J. Plasma Fusion Res. SERIES, Vol.7 (2006) 338-342

# Development of the Atomic and Molecular Data Markup Language for Internet Data Exchange

RALCHENKO Yuri, CLARK Robert E.H.[1], HUMBERT Denis[1],
SCHULTZ David R.[2], KATO Takako[3] and RHEE Yong Joo[4]

*National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD 20899-8422, USA*

[1] *IAEA, P.O. Box 100, Wagramerstrasse 5, A-1400 Vienna, Austria*

[2] *Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6372, USA*

[3] *National Institute for Fusion Science, Toki 509-5292, Japan*

[4] *KAERI, 150 Duck-Jin-Dong, Yusong-Ku, Taejon 305-353, Korea*

(Received: 5 October 2004 / Accepted: 10 January 2006)

**Abstract**

Accelerated development of the Internet technologies, including those relevant to the atomic and molecular physics, poses new requirements for the proper communication between computers, users and applications. To this end, a new standard for atomic and molecular data exchange that would reflect the recent achievements in this field becomes a necessity. We report here on development of the Atomic and Molecular Data Markup Language (AMDML) that is based on eXtensible Markup Language (XML). The present version of the AMDML Schema covers atomic spectroscopic data as well as the electron-impact collisions.

**Keywords:**

atomic and molecular data, Internet, data exchange, XML

## 1. Introduction

The amount and complexity of atomic and molecular data exchanged over the Internet necessitate development of new standards and verification tools that could guide the users in these activities. The previous standard for atomic and molecular data exchange, ALADDIN (from **A** **L**abeled **A**tomic **D**ata **In**terface), was developed in order to provide a standard and flexible format and interface for the exchange and management of atomic, molecular, and plasma-surface interaction and material properties data of interest to fusion research [1]. It was originally designed by R. Hulse at the Princeton Plasma Physics Laboratory and then adopted by the International Atomic Energy Agency (IAEA) and the Atomic and Molecular Data and Plasma-Material Interaction (A + M/PMI) Data Centre Network (DCN) for the exchange of data since 1988. ALADDIN consists of a set of Fortran-77 programs that provides for interactive data search, terminal display, and data manipulation capabilities. ALADDIN was designed on very general principles, which made it applicable to a wide variety of databases and their interfacing to a similar variety to data application codes.

Development of ALADDIN was an important step in providing users with modern strategies and tools for data exchange, and its significance has quickly become obvious for researchers working in fusion applications, plasma physics and diagnostics, and other disciplines. Nonetheless, an extremely fast advance of computer technologies, as well as the Internet in general and World Wide Web (WWW) in particular, calls for exploration and development of new approaches and techniques in atomic and molecular (A + M) data exchange, which would reflect the recent achievements in the aforementioned fields. The present paper describes our approach to development of the new A+M data exchange standards.

## 2. XML and A + M data exchange

The eXtensible Markup Language (XML) [2] was originally designed to meet the challenges of large-scale electronic publishing. Nowadays XML is also playing an increasingly important role in the exchange of a wide variety of structured data on the Web and elsewhere and, in fact, has become a cross-platform, extensible, and text-based standard for representing data. As a metalanguage, XML is used to develop new, more specific, markup languages for very different applications and realms of human activities such as, *e.g.*, government,



*Corresponding author's e-mail: yuri.ralchenko@nist.gov*

bibliography, physics, finances, etc. In addition to general rules on how to create documents within a specific markup language, there are validation tools that check conformance of documents to the relevant standards.

Among various XML methods and techniques, the XML Schema language (see, *e.g.*, [3]) is one of the most powerful and flexible ones. A schema can be used to validate such important features as the structure and order of the markup elements and attributes, and order of the markup elements and attributes, the data values of attributes and elements based on types, ranges, enumerations and pattern matching, and the uniqueness of values in an instance, to name a few. Besides, schemas can provide default and fixed values for the data elements and attributes as well as documentation about the data in an XML document. These and other important features make the XML Schema very appropriate for our purposes.

