



VAMDC

Virtual Atomic and Molecular Data Centre

D6.3

–

Interoperability Report 2

Version 0.1

Grant agreement no: 239108

Combination of Collaborative Projects & Coordination and Support Actions





Project Information

Project acronym: VAMDC
 Project full title: Virtual Atomic and Molecular Data Centre
 Grant agreement no.: 239108
 Funding scheme: Combination of Collaborative Projects & Coordination and Support Actions
 Project start date: 01/07/2009
 Project duration: 42 months
 Call topic: INFRA-2008-1.2.2 Scientific Data Infrastructure
 Project web sites: <http://www.vamdc.eu>

<http://voparis-twiki.obspm.fr/twiki/bin/view/VAMDC/WebHome>

Consortium:

Beneficiary Number *	Beneficiary name	Beneficiary short name	Country	Date enter project**	Date exit project**
1(coordinator)	Centre National de la Recherche Scientifique	CNRS	France	Month 1	Month 42
2	The Chancellor, Masters and Scholars of the University of Cambridge	CMSUC	UK	Month 1	Month 42
3	University College London	UCL	UK	Month 1	Month 42
4	Open University	OU	UK	Month 1	Month 42
5	Universitaet Wien	UNIVIE	Austria	Month 1	Month 42
6	Uppsala Universitet	UU	Sweden	Month 1	Month 42
7	Universitaet zu Koeln	KOLN	Germany	Month 1	Month 42
8	Istituto Nazionale di Astrofisica	INAF	Italy	Month 1	Month 42
9	Queen's University Belfast	QUB	UK	Month 1	Month 42
10	Astronomska opservatorija	AOB	Serbia	Month 1	Month 42
11	Institute for Spectroscopy RAS	ISLAN	Russian Federation	Month 1	Month 42
12	Russian Federal Nuclear Centre All-Russian Institute of Technical Physics	RFNC-VNIITF	Russian Federation	Month 1	Month 42
13	Institute of Atmospheric Optics	IAO	Russian Federation	Month 1	Month 42
14	Corporacion Parque Tecnologico de Merida	CPTM	Venezuela	Month 1	Month 42
15	Institute of Astronomy of the Russian Academy of Sciences	INASAN	Russian Federation	Month 1	Month 42



This project is funded under “*Combination of Collaborative Projects and Coordination and Support Actions*” Funding Scheme of The Seventh Framework Program of the European Union

Document

Deliverable number: D6.3
Deliverable title: Interoperability Report 2
Due date of deliverable: June 2011
Actual submission date: August 2011
Authors: M.L. Dubernet, F. Kosmala, and the WP6 team
Work Package no.: WP6-JRA1
Work Package title: Interoperability
Work Package leader: CNRS/KOLN
Lead beneficiary: CNRS/KOLN
Dissemination level: PU
Nature: Report
No of pages (incl. cover):

Abstract	The objective of D6.3 is to describe VAMDC Interoperability Report for Period 2. This report corresponds to Activities in WP6: JRA1 “Interoperability”. This report is included in the VAMDC Periodic Report for Period 2.
----------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Versioning and Contribution history

Version	Date	Reason for modification	Modified by
V0.1	June/July 2011	Compilation of Node contributions to WP6	F. Kosmala
V0.1	August 2011	D6.3 Document	M.L. Dubernet

Final Version (v0.1) released by		Circulated to	
Name	Date	Recipient	Date
M.L. Dubernet	24 th August 2011	Mrs Asero	24 th August 2011

Disclaimer: The information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

All rights reserved:

The document is proprietary of the VAMDC consortium members. No copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights.

