VOTECH PROJECT

The European Virtual Observatory VO Technology Centre

EC Sixth Framework Design Study : Project 011892

Euro-VO Reference Architecture

Editor : A.Lawrence, on behalf of VOTECH team

Final version 2009-06-30

Summary

The VOTECH project has completed its prime task of undertaking the technical preparatory work needed in order to build the European Virtual Observatory (Euro-VO). We describe the conceptual structure of Euro-VO and the VO in general, along with the key technical components in place, and the required continuing work. The VO is conceived as a service architecture, with the three kinds of service being data services, application services, and infrastructural services. The key infrastructural components are identified as Registries; VOSpace; Identity and Community; middleware to allow tools to speak to VO services and each other; and a framework to support workflow and scripting. As well as a need to deploy and operate such infrastructural services, the technical arm of Euro-VO needs to provide support and tools both for data centres in deploying components and publishing data services, and for third party applications writers. VO tools will mostly be developed by third parties, but Euro-VO itself should provide some core tools, and lead the development in key areas - workflow tools, semantics related tools, visualisation and data mining services. Even though the VO is reaching maturity, it is very important to continue a programme of assessing emerging technologies and prototyping tools and infrastructural components which test these technologies. We confirm that the "three arm" structure of Euro-VO - VO Facility Centre (VOFC), Data Centre Alliance (DCA), VO Technology Centre (VOTC) - makes good sense, and that there is a vital continuing role for Euro-VOTC. We provide sixteen recommendations for Euro-VOTC and Euro-VO more generally.

CONTENTS :

1. INTRODUCTION

VOTECH is a Design Study project funded by the European Commission under Framework 6, whose top level goal has been to complete all technical preparatory work necessary for the construction of the European Virtual Observatory (Euro-VO), Following on from the European collaboration started by the Astrophysical Virtual Observatory (AVO) project in 2001-2004, VOTECH was proposed during 2004 and commenced work in January 2005. The participating partners are listed in Appendix A. The originally expected completion was in December 2008, but was extended to June 2009 in order to better capture new sets of requirements from the community, and to allow overlap with the EuroVO-AIDA project, which is beginning the process of implementing the VOTECH design studies.

VOTECH has built towards its top-level goal through a series of methods - the assessment and testing of new technologies; creating designs of new infrastructure components based on those technologies; collecting requirements and creating designs for new science user tools and datamining services; developing trial versions of new infrastructure components and tools, and datamining services used to test them; and last but not least, using our experiences to bring forward requirements for VO standards to the International Virtual Observatory Alliance (IVOA), and to contribute towards developing those technical standards. This work has been carried out in four substantive Design Study areas : Infrastructure (DS3); New User Tools (DS4); Intelligent Resource Discovery (DS5); and Data Exploration (DS6). Each of these work areas has produced Study Reports as well as software prototype deliverables. The aim of this final report is to collate and summarise the key findings from those studies, and to arrive at recommendations for the deployment phase which Euro-VO is now entering. The recommendations are derived through the flow of the text, and summarised in Section 8. Note that the term "Architecture" is not meant to imply a detailed technical architecture, as for example in a UML design, but rather a conceptual structure that will be followed in creating Euro-VO.

The document is meant primarily for consumption within Euro-VO, but is also intended to be of interest to other international VO projects, and to e-Science projects more generally. The aim is to summarise the major conclusions of the VOTECH project. It does not describe the work carried out to arrive at these conclusions, which can be found elsewhere in detail - in individual DS reports, and in material held on the project twiki. It aims to describe how Euro-VO *should be*, which is partially achieved, and sets out a continuing roadmap. The *recommendations* made are intended for the continuing Euro-VO Technology Centre (Euro-VOTC), except where stated.

2. NATURE OF THE VIRTUAL OBSERVATORY

The aim of the VO is to produce a transparent data infrastructure. Just as every web page feels as if it is somehow inside your PC, the VO aims to make it feel as if the world's astronomical databases are just sitting inside your own computer, waiting to explore. The key requirement to achieve this goal is *standardisation* - of data formats, data services, software components, and user tools. This does not require all these things to be the same, but only that they offer standard *interfaces*, and follow standard *data models*, so that they can communicate with each other. So for example, a data centre can use whatever database management system it likes, as long as it agrees to accept queries in a standardised form, and export standard data types and metadata. In addition to this service standardisation, the VO requires new types of *infrastructural service*, such as Registries, and Virtual Storage, and some method of expressing and communicating user identity, in order to preserve data access rights.

