

Design of System for Atomic, Molecular and Plasma-Material Interaction Data Exchange with XSAMS

Jun-Hyoung Park^{1,1}, Mi-Young Song¹ and Jung-Sik Yoon¹

¹*Plasma Technology Research Center,
National Fusion Research Institute,
113 Gwahangno, Yuseong-gu, Daejeon, Korea
{pjh1126, mysong, jsyoon}@nfri.re.kr*

Abstract

A+M / PMI (Atomic, molecular and plasma-material interaction data) is important data in plasma physics and engineering part. It will give data reliability to be used in the modeling, simulation for plasma processing. New generated data by various studies is required to exchange for getting reliability of data and to be used in other experiments. Also, ALADDIN and AMBDAS of IAEA to provide a numerical data and bibliographic data has not updated for the past few years. As a result, Researchers have a difficulty in data acquisition to be used in research. We are here concerned with the design of system for A+M/PMI data exchange by XSAMS made from XML. The system provides the numerical data and bibliographic data (including DOI information), as well as data exchange between researchers for validation of data using XSAMS. XSAMS is a XML schema aimed at describing atomic, molecular and particle surface interaction data in distributed databases around the world which is created by IAEA, NIST, etc.

Keywords: *Atomic, molecular and plasma-material-interaction data; data exchange; XML schema*

1. Introduction

Advanced developments in computer technologies offer exciting opportunities for new distribution tools and applications in various fields of physics. A+M/PMI data are of very importance across a wide range of applications such as fusion, astrophysics, chemistry, etc.

A+M/PMI data include large collisional and spectroscopic data but recently, fusion energy development has begun to approach demonstration and engineering for practical reactors. For this reason, data research has progressed for similar data for an expanded range of species and in new energy regimes. This data are evaluated for use in the plasma facing components and for such applications as in the divertor. In addition, a significant need for materials data and plasma-material interaction properties has arisen, and planning for the construction of ITER has served to focus international efforts to coordinated the evaluation of the existing data and the production of new data.

However, reliable data among the A+M/PMI data are more difficult to find because it includes data from not trusted sources and not objective data more than enough. The enterprises and institutes have researched and developed data to secure the reliability. Also exchanging of data by enterprises and institutes will be significant research activities for qualitative improvement in research.

Therefore, since 2003, the A+M Data Unit in the IAEA initiated within the collaborative efforts of the DCN (Data Centre Network) a new standard for atomic, molecular and particle surface interaction data exchange based on XML. As we have said above, it is named XSAMS.

This paper is about the design of integrated information system for reliable A+M/PMI data exchange that has been used or is proposed in the atomic and molecular physics community using XSAMS. It has a function to register, manage and exchange A+M/PMI data available for institutes and enterprises. We will find overview, function and each process of system in this paper. Also we will find system architecture, design of database structure and application of XSAMS Schema. Finally, we find out the meaning and complement of system.

Before going on with question of system design, it should be mentioned that related study and other similar systems. We shall now examine in more detail.

2. Related Works

2.1. XSAMS

XSAMS (XML Schema for Atoms Molecules and Solids) is a XML schema for exchange of A+M data which is created by IAEA, NIST, etc. It has been developed for atomic and molecular physicists who need data. The goal of XSAMS are to develop a complete set of rules (XML tags, document structure, relations, etc.) for description of atomic and molecular, and later plasma-material interaction, data sets [1-4].

XSAMS employs the eXtensible Markup Language (XML) that was developed to facilitate sharing of heterogeneous data across different systems, particularly systems connected to the internet. XML was initially designed for the exchange of information electronic “documents”. Assorted by a set of supportive technologies for data presentation, transformation, querying and validation, XML is becoming the standard format for data exchange among distributed applications components or co-operating applications. The use of XML for information interchange among different enterprises and organizations evokes the need for a common schema that the information must follow [5, 6].

2.2. Typically Similar Database System

2.2.1. BASECOL

The BASECOL database was addressed initially by the Basemol group and a list of fundamental and urgent molecules was set. It provides collisional rotational and ro-vibrational excitation rate coefficients for molecules of astrophysical interest, the molecular energy levels and a complete description of the chain of errors. It gives information on current calculations and is designed as a reference database to be used by both astrophysicists and physicists [7, 8].

