



VAMDC

Virtual Atomic and Molecular Data Centre

D4.2

–

Infrastructure Deployment Report 1

Version 0.4

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	CTPM	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: D4.2
Deliverable title: Infrastructure Deployment Report 1
Due date of deliverable: June 2010
Actual submission date: 8 September 2010
Authors: G. Rixon, M.L. Dubernet
Work Package no.: WP4-SA1
Work Package title:
Work Package leader: CMSUC
Lead beneficiary: CMSUC
Dissemination level: PU
Nature: Report
No of pages (incl. cover): 12

Abstract	The objective of D4.2 is to describe VAMDC Infrastructure Deployment Report for Cycle 1. This report corresponds to Activities in WP4: SA1 “Infrastructure Deployment”. This report is included in the VAMDC Periodic Report for Cycle 1.
----------	---

Versioning and Contribution history

Version	Date	Reason for modification	Modified by
V0.1	20/07/2010	Section 5	G. T. Rixon
V0.2	17/08/2010	Draft of D4.2	M.L. Dubernet
V0.3	18/08/2010	Typographical corrections	G. T. Rixon
V0.4	23/08/2010	Added UCL-WP4 on HITRAN	M.L. Dubernet

Final Version (v0.4) released by		Circulated to	
Name	Date	Recipient	Date
M.L. Dubernet	8 September 2010	Mrs Asero	8 September 2010

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.

WP4 ACTIVITIES DESCRIPTION

Work package number	4		Start date or starting event:					3			
Work package title	SA 1: Infrastructure Deployment										
Activity Type	OTHER										
Participant id	1	2	3	4	5	6	7	8	12	13	14
Person-months per beneficiary: (Total = EU + Node Contributions)	66	36	36	12	6	6	8	11	5	9	24

Table of Content

WP4 activities description.....	5
1. WP4 Objectives.....	5
2. WP4 Milestones and Deliverables	5
3. WP4 Tasks Description.....	6
4. WP4 Tasks Description for Period 1.....	7
5. WP4 Tasks Reports for Period 1	10

1. WP4 Objectives

To provide Data Access via a homogeneous environment where the distributed user community can retrieve AM resources through a standard interfaces. This involves implementing standard outputs for the AM databases, finding the resources by interrogating registries, using querying and pipeline tools.

WP4 leader is CMSUC (2)

2. WP4 Milestones and Deliverables

Milestones

M4.1	Deployment of Data Access	WP4	CMSUC	Months 10, 22, 34, 42	
M4.2	Deployment of Infrastructure	WP4	CMSUC	Months 10, 22, 34, 42	
M4.3	Evaluation of Available Software	WP4	CMSUC	Months 10, 22, 34, 42	Testing by Users Panels of prototype software
M4.4	Open Call for New Resources	WP4	CMSUC	Month 24	Text of Call Available on Public Website

Deliverables

<i>D4.1 Infrastructure Deployment Plan (PM 3)</i>
--

D4.2 Infrastructure Deployment Report to be included in report to the EU – Year 1 (PM 10)

D4.3 Infrastructure Deployment Report to be included in report to the EU – Year 2 (PM 22)

D4.4 Infrastructure Deployment Report to be included in report to the EU – Year 3 (PM 34)

D4.5 Final Report of Service Deployment to be included in final report to the commission (PM41)

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

3. WP4 Tasks Description

WP4 Leader	G. Rixon (CMSUC)	
Task Number	Leader	Other Partners
1	G. Rixon (CMSUC)	All partners
2	L. Molina (CNRS)	CMSUC (2), IVIC (14), UU (6) + others TBD
3	K. Benson (UCL)	CNRS (1), UU (6), RFNC-VNIITF (12)
4	TBD (UCL) – not cycle 1	TBD
5	TBD (CNRS) – not cycle 1	TBD
6	TBD (CNRS) – not cycle 1	TBD