It has been clear for many years that a national approach, or even a European approach, to building such a data infrastructure is insufficient. The datasets and services which astronomers wish to use are spread all around the world. The key step is therefore to agree standards which apply globally. This is achieved through the International Virtual Observatory Alliance (IVOA). The resulting "global VO" can be seen as a kind of nurtured eco-system. It is not a spontaneous market-like system, but neither is it a rigidly designed monolith. The power of this co-operative standards based approach has been demonstrated clearly during the lifetime of VOTECH.

To some extent the emergence of globally agreed standards allows individual institutes, facilities, data centres and research groups to operate freely. However there is still a strong need for European co-ordination, for two main reasons.The first is to continue shared development - of new technologies and tools, new standards, community support and training - and also to share best practice and experience in deploying VO technologies. The second main reason is that we can establish background infrastructural services and components that are shared - for example, registries, virtual storage, and identification servers.

To achieve this European co-ordination, the planned Euro-VO has been structured with three arms. The VO Facility Centre (VOFC) is aimed at community support and central services. The Data Centre Alliance (DCA) is intended as a loose alliance of data and resource centres sharing best practice and policy. The VO Technology Centre (VOTC) is designed as a virtual organisation sharing work on technology development to feed to VOFC and DCA. The VOTECH project was seen as the prototype of Euro-VOTC.

3. VO SERVICES

The key lesson that VOTECH experience confirmed is that the VO should be constructed as a *service architecture*. Users can construct workflows to achieve the science they wish to undertake by combining calls to such services, either directly, or more often, using tools which understand the services. We can usefully divide services into three types - data services, application services, and infrastructural services.

3.1 Data Services

Data services are the bulk content of the VO. These can be straightforward access to data files; more structured access, for example to relational databases; or operations on data, such as requesting image cutouts, or running an SQL-like query on a database. An early conviction of VO projects worldwide was that the data provider must be left to manage their own holdings, and the workings of their services, as they see fit, with VO compatibility requiring only a relatively thin standard *interface layer*. A data service of a given type must (i) accept certain standard types of query, translating as necessary to its own required form; (ii) export certain standard data types, such as a VOTable; and (iii) describe what service it provides in a standardised way, so that this can be published in a Registry, and discovered.

Thousands of VO-compliant (or at the very least, VO-compatible) data services are now available worldwide. This has been an early success of the VO worldwide. However, it is clear that understanding the VO standard definitions, and constructing the VO-service layer, is a considerable burden on data providers. During VOTECH we have therefore put much effort into designing and prototyping software to assist data publishing - in particular the Data Set Access (DSA) software, and the Data Access Layer Toolkit (DALT). These tools enable professional data providers to run relatively simple configuration software that automatically sets up the service, and continues to run, acting as the necessary interface layer in action. Related prototypes are the Data Model Mapper (DMMapper), which helps to provide a simplified view of a complex database, and SAADA, an application which individual users can use to create publishable databases from their private datasets, a facility which is often requested by users at workshops. We therefore *recommend that operational versions of these toolkits be developed, expanded, and simplified, for both professional data providers and individual end users.*

3.2 Application Services

Any application that can be run in a "black box" fashion - i.e. which takes a standard set of input parameters and files, runs without attention, and then returns a standard set of output files or values - can be invoked as a remote service. The key requirement is for such a service to describe itself in a standardised way, for example stating the name and type of the input and output parameters, and specifying its access URL. It is then possible for such services to be published in a Registry and discovered, just like data services. Within VOTECH we have designed and prototyped a method to standardise application services in this way, known as the Common Execution Architecture (CEA). So far, this has been used only in Europe, but the concepts are now being re-developed for international standardisation as the "Parameterised Application" component of the Universal Worker Service (UWS) standard, known as UWS-PA. We believe that application services will be a major growth area, so *we recommend that Euro-VOTC vigorously pursue the UWS standard*.

VOTECH was very active in setting requirements and developing prototypes to demonstrate the feasibility of performing data exploration through the usage of visualisation tools as services. The concepts of the TOPCAT plotting tool were expanded to allow server-side computing (STILTS); similarly, advanced prototypes of a 3-D visualisation package were adapted to access VO data (VisIVO), together with two additional versions: a server-based one and a fully Web-oriented one (VisIVO-server and VisIVO-web).

