"Virtual Atomic and Molecular Data Centre" and Astrophysics: Level 2 Release

M. Doronin,[1] M. L. Dubernet,[1,2] N. Walton,[3] N. Mason,[4] G. Rixon,[3]
P. Le Sidaner,[5] S. Schlemmer,[6] N. Piskunov,[7] J. Tennyson,[8] A. Akram,[3]
C. Endres,[6] C. Hill,[8] T. Marquart,[7] L. Nenadovic,[1] K. Smith,[9] and
VAMDC Consortium[10]

[1]*Laboratoire de Physique Moléculaire pour l'Atmosphère et l'Astrophysique, UMR7092 CNRS/INP, Université Pierre et Marie Curie, Case 76, 4 Place Jussieu, 75252 Paris Cedex 05, France*

[2]*Laboratoire Univers et Théories, UMR8102 CNRS/INSU, Observatoire de Paris, Section Meudon, 5 Place Janssen, 92195 Meudon Cedex, France*

[3]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB30HA, UK*

[4]*Open University, Faculty of Science, Walton Hall, Milton Keynes, MK7 6AA, UK*

[5]*Division Informatique de l'Observatoire, VO-Paris Data Centre, UMS2201 CNRS/INSU , Observatoire de Paris, 5 Place Janssen, 92195 Meudon Cedex, France*

[6]*University of Cologne, I. Physikalishes Institut, Zuelpicher Strasse 77, Koeln 50937, Germany*

[7]*Uppsala University, Department of Physics and Astronomy, Lägerhyddsvägen 1, Uppsala 75120, Sweden*

[8]*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*

[9]*School of Mathematics and Physics, Queen's University Belfast, University Road, Belfast BT7 1NN, UK*

[10]`http://www.vamdc.eu`

Abstract. The Virtual Atomic and Molecular Data Centre (VAMDC, `http://www.vamdc.eu/`) is a consortium between groups involved in the generation, evaluation, and use of atomic and molecular data, funded by the European Union. VAMDC aims to build a reliable, open, flexible and interoperable e-science interface to existing atomic and molecular data. The project will cover establishing the core consortium, the development and deployment of the infrastructure and the development of interfaces to the existing atomic and molecular databases. This paper describes the organisation of the project and the achievements at the end of its second year.

1.   Introduction

Atomic and molecular (A&M) data are of critical importance across a wide range of applications such as astrophysics, atmospheric physics, fusion, environmental sciences, combustion chemistry, and in industrial applications from plasmas to lighting. Currently these vital and fundamental A&M data resources are highly fragmented and only available through a variety of highly specialized interfaces, thus limiting the full exploitation of their scientific worth. This in turn hinders research across a wide range of topics including space exploration (the characterization of extrasolar planets, understanding the chemistry of our local solar system and of the wider universe); the study of the terrestrial atmosphere and quantification of climate change; the development of the international fusion programme for energy, and our understanding of radiation damage within biological systems, to give just a few examples.

The Virtual Atomic and Molecular Data Centre (VAMDC: `http://www.vamdc.eu`) is a major new European initiative now building a unified, secure, documented, flexible and interoperable e-science environment-based interface to existing A&M data. VAMDC combines the expertise of existing A&M databases, data producers and service providers with the specific aim of creating an infrastructure that is easily tuned to the requirements of a wide variety of users in academic, governmental, industrial or public communities. The project encompasses the construction of the core consortium, the development and deployment of the infrastructure and the development of interfaces to existing A&M databases. VAMDC partners (Dubernet et al. 2010) are responsible for many of the world's major A&M data resources.

The paper describes the current 'Level 2' service deployment of the VAMDC data infrastructure across a wide range of VAMDC partner provided resources, outlines our objectives and describes our softwares in connection to Astrophysics.

2.   A Level-2 Release

A level-n release is a combined release of: standards for data access, software to prepare the databases and implement the web services, VAMDC nodes, each containing a database and web service following the standards, a registry of the services, a web portal as user interface (UI) for the system, a set of tools following the standards.
A preliminary Level-1 release (Rixon et al. 2011) was made in September 2010 including some prototype services deployed on 4 databases CHIANTI (Dere et al. 2009), VALD (Heiter et al. 2008), CDMS (Müller et al. 2005), BASECOL (Dubernet et al. 2006), HITRAN (Rothman et al. 2009).

Level-2 release in September 2011 is based on standards for data access version 11.05 that included the VAMDC-XSAMS XML schema version 0.2, the query language version 11.05, the dictionaries version 11.05, the registry standards version 11.05, the access protocol VAMDC-TAP version 11.05 and the node softwares in Python version 11.05. The Level-2 release deploys an infrastructure including 13 databases, the registry to find those resources, a web portal as UI for the system, a tool called TAPValidator to validate VAMDC-TAP services (the application supports both Graphical user interfaces and command-Line mode operations) and the SPECTCOL tool designed to handle and cross-federate spectroscopic and collisional data for use in the analysis of spectra from the interstellar medium (Dubernet, M. L., Nenadovic, L. and Doronin, M.