Description of work

The VAMDC infrastructure will be designed as an homogeneous environment where any AM producer or "community" users will be able respectively to publish their AM data or to retrieve and manipulate those data. The AM producers range from atomic physics to molecular physics handling complex molecules, solids and surfaces. The communities encompass astrophysics users from very different areas: stellar, galaxies, interstellar medium (those application areas are handled by the IVOA and Euro-VO projects), planetology and small bodies of the solar system (EuroPlanet Project), solar-earth system (EGSO and SPASE projects), atmospheric users (studies of earth atmosphere), environmental and combustion chemistry, fusion physics and industrial applications. The IVOA community is the most advanced project as far as building an interoperable infrastructure for astronomy and we will use some of their achievements, i.e. standards, tools, services when those are relevant to the project.

Task 1: Standard access to AM data (lead by CMSUC(2), all SAI partners)

We will provide standard service interfaces to AM databases. JRA1 will define these interfaces and this task is to implement them on the existing databases held by each VAMDC node. Participants at all nodes will be involved.

Task 2: Standard access to numerical codes (lead by CNRS(1), with partners (2), (6))

Where a VAMDC node has a useful numerical code for AM analysis or modelling, we will make it available as a service. These services will provide a uniform way of launching the codes and recovering their outputs. This task is complementary to the grid adaptation of code in SA2: the codes run on resources contributed by the node owners and need not be made portable to an external grid.

Task 3: Implementing registries (lead by UCL(3) with partners (1), (6), (12))

The registry facilities defined by JRA1, and implemented with the software produced by JRA2, must be populated with information. This task gathers the meta data for the services at each node and adds it to the registries.

Task 4: Augmenting VODesktop (lead by UCL(3))

The EuroVO's VODesktop is a generic interface for the virtual observatory. It allows access to all VObs data, plus launching numerical codes and sharing of data between desktop visualization tools. We will adapt A-M desktop applications to work with VODesktop and the underlying VObs applications environment.

Task 5: Publishing desktop software (lead by CNRS(1), with partner (5))

We will collect and make available to end users chosen A-M applications for the desktop.

Task 6: Expansion of the infrastructure (co-lead by CNRS(1) and CMSUC(2) with (5), (6))

Once the core infrastructure is deployed, new resources will be included in the infrastructure via an open call to producers of AM resources. Those new resources will need to be deployed and tested within the infrastructure. Task 6 will be devoted to the technical inclusion and testing of these new AM resources. The choice of these resources will be made in NA1 by the VPB by the EPT.

4. WP4 Tasks Description for Period 1

Task 1: Standard access to AM data

Task 1.1 TAP access to A-M data

Using existing software from the Virtual-Observatory movement, implement Table Access Protocol (TAP) services for sample nodes with data in relational databases.

TAP is the IVOA standard for remote access to relational databases. It is a low-level protocol, applicable to A-M physics as well as astronomy, suitable for prototyping VAMDC applications and workflows. Because complete, reusable, TAP software is available to us at no cost, TAP is a cheap way to make data available in year 1 of VAMDC before the VAMDC-specific web-services are defined.

Various databases will be selected for this trial according to schedule: CHIANTI, XSTAR, VALD, BASECOL, CDMS.

Task 1.2 Ingest data into databases

Ensure that all data-publishing nodes in VAMDC have their data copied into database systems suitable for supporting the VAMDC-standard web-services to be defined by WP6 and WP7.



Some VAMDC partners already have their data-sets stored in relational databases; others keep their data in flat files. The web services foreseen in the early planning require all data to be loaded into some kind of database. The exact kind of database depends on the design of the services.

Task 1.3 Install early prototypes of VAMDC-standard data-access services

Acquire an implementation of the VAMDC-standard data-access service, as defined by WP6 and implemented by WP7. Install this on all VAMDC node that have achieved a suitable database (as per task 1.2). Register the data-access service in the VAMDC registry.

These VAMDC-standard services will probably be based on the XSAMS data-model, possibly using the XQuery query-language. They operate at a higher semantic-level than the TAP services.