Most implementations of applications services produced to date are not just calculations, but more typically operations on data - for example running a tool to detect objects within a supplied image file by running Sextractor at a remote location, or running a user-tuneable pipeline on interferometer fringes in the Jodrell Bank archive to make radio maps. In VOTECH we have undertaken considerable prototyping and analysis of high-powered datamining applications within a VO context - large database visualisation, neural net calculations, and multi-algorithm packages such as Astroneural and Weka. While these have not yet taken off worldwide in the way that simple data services have done, we are confident they are crucial to the future of Euro-VO.

3.3 Infrastructural Services.

Data services and application services are the things that users ultimately want, but infrastructural services are what make these join up into a powerful whole. Five types of service have emerged as necessary : queryable Registries of Resources; Virtual Storage Services; services managing Community based Identity; Middleware services to interface between applications and other VO services; and services which manage workflows. In the near future, we will probably also require ontology services, to transport the meaning of metadata. These key infrastructural components are summarised in the next section.

4. VO INFRASTRUCTURE COMPONENTS

4.1 Registry.

Possibly *the* central concept of the VO is that of the *Registry*, which is essentially the yellow pages of the VO, containing the standardised descriptions of those VO services. This makes it possible for users to discover the resources they wish to use, but also for software tools to automatically correctly invoke those services. Considerable effort has been expended worldwide into defining and developing the structure and standardisation of such Registries, with VOTECH staff playing key roles. The key concept is defining a standard set of Resource Metadata, including both general metadata, such as a unique identifier specifying the service, and metadata specific to the particular service type, defined through a series of extension schemas. Registries have been defined and developed in a way that maintains a core of global standardisation, while allowing local flexibility. For example, registries in Europe (in the UK) are "fine grained", holding information on the columns of database tables, whereas those in the US are "coarse grained", specifying where such extra information can be obtained from the data service itself. Registries around the world can "harvest" each other, so that although they do not contain identical information, they are all up to date. However there remains a continuing danger that Registries will drift apart. *We recommend that Euro-VO place a high priority on the interoperability of registries*.

4.2 VO Space

The second key concept is that of *Virtual Storage Services* or *VOSpace*. On the web, users are already used to their content being stored for them remotely by service providers, but the VOSpace concept is more than just a kind of extra scratch space. There are three main aims, the first of which is already achieved, and the second two of which are conceptually agreed but not yet implemented. (i) VOSpace acts as a kind of staging post that avoids security issues involved in writing to machines not owned by the software agent concerned. This can be used by VO services internally, storing intermediate results, managing workflows and so forth; or by users, for example storing a subset of a database to run further queries on, before downloading a final result. (ii) VOSpace will be able in principle to present and manipulate various kinds of standard *data structure*, rather than simply being a view to a filestore. (iii) VOSpaces can be *shared*, and so act as a collaboration space. This is closely tied to the concept of *Community* below.

The technical and standardisation issues concerning VOSpace are now largely solved, so that it is now possible for full implementations to proceed, and to use the technology to enhance the functionality of data services and user tools. There are some issues in the re-engineering of VOSpace to use REST rather than SOAP, but this should not delay deployment. We therefore *recommend that Euro-VOFC, and members of Euro-VO DCA move to deploy VOSpace services as soon as possible, and design new user tools that use the power of VOSpace*.

4.3 Identity and Community

It is necessary for the *Identity* of individual users to be expressed in a standardised and structured way. Although many datasets are completely public, and so can be accessed anonymously, others are proprietary, restricting access either to specific individuals (e.g. the PI and co-Is of a space mission or a telescope proposal) or to classes of individuals (e.g. only staff members of Universities A, B, and C). The traditional solution is to *authorise* specific individual users of each service by issuing a username and password, but complex VO workflows will then be interrupted many times to ask users to type these in, and there is no way to specify "anybody from University X". However if a standard string specifying a user uniquely is carried with every service request, then any service can compare to its own database and automatically tell whether the request is an authorised one.

Such issues have also been of concern in the Grid world, where the question is usually about authority to consume resources (e.g. CPU cycles), and in education, where the need is to allow students and staff to access a wide variety of services. The solutions explored within VOTECH aim to minimise the technical burden on users, avoiding the use of global "certificate authorities", and rather relying on initial login to a "community server", which is administered by some competent organisation, such as a university department or a national agency funded project. This creates a digital object representing the user which can then be passed on as required to services which follow the VO standard. This approach also in principle allows the "group" flexibility that we wish for the VO, allowing users to set access control to their own datasets for individually tailored lists of collaborators, although the precise mechanism for this "group" definition has not yet been agreed. This solution is in line with similar ones adopted by the networking community (i.e. NRENs).