2.2.2. NIST ASD

The Atomic Spectra Database (ASD) is the most extensive and comprehensive numerical atomic database at NIST. It contains critically evaluated NIST data for radiative transition and energy levels in atoms and atomic ions. Following general principles for database at NIST only published and evaluated data are included and accuracy must be reported. The data are consistent between spectral lines and atomic levels, which are both in the ASD. Compiling new data involves a critical evaluation of the literature, the assessment and possibly changes of line identifications, the optimization of energy levels based on multiple sources, the assignment of accuracies, and internal NIST quality review. In addition to the ASD contains fully integrated with the bibliographical atomic database at NIST. The content has been enriched with direct HTML link to online papers through Digital Object Identifiers [4, 9-11].

2.2.3. VAMDC VALD

The VAMDC (Virtual Atomic and Molecular Data Center) is an EU-FP7 e-infrastructure project devoted to building a common electronic infrastructure for the exchange and distribution of atomic and molecular data. VALD (Vienna Atomic Line Database) as one part of VAMDC platform has developed into a well-known resource of atomic data for spectroscopy particularly in astrophysics. This particularly relates to the data contents where new sets of atomic data for both precision spectroscopy as well as opacity calculations have been included. Also an intended outcome concerned with molecular spectroscopy microwaves and ultrawaves [12-14].

2.2.4. IAEA ALADDIN

ALADDIN is the online numerical atomic physics database by the IAEA Atomic and Molecular Data Unit, and it providing atomic, molecular and plasma-material interaction data of fusion research. The contents are largely based on IAEA Coordinated Research Projects. The data are expected to be the best available ones at the time that they are first entered into the database, but the coverage is not at all comprehensive and older data may not be the best that is available at present. In order to provide convenient access to many numerical databases the A+M Data Unit maintains the GENIE search engine. It provides access at this time to 8 databases for radiative properties and 4 databases for electron-impact collisional process, also including heavy particle collisions and molecular processes, and bibliographical data [15-17].

However, ALADDIN data format was not generally adopted by plasma modelers who rather invented data format specific to their own source codes, Also, as more and more data are available through web interface, a common exchange format for the data transfer through web applications gained attention and the XSAMS project was started by DCN members and coordinated through IAEA Consultants' meeting since 2003. Although, the XSAMS may not be mature enough yet to be implemented. Nevertheless, it is good time to start implementation even for limited data sets.

Therefore, we propose design of integrated information system for A+M/PMI data exchange using XSAMS and contribution of numerical and bibliographical data sets to ALADDIN, AMBDAS of IAEA.

3. Design of System

3.1. Purpose

A+M/PMI data in system have been collected and assessed in various sources which support a wide range of physics in applied research and industrial development. Many of them have been built to serve specific needs. These heterogeneous data have been collected in different formats and with varied degree of completeness.

Therefore, the main purpose of the system is to

- apply and develop exchange format extend XSAMS for interoperability of resources on A+M/PMI data
- allow easy access to data resources
- query those resources with dedicated protocols and query languages
- create a safe environment for publishing the latest sets of A+M/PMI data
- connect data producers more closely with data users (modelers, simulators, evaluators, etc.)

The key benefits expected from using systems are the possibilities to

- search any type of A+M/PMI data with one click
- have common formal access to the published data
- allow cross matching of different data set
- allow wide cross to latest published data

