processes for surface interactions.

For the collisions part, Dave will simplify the list of processes proposed in the IAEA document "Classification of Processes".

4. Namespace, Registry, DAL, Schema Name

These topics are still pending – a decision was made to postpone further discussion to the next meeting. As a reminder, these topics should be highlighted through a recognized international group. The problem is how to create and organize such an international body.

Find a logo and a proper name, since AMDML does not refer to PSI.

Actions:

- Web site: *Yuri* to develop a home page at NIST
- Everyone to make proposals for a name

Information on our meetings is posted on the AMD Unit web site.

Considered important to release a first version of the XML schema (version 1.0) by June 2008.

5. Draft Document

Yuri will prepare an overview document of the schema by the end of March. A draft document for each container, with the element description, is also recommended. The IOVA document presentation will be used (e.g. M-L document).

Action: Marie-Lise to send a Latex template

6. Collaboration

Yuri reported continued strong interest from Russia and China to participate in the project. The situation with Korea is unclear. China and Russia will be invited to our next meeting.

7. ICAMDATA, China, October 2008

A request was made to Graeme Lister to include an invited talk on the schema at the Beijing ICAMDATA. Posters on different applications will be presented.

8. Working Method, Working Plan and Time Schedule

Yuri opened a Google group, AMDML (http://groups.google.com/group/amdml). Every change in the schema should be posted at this group and reviewed by everyone, increasing the traceability of each modification. All containers should be uploaded by the end of the year

Proposed actions and milestones are:

- 31 Dec 2007 Version 0.1, posted on AMDML Google group *(everyone)*
- January 2008 Draft documentation for solid schema and bibliography *(Dave)*
- Mid-Feb 2008 XML application at NIST, ASD lines *(Yuri)*
- Mid-Feb 2008 XML application at IAEA, ALADDIN collisions *(Denis)*
- End March XML Application at Observatoire Meudon, BASECOL (spectroscopy and collisions) *(Marie-Lise)*
- End March Overview documentation *(Yuri)*
- XML meeting, 15-16 May 2008, Vienna
- **Version 1.0 of the XML schema, June 2008**
- XML meeting, 23-24 October 2008, China
- ICAMDATA, 28-31 October 2008, China

9. Future Meetings

- 15-16 May 2008, IAEA Vienna
- 23-24 October 2008, before the ICAMDATA, China

10. Data Consistency (from previous report INDC(NDS)-0520)

Types:

- Types may be labelled with *extension* when referring to a global type. Consistency must be checked. Data types in transitionProbabilities are with the extension label (oscillatorStrength) and without (lineStrength).
  Action for each container:

Names:

- name syntax: first letter lower case, uppercase for each new word (*atomicState*). Types begin with an uppercase in a text document.
- Full name or abbreviation: *oscStrength* vs *oscillatorStrength*

Structure:

- All groups must be defined, with a begin tag and an end tag:
  particles → *particle* → ....
- Unbounded and bounded tags appear in the following order: unbounded and bounded

Appendix

IAEA Consultants' Meeting: XML Schema for Atomic and Molecular Data

6-7 December 2007, building-A, floor-6, room-601, Observatoire de Paris, Paris, France

---

MEETING AGENDA

Thursday, 6 December

09:15 Opening, Evelyne Roueff

Morning

1. Last meeting minutes review and adoption of the agenda
2. Report on:
   IAEA CRP on surface erosion: 1st RCM, 17-19 October 2007
   IAEA DCN Meeting, 3-5 October 2007
3. Review of present "AMDML" status
   - Overview *Yuri*
   - Atomic data *Yuri*
   - Solid *Dave*
   - Molecular data *Marie-Lise*
   - Single reactant and multi reactants processes *Denis, Bob*
   - Biblio *Dave*
   - Functions, methods *Denis*
   - Types and attributes *Yuri*
   - Actions from last meetings

Afternoon

4. Classification of processes
   a. Atomic and molecular collisions (document "Classification of processes")
   b. Particle surface interactions, new document (*Denis*)
5. Comments, corrections, new developments and unresolved issues from last meeting:
   - element "dataXY"
   - list of data types: cross sections, reaction rates…
   - structure consistency, rules: draft document *(Denis)*
   - units, link with UNITSML
   - isoelectronic series, isonuclear sequences, quantum defects, average atom
6. AMDML applications

Estimated end of first day 17h

Friday, April 3

09:00 Meeting Continued

Morning

7. Web site: IAEA *(Denis),* NIST *(Yuri)*
8. Draft document
9. Space name for "containers"*(pending item)*
10. Data Access Layer *(pending item)*
11. AMDML web registry *(pending item)*

Afternoon

12. Collaboration
13. Work plan and milestones
    a. Version 1.0
    b. ICAMDATA "AMD/PSI data Exchange", spring 2008
    c. Dates for next meetings

Estimated closing time 16h

Nuclear Data Section
International Atomic Energy Agency
P.O. Box 100
A-1400 Vienna
Austria

e-mail: services@iaeand.iaea.org
fax: (43-1) 26007
telephone: (43-1) 2600-21710
Web: http://www-nds.iaea.org