There have recently been some attempts to develop either very general specifications for scientific data, *e.g.*, eXtensible Data Format [4], or very specific schemas, *e.g.*, SpectroML for molecular spectrometry data [5], neither of which can adequately serve the purpose of the A + M data exchange. Therefore, in October 2003, the IAEA DCN decided to form a working group on development of the XML-Schema-based Atomic and Molecular Data Markup Language (AMDML) that is supposed to replace ALADDIN as the standard for the A + M data exchange.

## 3. AMDML schema

The present development version of ADMDL Schema covers atomic spectroscopic data and electron-impact collisions with atoms and ions while the heavy-particle collisions will be dealt with in the next versions. An XML data file must have a mandatory block describing the atomic states and several optional blocks describing various radiative and collisional types of transitions between those states. The schema allows different levels of generalization for presentation of the atomic states, *e.g.*, levels, terms, configurations, etc., which are complemented by detailed information on state properties including energies and various quantum numbers. The radiative block provides data on transition probabilities, transition type, etc. The collisional blocks may contain both original experimental and/or theoretical data and fits of the original data. For the former, the fitting function expression and fitting parameters are provided as well. In all cases, the numerical data may be supplemented by accuracy estimates and relevant bibliographic references and/or comments. An example of an AMDML-valid data file for radiative and collisional transitions between the $2s^2\ {}^1S$ and $2s2p\ {}^1P^o$ terms in C III is presented in the Appendix. The draft version of the AMDML Schema will be shortly distributed on the DCN web sites for user comments and suggestions.

AMDML will offer a number of advantages over ALADDIN. One of the most important is the possibility of automatic data validation which can be exemplified by several entries in the C III file in Appendix. For instance, the spectroscopic charge of an ion given by the `spectrCharge` parameter is allowed to take only positive integer values, while the `parity` parameter can only take values "−1" or "1". For the complex data type `orbitalMomentum`, its character component must only be a letter while the numerical component given by the `l_Value` parameter must be a non-negative integer. The AMDML data validation approach adds solid physical considerations to the contents of the data file thereby correcting possible errors and misprints. Also, the AMDML data files may include an explicit expression for the fitting function with a detailed identification of the fitting parameters. This would allow a direct generation of fitted data by applications.

## 4. Conclusions

The importance of the A + M data exchange via the Internet is widely recognized. We presented here the first attempt to develop a new standard which is based on the most advanced Internet technologies for structured data exchange.

## References

[1] R.A. Hulse, in: *Atomic Processes in Plasmas*, ed. by Y.-K. Kim and R.C. Elton, Gaithersburg, MD 1989 (AIP 1990), p.63.

[2] URL http://www.w3.org/XML/.

[3] P. Walmsley, *Definitive XML Schema* (Prentice Hall PTR, Upper Saddle River, NJ, 2002).

[4] URL http://xml.gsfc.nasa.gov/XDF.

[5] M. A. Rühl *et al*, Journal of the Association for Laboratory Automation, **6**, 76 (2001).