This task presumes that WP6 and WP7 produce a design and implementation suitable for wide deployment. It is possible that the new software does not become sufficiently mature for full deployment by the end of period 1. To conserve resources, two sites should be chosen for test deployment and evaluation before the software is deployed on all sites.

Task 2: Standard access to numerical codes

Task 2.1: Make a census in order to inventory numerical codes likely to be turned into public services

Task 2.2: Propose solutions to establish services in order to launch codes in VAMDC environment – test solution on one or two codes

Task 3: Implementing registries

Task 3.1 Choose a host for the VAMDC registry

Select a VAMDC partner to run the registry.

In period 1, we need only one registry on a single site. Towards the end of the VAMDC project we may set up a mirror of this registry on a separate site.

Task 3.2 Install registry software

Install the AstroGrid registry-service component at the chosen registry site. Establish registry access for all VAMDC partners who need to register resources.

The registry software for this task is a web service available ready-to-run from AstroGrid. The task does not include populating the registry with information; when the task is complete, the registry will be available to publishers but will not contain all the metadata needed by end users.

Task 3.3 Register existing resources

Register the existing, web-browser interfaces to data sets.

Task 4: Augmenting VODesktop

Task 4.1 Distribute TAP-capable copy of VODesktop

Ensure that all VAMDC staff have access to a version of VODesktop that works with the early-access TAP services. Make this software available for download from the VAMDC site.

By the time that VAMDC needs it, the default version of VODesktop should be suitable.

Task 4.2 Extend VODesktop to work with VAMDC-standard data services

Add code to VODesktop such that it can query the new prototypes of the VAMDC data-services deployed in task 1.3.

This task is only possible once WP6 designs the service protocol. If that design is not stable during period 1, it is not worth altering VODesktop; instead, we should produce simpler, disposable clients to test the services.

Internal deliverables

- TAP services for selected sites
- VAMDC-standard data-access services
- Registry service, ready for publishing, not fully populated
- Registration of web-browser (pre-VAMDC) interfaces
- VODesktop version for TAP services

5. WP4 Tasks Reports for Period 1

Period: 01/07/2009 – 30/06/2010

Work Package: WP4 – deployment

Work Package Leader and co-Leader: G. Rixon, A. Shih

Participants in the Work Package: CNRS, CMSUC, UCL, OU, UNIVIE, UU, KOLN, INAF, RFNC-VNIITF, IAO, IVIC

Part 1

A summary of progress towards objectives and details for each tasks

Task 1: access to data

Data access in VAMDC is by queries to relational databases mediated by web services. Using the partners' existing databases, we have deployed web services supporting both established protocols and new protocols developed within VAMDC.

Web services implementing the Table Access Protocol (TAP) of the International Virtual Observatory Alliance were deployed and registered on four data-sets: XstarDB (IVIC/CMSUC), BASECOL (CNRS), CHIANTI (CMSUC/UCL) and CDMS (KOLN). These deployments used the existing, relational databases and existing, free software by AstroGrid. The TAP services are available to users via VODesktop and the VAMDC web-portal.

TAP-XSAMS is a web-service protocol by VAMDC that extends TAP with standard views of the database and data output in the XSAMS format. Prototypes of TAP XSAMS were deployed on the data sets: VALD (UU), CDMS (KOLN) BASECOL (CNRS), and UDFa (QUB). These deployments are preliminary and are not yet registered; thus they are not available for general use. Full deployment of TAP-XSAMS is deferred to cycle 2.

Progress was made in creating new relational-databases for the data-sets, as required for the general deployment of TAP-XSAMS in cycle 2. The XstarDB data-set (curated by IVIC) was successfully established in a remote database (hosted by CMSUC). The CHIANTI data were exported to a database at UCL. CDMS (KOLN) has been ingested into MySQL. A solid-spectroscopy data-set (CNRS) was prepared for ingestion. The CDS-100 data-set (IAO) was ingested into a database. The rest of the data will be ingested during cycle 2, when tools and a standard schema are available from WP6 and WP7. A relational database model for HITRAN has been designed and a subset of molecules from the catalogue has been imported into a Relational Database Management System (MySQL was the chosen RDBMS); the process of importing molecules from the HITRAN catalogue to the RDBMS is ongoing.