This document reflects only the authors' view. The European Community is not liable for any use that may be made of the information contained herein

WP6 ACTIVITIES DESCRIPTION

Work package number	6		Start date or starting event:				3			
Work package title	JRA1: Interoperability									
Activity Type	RTD									
Participant id	1	6	7	8	9	10	12	14		
Person-months per beneficiary: (Total = EU + Node Contributions)	36	18	10	5	12	6	3	6		

Table of Content

WP6 activities description.....	5
1. WP6 Objectives as in Annexe I.....	5
2. WP6 Milestones and Deliverables as in Annexe I.....	5
3. WP6 Tasks Description as in Annexe I.....	6
4. WP6 Tasks Plan for Period 2.....	7
5. WP6 Tasks Reports for Period 2.....	9

1. WP6 Objectives as in Annexe I

Define all standards necessary to build an interoperable infrastructure

WP6 Leader is KOLN with co-leader CNRS

2. WP6 Milestones and Deliverables as in Annexe I

Milestones

M6.1	Technical Meetings	WP6	UU	Months 5,10, 16, 22, 28, 34, 40, 42	Minutes. Presentations on internal Website
M6.2	Evaluation of standards releases	WP6	UU	Months 10, 22, 34	

Deliverables

D6.1 Interoperability Plan (PM 3)

D6.2 Interoperability Report to be included in report to the EU – Year 1 (PM 10)

D6.3 Interoperability Report to be included in report to the EU – Year 2 (PM 22)

D6.4 Interoperability Report to be included in report to the EU – Year 3 (PM 34)

D6.5 Final Report of Interoperability to be included in final report to the commission (PM41)

Annual Interoperability Plan revisions included in Revised Annual VAMDC Project Plans – Year 1,2,3

3. WP6 Tasks Description as in Annexe I

WP6 Leader (co)	S. Schlemmer (KOLN)/ M.L. Dubernet (CNRS: LPMAA/LUTH)	
Task Number	Leader	Other Partners
1	M.L. Dubernet/J. Bureau(CNRS:LPMAA/LUTH)	All partners
2	T. Millar (QUB)	All partners
3	N. Piskunov (UU)	CNRS:LPMAA/LUTH
4	M. Doronin (CNRS:LPMAA)	UU
5	S. Schlemmer (KOLN)	All partners

Description of work (possibly broken down into tasks)

Task 1: Data Models and XML Schema (lead by CNRS(1), all JRA1 partners)

The current data models and XML schema are a description of atomic and molecular line lists for use in an astrophysical context and a description of atomic and molecular elementary processes. The documents have been designed by a small number of people, those documents are still in draft mode, do not cover all application fields and have not been discussed extensively among users and producers. These preliminary versions will be completed and extended in order to cover a wider range of species, a wider range of processes and will include the effect of the environment. For now we identify the following extensions:

- inclusion of solid, surface spectroscopy for interstellar medium and planetology
- inclusion of larger molecules such as PAH
- description of atomic and molecular line shapes arising from different sources

In connection to the user & producer board of NA2, more extensions will be considered if necessary. Through the Standards and Processes Committee all standards will be connected to International efforts of standardisation.

Task 2: Dictionaries (lead by QUB(9), all JRA1 partners)

In order to uniquely identify resources we will need to define and build dictionaries both general and specific to applications. At present we identify the following dictionaries:

- single identification of databases and services
- list of conventions (link to IUAPAC and other convention bodies)
- list of names of species (in relation with other fields such as chemistry)
- list of processes and coding of processes
- list of quantum numbers

Other lists will be identified during the course of the project in relation with activities in WP2 (NA2)

Task 3: Access Protocols and Query/Retrieval Languages (lead by UU(6), all JRA1 partners)

We will define protocols retrieving different types of resources: numerical data, libraries, documentation, references. Those protocols will cover asynchronous queries and the retrieval of huge sets of data. In a second step we will design a general query language allowing to access and retrieve any atomic and molecular data.

Task 4: Registries (lead by CNRS(1) with partner 6)

Registries provide a mechanism with which applications can discover and select resources-- e.g. data and services--that are relevant for a particular scientific problem. We will start from the registries defined in the IVOA and see how to adapt and/or extend the documents to our own purpose. In particular we wish to implement ways of finding resources at various levels of granularity.