## APPENDIX

```
<?xml version="1.0" encoding="UTF-8"?>
<!--Sample XML file generated by XMLSPY v2004 rel. 4 U (http://www.xmlspy.com)-->
<atomicTransition xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="C:\temp\adml1.xsd">
 <atomicStates>
  <atomicState>
   <element nuclearCharge="6">C</element>
   <spectrCharge>3</spectrCharge>
   <stateID>ID1</stateID>
   <ionCharge>2</ionCharge>
   <parity>1</parity>
   <energy units="eV">
    <value>0</value>
   </energy>
   <ionEnergy units="eV">
    <value>47.864</value>
   </ionEnergy>
   <state stateType="term">
    <component>
     <configuration>
      <core>
       <elementCore nuclearCharge="2">He</elementCore>
       <term>
        <LS>
         <L>
          <l>S</l>
          <l_Value>0</l_Value>
         </L>
         <S>0</S>
        </LS>
       </term>
      </core>
      <shells>
       <shell>
        <principalQN>2</principalQN>
        <orbitalMomentum>
         <l>s</l>
         <l_Value>0</l_Value>
        </orbitalMomentum>
        <numbElectrons>2</numbElectrons>
        <parity>1</parity>
        <shellTerm>
         <LS>
          <L>
           <l>S</l>
           <l_Value>0</l_Value>
          </L>
          <S>0</S>
         </LS>
        </shellTerm>
       </shell>
      </shells>
     </configuration>
     <term>
      <LS>
       <L>
```

```
      <l>S</l>
      <l_Value>0</l_Value>
     </L>
     <S>0</S>
    </LS>
   </term>
  </component>
 </state>
 <statWeight>1</statWeight>
</atomicState>
<atomicState>
 <element nuclearCharge="6">C</element>
 <spectrCharge>3</spectrCharge>
 <stateID>ID2</stateID>
 <ionCharge>2</ionCharge>
 <parity>-1</parity>
 <energy units="eV">
  <value>12.690</value>
 </energy>
 <state stateType="term">
  <component>
   <configuration>
    <core>
     <elementCore nuclearCharge="2">He</elementCore>
     <term>
      <LS>
       <L>
        <l>S</l>
        <l_Value>0</l_Value>
       </L>
       <S>0</S>
      </LS>
     </term>
    </core>
    <shells>
     <shell>
      <principalQN>2</principalQN>
      <orbitalMomentum>
       <l>s</l>
       <l_Value>0</l_Value>
      </orbitalMomentum>
      <numbElectrons>1</numbElectrons>
      <parity>1</parity>
      <shellTerm>
       <LS>
        <L>
         <l>s</l>
         <l_Value>0</l_Value>
        </L>
        <S>0.5</S>
       </LS>
      </shellTerm>
     </shell>
     <shell>
      <principalQN>2</principalQN>
      <orbitalMomentum>
       <l>p</l>
       <l_Value>1</l_Value>
      </orbitalMomentum>
      <numbElectrons>1</numbElectrons>
      <parity>-1</parity>
      <shellTerm>
       <LS>
        <L>
         <l>p</l>
         <l_Value>1</l_Value>
        </L>
        <S>0.5</S>
       </LS>
      </shellTerm>
     </shell>
    </shells>
   </configuration>
   <term>
    <LS>
     <L>
      <l>p</l>
      <l_Value>1</l_Value>
     </L>
     <S>0</S>
    </LS>
   </term>
  </component>
 </state>
 <statWeight>3</statWeight>
</atomicState>
</atomicStates>
<radData>
 <transition>
  <initialStateID>ID2</initialStateID>
  <finalStateID>ID1</finalStateID>
  <transitionProbability>
   <oscStrength>
    <value>0.7586</value>
   </oscStrength>
  </transitionProbability>
  <wavelength lineSource="Ritz" units="Angstrom">
   <value>977.020</value>
  </wavelength>
 </transition>
</radData>
<collData>
 <transition>
  <process>collExcitation</process>
  <projectile>
   <electron/>
  </projectile>
  <initialStateID>ID1</initialStateID>
  <finalStateID>ID2</finalStateID>
  <energyUnits>eV</energyUnits>
  <threshold units="eV">
   <value>12.690</value>
  </threshold>
  <dataType>CrossSection</dataType>
  <collDataUnits>cm^2</collDataUnits>
  <data>
   <fitData>
    <fitValidityLimits>
```

```
      <lowerLimit units="eV">
       <value>12.690</value>
      </lowerLimit>
     </fitValidityLimits>
     <fittingFunction>
      <name>Function1</name>
      <expression>y = A/X + B/X^2 + C/X^3 + D/X^4+E*ln(X)/X</expression>
      <indepVariable>y</indepVariable>
      <depVariable comment="energy in threshold units">X</depVariable>
      <parameters>
       <parameter name="A">
        <paramValue>-1.974e-16</paramValue>
       </parameter>
       <parameter name="B">
        <paramValue>8.347e-16</paramValue>
       </parameter>
       <parameter name="C">
        <paramValue>-3.640e-16</paramValue>
       </parameter>
       <parameter name="D">
        <paramValue>5.952e-17</paramValue>
       </parameter>
       <parameter name="E">
        <paramValue>4.539e-16</paramValue>
       </parameter>
      </parameters>
     </fittingFunction>
     <comments>This function provides correct asymptotics for optically allowed transitions.
Original Data were calculated with the Coulomb-Born-exchange method</comments>
    </fitData>
   </data>
  </transition>
 </collData>
</atomicTransition>
```