VOTECH has made working prototypes of these concepts, and has made progress in contributing to the definition of the necessary standards on the international scene, but the situation is still quite fluid, because of changing background technology, and the needs and perceptions of other projects and users worldwide. *We recommend that Euro-VOTC work with international partners to complete standardisation as rapidly as possible, so that implementations can be deployed*.

4.4 Middleware

Use of the VO is much simplified by provision of interfaces between user, applications, and VO services. One method is to provide a *VO web portal*, such as the US NVO portal or the CDS portal developed while VOTECH was active, where all the necessary work is done by the portal. Alternatively, many single-purpose applications can access VO services directly; but for such an *ecology* of applications to provide a rich experience for the user, and for developers to be able to build new applications quickly and easily, it is strongly preferable to provide *middleware* that enables them to access VO services, and to communicate with each other.

VOTECH prototyped a simple but but effective "message hub" system (called PLASTIC) for application communication on the same desktop. This has been successfully standardised and broadened within IVOA, creating the "SAMP" standard. Increasing numbers of VO tools now intercommunicate in this fashion. This enables for example an image found by VO Desktop to be "piped" directly to Aladin for analysis, rather than having to add image analysis capability to VO Desktop.

For applications and web pages to exploit the VO, they need to know how to access those services - how to query the registry and return sensible results; how to invoke various kinds of data service; how to launch application services; how to examine and change VOSpace filestores; how to log into and out of community servers, and pass round users' credentials to services that request them; and how to access other related services, such as the Simbad name resolver. IVOA standards are defined explicitly so that competent applications can code these interactions, and many applications simply do this. However, it makes sense to package these interactions as a standard library, and preferably to present this not just as a code collection, but as a runnning service, so that applications make simplified calls which are routed through the middleware service. This is achieved by the "Astro Runtime (AR)" package, which also provides some standardised GUI components. The AR comes bundled inside the VO Desktop application, but can be also be run independently. During VOTECH, we also prototyped a middleware package for accessing Grid applications (VOTECHBroker). This remains experimental. *We recommend further exploration and implementation of methods of joining the VO and Grid worlds.*

Both PLASTIC/SAMP and the AR are client-side middleware. SAMP will remain so, but there is a clear need for server-side middleware, which will make it easier for applications and web pages to provide VO access seamlessly without having to instruct them to download the AR. *We recommend the development and deployment of server side middleware*.

4.5 Workflow / Job Execution

The power of the VO can be much amplified if it is possible to put together a *sequence* of actions, in which the output of one task is piped into another task. This can be achieved by graphical interfaces or by scripting, as described in Section 5. In either case, such a workflow needs to be able to run without user interaction, and to maintain a state, so that it can proceed without requiring to be permanently connected to the services called. (In other words they need to be *asynchronous*.) This workflow management can be done either client-side or server-side. On the client side, one can use applications (like Taverna, or the AIDA/JLOW system, which became available during VOTECH) or programming languages, such as Python. In this case, what is required is to write classes / routines / plug-ins that undertake the VO actions, and this approach has been successfully demonstrated. VOTECH also experimented with a server-side "Job Execution Service", but this proved harder than expected. *We recommend that the server-side approach be re-investigated in due course*.

4.6 Ontology Services

Our investigations of ontology (see below) indicate that at some point in the future Euro-VO may need to deploy some kind of OWL-compliant ontology service. *We recommend that investigations of ontology by Euro-VOTC continue, with a view to eventual deployment of ontology services by Euro-VOFC and members of Euro-VO DCA.*

5. VO TOOLS

5.1 General points

The Euro-VO infrastructure makes it possible for compatible tools to be built by a variety of providers - by Euro-VOFC, by data centres within the Euro-VO DCA, by other VO projects worldwide, and by third party providers. As explained in section 4, this is made possible by two middleware technologies - PLASTIC/SAMP, for application intercommunication, and the Astro Runtime (AR), for simplifying communication with worldwide VO services. However, we should not leave the construction of VO tools entirely to chance. Euro-VOFC has gathered VO tools requirements from the community in two ways : by running hands on workshops, and by operating a Science Advisory Committee. *We recommend that Euro-VOFC continue these methods*.