The tools developed in WP7 have been wrapped into a web application which has been deployed at UCL as a prototype HITRAN web service; this web service can be found here: <http://msslsc.mssl.ucl.ac.uk/index.html>

Task 2: access to numerical codes

Large-scale numeric computations need high-performance hardware or a computational grid, neither of which is provided within VAMDC; WP5 covers access to external grids. Smaller computations can be run on demand on hardware provided by VAMDC partners via web services. In cycle 1, we investigated two protocols for these services.

Using Soaplab2 libraries and Perl wrappers, the stand-alone Xstar code has been transcribed to work as a web service (IVIC). With the help of any web browser, single-model calculations of XSTAR can be performed through the Spinnet web client. These services are also compatible with the Taverna workflow-tool. Soaplab2, its web client and Taverna are free software developed outside VAMDC.

The Xstar code, together with spectral-fitting codes (SPFIT and SPCAT from the CHIANTI package), was also made available (UCL) as a web service in the Common Execution Architecture system (CEA) of AstroGrid. This service is registered and can be used from VODesktop and Taverna.

From a preliminary survey of available codes (CNRS), we found that most codes are consumers of atomic and molecular data in other fields of science. There is no urgent need for VAMDC to execute these codes as services. However, there is great potential advantage in adapting these codes to consume data direct from the VAMDC data-services, instead of packaging the data with the code for execution. This changes the emphasis of this WP4-task for cycle 2.

Task 3: registry

VAMDC uses a central registry of metadata, established as a web service, to describe its resources. The information in the registry can guide scientists in their use of VAMDC and informs the connection between user interfaces, applications and VAMDC's web-services. Registration of a service makes it generally available to VAMDC users via VAMDC's standard user-interfaces.

A VAMDC registry was established (UCL) using free software from AstroGrid.

Available services were registered:

- web-browser interfaces to data that pre-date VAMDC (CMSUC)
- TAP services (CMSUC, IVIC, CNRS, UCL, KOLN);
- codes available through CEA (UCL).

Task 4: VODesktop and associated user-interfaces

VODesktop is a generic, desktop UI for IVOA web-services inherited from AstroGrid. We also support the Taverna tool for workflows and web-browser access via a web portal. All these UIs work the VAMDC registry

Minor upgrades were made to VODesktop (CMSUC) to ensure compatibility with the TAP services of task 1. Further upgrades to support TAP-XSAMS are planned. The current VODesktop can also drive the CEA services produced in task 2.

A prototype web-portal was developed (CMSUC) to demonstrate use of the registry and access to VAMDC services from a web site. The current portal can drive the TAP services and links to the legacy web-sites of the VAMDC partners. This first portal reflects, intentionally, the technical structure of the VAMDC registry and services rather than the scientific use of these resources. The portal will be rearranged with a science focus during cycle 2.

The Taverna workflow-tool is well established in several areas of science. It can be used with several VAMDC services and prototypes:

- TAP services (with AstroGrid plug-in for IVOA protocols);
- CEA services (with AstroGrid plug-in for IVOA protocols);
- SOAPlib2 code-execution services.

Significant results (Activities and Deliverables)

- Preliminary plan for deployments
<http://voparis-twiki.obsmpm.fr/twiki/bin/view/VAMDC/WP4>
- VAMDC registry: see http://registry.vamdc.eu/vamdc_registry
- Web portal: see <http://131.111.70.87:8080/registrybrowser/home.seam>
- Web services.

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

No deviations.

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

- a) Ingestion of data-sets into relational databases is slipped to cycle 2. The required schema has not yet been finalized by WP6.
- b) General deployment of TAP-XSAMS is slipped to cycle 2. The databases are not ready, and the web-service implementation, by WP7 is not stable yet.



Proposed corrective actions (if applicable)

The minor slippage can be absorbed in the cycle-2 plan and does not effect the overall outcome, schedule or budget of VAMDC.