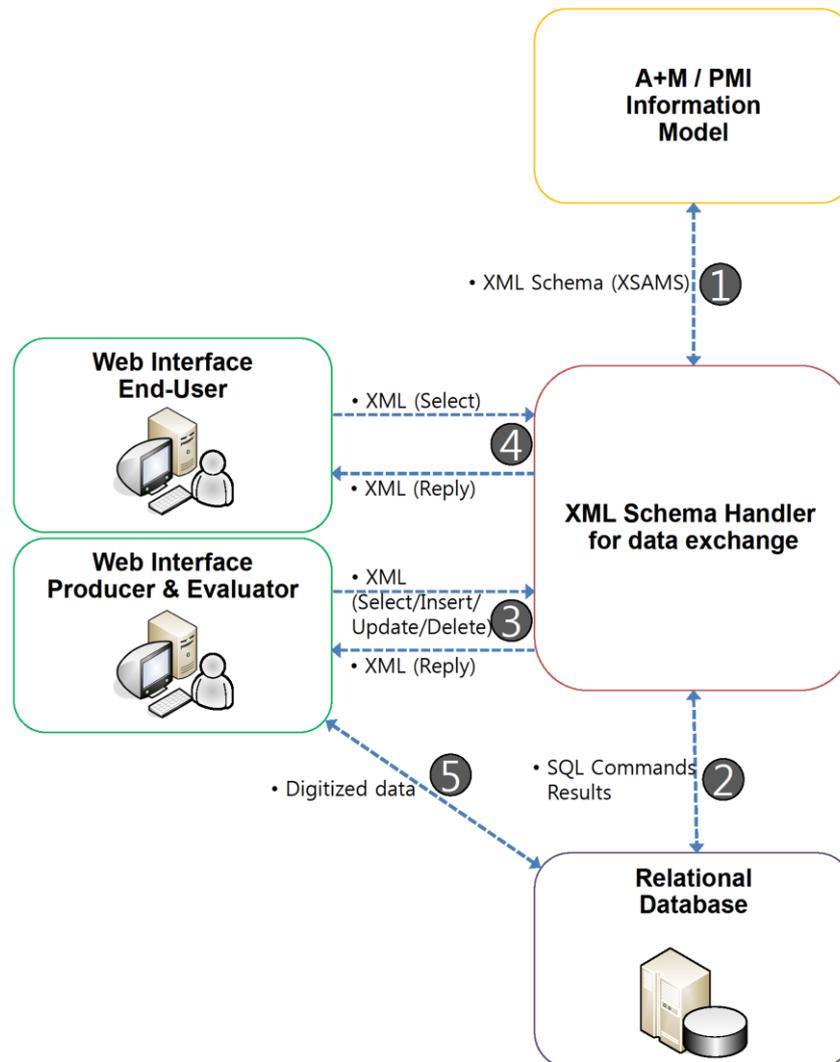


Figure 1. Information Flow

3.2. Information Flow

The information flow of A+M / PMI data system for data exchange is presented in Figure 1. The information flow among the various components is described in the following.

- ① The administrator of the A+M/PMI information model, analyses process results to an XML schema file. This file is used as a validating mechanism for the XML documents the will be created.
- ② The XML schema handler module receives the XML schema file and automatically generates the database. The module receives XML documents created during ③ and ④ and converts them to SQL commands. It also retrieves the query results and

constructs the XML reply documents, which are sent either to the producer and evaluator or to the web users.

- ③ The producer and evaluator can insert, update, delete and retrieve information from the database and store the digitized reference file into the database.
- ④ The web-users retrieve information from the database. User queries are translated to the appropriate XML documents before being sent to the XML schema handler module.
- ⑤ The producer and evaluator confirm the digitized reference descriptions with real numerical data or copies and provide all the data information related to each property object.

3.3. System Architecture

The system architecture is presented in the following Figure 2. The included components and others are described in the following paragraphs.

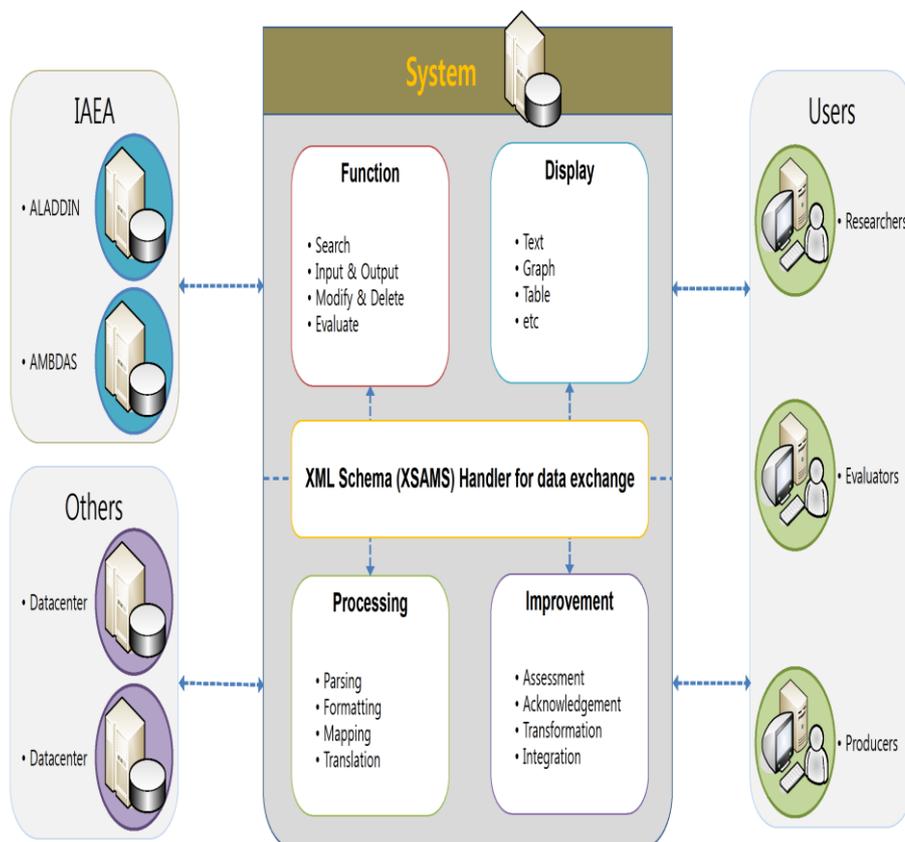


Figure 2. System Architecture

3.3.3. XML Schema (XSAMS) Handler

XML schema handler concerns the analysis of the XML schema file (XSAMS) that contains information about the structure of the exchanged XML documents. This information is used to generate the relational database structures. This occurs in the initialization step of the module. It also parses XML documents and comprises the

appropriate SQL commands that processed by the database. It takes database results and formulates valid XML documents as a reply to user queries.