2012). The VAMDC website on standards and softwares provide the documentation for the standards and softwares as well as the different softwares for download.

2.1. Standards

- Registries: The registries present the yellow pages of the infrastructure, i.e. a single point that allows discovery of all resources. They are based on IVOA Astrogrid registries, with a special identifier ivo://vamdc/ and with an extended service registration supporting the query of keywords, registration of software/standards version, and sample queries, suitable for service monitoring.

- Access Protocol: VAMDC-TAP. The access protocol is based on the synchronized version of the IVOA Table Access Protocol with additional or modified features: it includes a HTTP HEAD request to estimate the result of the query, it outputs VAMDC-XSAMS, implements VSS# query language (VSS1 for Level-2 release) and use data compression during the transfer.

- Data Format: XSAMS and VAMDC-XSAMS. The VAMDC activities are based on the use of an XML schema XSAMS, which has been developed under the auspices of the IAEA and aims to describe atomic, molecular and particle-surface interaction data in distributed databases around the world. XSAMS is currently being improved and extended within the VAMDC consortium and successive VAMDC version of XSAMS are called VAMDC-XSAMS. The VAMDC website on standards provide the modified VAMDC-XSAMS schema version 0.2, its documentation and the change log document that follows modifications from version 0.1 of XSAMS.

- Query Language: VSS# is a small sub-set of SQL with query keywords (called Restrictables) derived from the elements of the VAMDC-XSAMS schema. In Level-2 release VSS# is called VSS1, allows to query one molecule at a time and therefore does support correct queries of reactive and collisional databases. VSS2, to be released in December 2011, allows more sophisticated queries. The dictionaries used for selecting query keywords are provided on the VAMDC standards website.

2.2. Implementation

Data publishing in VAMDC is realized by setting up a VAMDC node, consisting of a database engine, a web server and a framework/interface providing the connection to the central VAMDC server (registry) and allowing standardized data queries. Two software packages which provide this functionality have been developed, implemented and tested for Level-2 release:

- The first one, called "NodeSoftware", is based on the Python framework Django. It allows a new data provider to build a new VAMDC node from scratch, following the released standards. This NodeSoftware package has been deployed for Level 2 for 11 of the core VAMDC databases (Dubernet et al. 2010). The implemented databases can be visualised from the VAMDC portal (`http://portal.vamdc.eu/`). In addition, a completely new VAMDC node was created in Lund from data distributed in various publications and unpublished data files.

- The second one is a modularized Java implementation compliant with VAMDC standards in order to publish atomic and molecular data. This software package has been deployed for the BASECOL and KIDA (⟨`http://kida.obs.ubordeaux1.fr`⟩) databases, and had the advantage of relying on more mature technologies. The tool understands the standard queries from VAMDC and returns data in VAMDC-XSAMS compliant format. It is a full alternative to the Python/Django node software platform.

2.3. User Interface = VAMDC Portal

The VAMDC Portal is in continuous evolution and the entry can be found on `http://portal.vamdc.eu/`. Several UI are available: a frozen Level-2 version, used for the European Community Review in September 2011, that supports VSS1 only (VSS2 is a non working option) and that will be shut down after few months, a Stable Version (albeit with known bugs) that evolves slowly once major changes have been agreed upon, and the Test Version aimed at the VAMDC collaboration for testing both the UI and their databases deployments. All these User Interfaces are public, but users should be aware that they are in a development phase up till June 2012.

References

Dere, K. P., Landi, E., Young, P. R., Del Zanna, G., Landini, M., & Mason, H. E. 2009, A&A, 498, 915

Dubernet, M.-L., Grosjean, A., Flower, D., Roueff, E., Daniel, F., Moreau, N., & Debray, B. 2006, Journal of Plasma Research Series, 7, 356

Dubernet, M. L., et al. 2010, JQSRT, 111, 2151

Dubernet, M. L., Nenadovic, L. and Doronin, M. 2012, in ADASS XXI, edited by P. Ballester, D. Egret, & N. P. F. Lorente, vol. 461 of ASP Conf. Ser., 335

Heiter, U., et al. 2008, Journal of Physics Conference Series, 130, 012011

Müller, H. S. P., Schlöder, F., Stutzki, J., & Winnewisser, G. 2005, Journal of Molecular Structure, 742, 215

Rixon, G., et al. 2011, in American Institute of Physics Conference Series, edited by A. Bernotas, R. Karazija, & Z. Rudzikas, vol. 1344 of American Institute of Physics Conference Series, 107

Rothman, L. S., Gordon, I. E., & Collaborators 2009, JQSRT, 110, 533