Task 5: Other Documents (lead by KOLN(7), all JRA1 partners)

Other basic definitions/standardisations might be necessary in order to find/identify resources and documents produced by the IVOA will be assessed in order to keep/adapt to our own needs.

4. WP6 Tasks Plan for Period 2

Year two WP6 JRA1 Interoperability plan

Period: 01/07/2010 – 30/06/2011

WorkPackage: WP6 Interoperability

WorkPackage Leader and co-Leader: M.L. Dubernet (CNRS: LPMAA/LUTH)

/ S. Schlemmer (KOLN)

Participants in the WorkPackage: All Partners

Part 1

Objectives and details for each task in Year 2.

Period 2 will be the period where standards documents will be officially available through a public with regular updates.

Task 1: Data Models and XML schema Documents (CNRS/LPMAA/LUTH)

1.1 Investigation of XSAMS

This part continues while implementation of XSAMS is happening.

Period 1 has seen development of "case-by-case" - Period 2 will have 3 aspects: 2 technical and 1 political

- technical aspects:
 1. continue to improve XSAMS using inputs from VAMDC partners (over period 2)
 2. choose between the different versions of molecular XSAMS: this issue should be solved about November 2010
- political issue:
 1. who should be the reference body for XSAMS (IAEA or VAMDC). This issue should be solved about November 2010

1.2 Schema for Solid Spectroscopy (CNRS/LPG)

The Solid Spectroscopy Data Model (SSDM) will continue its development and structure.

In Period 1 we have added most of the keywords to fully describe solid samples through their layers, materials, constituents and species as well as some types of spectroscopic experiments and their products. And we started the optimization of the structure of SSDM.

In period 2 we will further improve the overall structure of SSDM, complete the band list DM and develop some extensions (natural materials, ...). A meeting is planned to finalize a first stable version.

It is now clear that the SSDM cannot be unified with XSAM(S). So we will study "mapping" between SSDM and XSAMS to be able to make some intercomparison between gas and solid phase data. And make scientific specifications for intercomparison and interoperability tools.

1.3 Schema for PAH (leader: INAF + CNRS/CESR)

Investigation of suitability of XSAMS for PAH data, Organisation of a workshop together with CESR partners, Identification of missing/desirable functionalities, Draft document have been done in Period 1.

Period 2: Inclusion of PAH in XSAMS to be finalized in order to allow start of implementation by December 2010

1.4 Schema for Line Shapes related to pressure broadening (Leader UCL with participants CNRS/GSMA, ICB, IAO, CFA)

Preliminary studies have been done in Period 1

Period 2: Propose a standard after discussion with wide panel of specialists ; discussion whether it should be combined with Atomic line shape standard

1.5 Schema for Atomic Line Shapes (Leader Yuri Ralchenko from NIST with participants AOB, LERMA/CNRS)

Period 2: Propose a standard and discussion with molecular physicists

Task 2: Dictionaries Documents (QUB)

2.1 List of Species (QUB, CNRS/LPMAA and partners)

Names: make decision about standardized names

In Period 1, Studies of InChI was made for identification of species, but some drawbacks were identified

Period 2: continue investigation of InChI in order to uniquely identify species

SSDM will check compatibility with solids and will adopt and integrate this nomenclature

2.2 List of Processes

Period 1: list of processes of IAEA was adopted - Period 2: improve the list if necessary

2.3 List of Conventions

Period 2: Nothing was done - Period 2: to be done if necessary

2.4 List of Quantum Numbers (UCL)

Period 2: a proper list of quantum SHOULD be made and adopted across all schemas

The specific problem of definition and quantification of vibration modes in solids (some types being fully specific) will be addressed for inclusion in SSDM and drawing an equivalence schema with Quantum numbers in XSAMS.

2.5 List of cases (UCL)

In Period 1, list of cases was designed (see report P1)

Period 2: check the list, improve and complete while implementation is carried out -

Task 3: Access Protocols and Query/Retrieval Language Documents (CMSUC)

3.1 TAP-XSAMS

leader of Document : G. Rixon

Period 2 content: refine the existing standard to deal with issues found in implementation.