3.3.3. Users

The system users are divided into two categories: professional users, such as producer and evaluator, who access the A+M/PMI data for generate or assess the new valid data and non-professional users who search the information using a web. Their queries are converted into proper XML documents and set to the XML schema handler. The retrieved information is sent back to the application, as an XML document and is presented to the users.

3.3.4. Database

In an effort to keep the system generic, we addressed ANSI SQL commands to the database so that it works with any relational databases. All the tables and relationships are created by the XML schema handler according to the XSAMS structure. Also all applications communicate with each other and the database using XML document.

4. Conclusions and Future Works

XML was used as standard of data exchange was used in transmission, which resolved the problems of data mapping and data transfer among heterogeneous data sources. XSAMS (XML Schema for Atoms Molecules and Solids) is a XML schema for exchange of A+M data which is created by IAEA, NIST, etc. It has been developed for atomic and molecular physicists who need data. Currently, we are modifying the XML schema to include most of plasma chemistry data such as radiative, non-radiative etc.

This article has attempted to sketch out the main characteristics of integrated information system for atomic, molecular and plasma-material interaction data exchange using XSAMS. This is focus on background and difficulty of heterogeneous data exchange among sources in web environment. In general, it is believed that further implementation with the methods outlined is worthwhile.

References

- [1] <http://www-amdis.iaea.org/xsams/>.
- [2] R. Yuri, "Development of the Atomic and Molecular Data Markup Language for Internet Data Exchange", Plasma Fusion Res. SERIES, vol. 7, (2006), pp. 338-342.
- [3] Y. Ralchenko, "Development new standards for exchange of atomic and molecular data", ICAMDATA-2008 : 6th International Conference on Atomic and Molecular Data and Their Applications. AIP Conference Proceedings, vol. 1125, no. 1, (2009), pp. 207-216.
- [4] Summary Report of an IAEA Consultants' Meeting on "XSAMS: XML Schema for Atoms, Molecules, Solids", IAEA Publication INDC(NDS)-0570, (2010) March 24-26.
- [5] I. Varlamis and M. Vazirgiannis, "Bridging XML-Schema and relational database. A system for generating and manipulating relational databases using XML documents", Proceedings of the 2001 ACM Symposium on Document engineering, (2001), pp. 105-114.
- [6] B. Thuraisingham, "XML databases and the semantic web", CRC Press, (2002).
- [7] <http://basecol.obspm.fr>.
- [8] D. Marie-Lise, "Ro-vibrational Collisional Excitation Database: BASECOL", Plasma Fusion Res. SERIES, vol. 7, (2006), pp. 356-357.
- [9] <http://physics.nist.gov/asd3>.
- [10] Y. Ralchenko, "NIST atomic spectra database", Memorie della Supplementi, vol. 8, (2005), pp. 96-102.
- [11] Y. Ralchenko, "New Generation of the NIST Atomic Spectroscopic Databases", Atomic and Molecular Data and Their Applications. AIP Conference Proceedings, vol. 771, (2005), pp. 276-285.
- [12] <http://www.vamdc.eu/>.
- [13] U. Heiter, "VALD - an atomic and molecular database for astrophysics", Journal of Physics: Conference Series, vol. 130, no. 1, (2008), pp. 1-4.
- [14] F. Kupka and the VAMDC Collaboration, "VAMDC as a Resource for Atomic and Molecular Data and the New Release of VALD", Baltic Astronomy, vol. 20, (2011), pp. 503-510.

- [15] <http://www-amdis.iaea.org/ALADDIN>.
- [16] R. A. Hulse, "The ALADDIN atomic physics database system", AIP Conference Proceedings, vol. 206, (1990), pp. 63-72.
- [17] D. P. Humbert, "IAEA Nuclear Data Section: provision of atomic and nuclear databases for user applications", Proceedings of the 14th International Conference on Radionuclide Metrology and its Applications, vol. 60, (2003), pp. 311-316.