Add any optional query-languages and formats (e.g. the format for solid-spectroscopy data, if that is to be handled by this kind of service). Note any allowed variations between services implementing this protocol

3.2 SQL based Query

leader of Document: G. Rixon

Period 2 content: formalize the definition of VSS1 currently on the wiki web. Consider a possible VSS2 which allows a greater sub-set of SQL.

Task 4: Registry Documents (UCL/MSSL)

4.1 Registration format for TAP-XSAMS

leader of document: K. Benson

Period 2 content: if variations in features between TAP-XSAMS instances are allowed by the protocol standard, define how these differences are annotated in the registration. Define an XML schema, for a special type of IVOA Capability structure, to hold these annotations. Consider how the data content might be described in these annotations.

5. WP6 Tasks Reports for Period 2

VAMDC Periodic Report P2

Period: 01/07/2010 – 30/06/2011

WorkPackage: WP6 – Interoperability

WorkPackage Leader and co-Leader: M.L. Dubernet (CNRS: LPMAA/LUTH)
/ S. Schlemmer (KOLN)

Participants in the WorkPackage: UCL, UU, CMSUC, CNRS are the most active participants in WP6. Other partners provide some punctual inputs.

Part 1

A summary of progress towards objectives and details for each tasks

Task 1: Data Models and XML schema Documents

Period 1 was a period of discovery of XSAMS by the different partners and of new propositions for improvement of XSAMS. In Period 2 we were meant to decide about how to handle the political and scientific aspects of XSAMS. We were meant to meet requirements in order to describe PAH and line parameters for atoms and molecules.

Studies of connections between the SSDM schema and XSAMS were supposed to be studied in Period 2.

All those goals have been achieved in Period 2.

1° VAMDC has decided to split from IAEA during the course of the project and all improvements are now made within VAMDC consortium. New version of XSAMS is called VAMDC-XSAMS. IAEA accepts changes at a very slow pace and we will converge at one point.

2° Repository has been created (not public outside VAMDC) for the different developpements of VAMDC-XSAMS

- Agreement of the proposed Schema is currently held in the vamdc-working branch of the repository <http://xsams.svn.sourceforge.net/viewvc/xsams/branches/vamdc-working/>
- Having proposals in place for Environments, Lineshapes (ie spectral line broadening), Cross Sections, Partition Functions, Normal Mode description

(CH working with Giacomo Mulas for the TSCD PAH database), Molecular Structure, and Basis Set Expansions; these are fully documented <http://xsams.svn.sourceforge.net/viewvc/xsams/branches/vamdc-working/docs/>

- Other branches exist for UCL, NIST, etc.. in order to make improvements visible before going to vamdc-working branch.

3° Official Releases of VAMDC-XSAMS standards are on <http://www.vamdc.eu/documents/standards/>

4° Work on SSDM

Extensions of SSDM:

- Major developments and restructuration of the SSDM (Solid Spectroscopy Data Model): Samples, matters, species
- Major rebuild of the sample and material section of the SSDM, allowing for more complex samples structures to be described. New tables for external matters database link, allowing the development of this database and its content description with fundamental matters and samples.
- Splitting of the experiment, instrument and laboratory tables. Each now having its own set of tables and parameters, allowing to link together experiment by instrument and laboratories in a more efficient way.
- Upgrade of the species section of SSDM. Check of compatibility with molecular solids and implementation of external keywords to species indexing: Inchi, InchiKey, IUPAC name, CAS registry number. SMILES formalism adopted to describe the structural formula of molecules. New definition of the relational database structure for isotopes and natural composition of atoms and molecules.
- Delivery of the pre-stable version of the core of SSDM (v0.3.4)
- Writing of the dictionary and documentation of SSDM (v0.3.4, 130p.)
- Meeting of the SSDM expert group (23-24/06/2011 Grenoble, 17 people) to finalize the first stable version (v0.4) of SSDM and define units, accuracies and enumerations. Test of the data import on wide types of solid data from various data producers.

The Solid Spectroscopy developments are at

http://ghosst-prod.obs.ujf-grenoble.fr/wiki/index.php/SSDM_data_model

Task 2: Dictionaries Documents

Different dictionaries for species, quantum numbers, cases, processes were meant to be done or improved in Period 2. More work on Inchi and Inchikeys was planned.

Work has been performed on all those aspects.

2.1 Species :

- Inchi is the unique identifier for communication within VAMDC.
- Nevertheless a database of species is currently being built in order to facilitate the users search on the portal. This will allow “auto-suggestion” for search of species.

2.2 A process list has been edited using the process list indicated in version 0.2 of

VAMDC-XSAMS (release [11.05]), and completed. This list will be used at the portal for user's search allowing "auto-suggestion".

2.3 Official Cases and Quantum Numbers are given in documentation of "Data Model" on <http://www.vamdc.eu/documents/standards/>

Task 3: Access Protocols and Query/Retrieval Language Documents

Period 2 planned to work on TAP-XSAMS, formalize VSS1 and define VSS2. Work has been performed on all those aspects. Documents have been edited for VAMDC-TAP and VSS1 in release [11.05] on

<http://www.vamdc.eu/documents/standards/>

Note = TAP-XSAMS has been renamed VAMDC-TAP

Task 4: Registry Documents

Period 2 planned to work on improving registries to include VAMDC-TAP capabilities. Work has been performed on the subject. Documents have been edited in release [11.05] on <http://www.vamdc.eu/documents/standards/>

Task 5: Other Documents

A document is currently being drafted describing the process of Release and of version numbers. This document will be available in Period 3.

Significant results (Activities and Deliverables)

Deliverables to EU

D6.1 Interoperability Plan – DONE –

See <http://www.vamdc.eu/public-deliverables/18-deliverables-wp6>

D6.2 Interoperability Report to be included in report to the EU – Year 1 – Done –

See <http://www.vamdc.eu/public-deliverables/18-deliverables-wp6>

D6.3 Interoperability Report to be included in report to the EU – Year 2 – Done –

See <http://www.vamdc.eu/public-deliverables/18-deliverables-wp6>

Annual Interoperability Plan revisions included in Revised Annual VAMDC Project Plans – Year 1,2

See D1.2 and D1.5 <http://www.vamdc.eu/public-deliverables/12-deliverables-wp1>

Internal Deliverables

Standards are released every 3 to 4 months with versioning numbers on the www.vamdc.eu website which is the ONLY official source for the standards in VAMDC.

In Period 2, Release [11.05] of standards and softwares (in html and PDF) has been

D6.3 Interoperability Report 2

published. It should be noted that there has been a tremendous effort to get those standards coherently released.

Data access protocol, query language and dictionaries

- [Data access protocol \[v 11.05\] \[r 11.05\]](#)
- [Query language \[v 11.05\] \[r 11.05\]](#)
- [Dictionaries \[v 11.05\] \[r 11.05\]](#)

Data model

- [VAMDC-XSAMS reference guide \[v 0.2\] \[r 11.05\]](#)
- [VAMDC-XSAMS changelog \[v 0.2\] \[r 11.05\]](#)
- VAMDC-XSAMS schema files are available for download
- VAMDC-XSAMS schema documentation can be [viewed](#) in the browser or downloaded.
- Case-By-Case schema documentation can be [viewed](#) in the browser or downloaded from below.

Registry

- [Registry user's guide \[v 11.05\] \[r 11.05\]](#)

Deviations from the contract (Annex I) and reasons for them (if applicable)

No deviation

Failures to achieve critical objectives and/or not being on schedule and reasons for them (if applicable)

Not applicable

Proposed corrective actions (if applicable)

Not applicable

(approximate length of Part 1: 2 pages)