There are several areas where Euro-VO should continue to take a lead responsibility in tools development.

5.2 Core tools

Some tools are intrinsically connected to the VO infrastructure - for example those which query VO registries, or view and manipulate VOSpace (e.g. VOExplorer and File Explorer within VO Desktop). Others have historically been developed by Euro-VO partners, and are simply very popular - for example Aladin and Topcat. In all these cases, the tools concerned can eventually become independent, but in the short term, the Euro-VO partnership needs to guarantee their availability, maintenance, and continuing development.

A key VO science requirement that is still only partially met is the ability to perform distributed cross matching of large catalogues. Some of the issues in the cross-matching problem are scientific or algorithmic - calculation of likelihoods, use of prior information etc - and so are appropriate for individual application writers to address. Other issues - like the ability to link catalogues and pixel stores flexibly, for example to calculate upper limits in the absence of a match - are within the remit of data centres within Euro-VO DCA. But there are also cross-cutting infrastructural issues - solving the last mile problem, setting up dedicated fast cross-match links between data centres, to offer collaborative services, and so forth - which are within the remit of Euro-VOTC. *We recommend a combined Euro-VO campaign to address the cross-matching problem*.

5.3 Workflow and scripting tools

This is a very new area, and the related infrastructure components are still evolving. There are successful but provisional scripting implementations (eg AstroGrid Python, and, in the US, the "VO CLI" package) and graphical workflow builders (e.g. JLOW from CDS, and the Taverna application coming out of the Bio-informatics world). The key requirement, as discussed in Section 4, is for Euro-VOTC to develop and provide one or more APIs that make this possible, and also to consider development of a server side Job Execution Service. However, given that this is a new area, it is also advisable that Euro-VOTC work with other Euro-VO partners directly to develop next generation tools.

5.4 Semantics related tools

Semantics is an area where the background technology is evolving rapidly (see Section 6), so that constructing experimental tools that utilise this technology is very much a job for Euro-VOTC. (The exception is the use of "Unified Content Descriptors (UCDs)" which within the VO world is now a fairly mature technology). VOTECH made experimental versions of various tools related to resource discovery, semantics and ontology. These included an ontology explorer; a SIMBAD consistency checker; MEx, a metadata extraction tool; DJIN, a object name recognition tool; and SAADA (Systeme Automatique d'Archivage de Données Astronomiques) for ingesting metadata and automatically creating databases. These tools are at various stages of maturity and useability. *We recommend that Euro-VOTC continues constructing prototype semantics tools alongside its continuing exploration of semantics technologies*.

5.5 Lessons from data mining application services

Data exploration and data mining of very large databases is of increasing scientific importance. VOTECH took several existing packages and investigated how to integrate these into a scientific environment - Weka, Visivo, and Astroneural. In addition, a new experimental multi-dimensional visualiser (Eirik) was constructed. Because these techniques are generally computationally very intensive, they are not typically client-side applications, but rather remotely invoked application services, sometimes on a single remote server, and sometimes running on a Grid of machines. These explorations brought out several requirements for the Euro-VO infrastructure, and/or implementation at data centres. (i) Services which maintain the state of jobs while offline are crucial. (ii) Provision of VOSpace, and especially the facility to make temporary database extracts (referred to as "MyDB" in the US VO) is necessary. (iii) Some kind of broker is required which links standard VO interfaces and standard Grid interfaces (e.g. VOTECHBroker). (iv) Fast manipulation of very large databases needs several techniques to become standard - indexing, multi-threading, file mapping, and the use of column-oriented structures. (v) File mapping in particular will require data centres participating in Euro-VO to move to the use of 64 bit machines as soon as possible.

With these infrastructural elements in place, the emphasis will be on integrating existing data mining packages. We do not expect Euro-VOTC to develop its own data mining packages.

5.6 Inclusion of numerical simulations in the VO

The possibility of providing the community with advanced tools for the comparison between observational data and numerical simulations was pursued. This plan was coherent with the outcome of the EU/FP5 project AVO, in which a use case submitted to the project asked for this specific capability. The analysis carried out on the topic led to a European-led activity at the IVOA international level on the definition of standards to include numerical simulations in the VO. Two prototype standard protocols to access the simulations are currently being discussed, which have been demonstrated to be effective on different classes of numerical simulations derived from theory. *We recommend that Euro-VOTC continues activities related to the inclusion of theory data in the VO, actively pursuing the definition of an internationally-agreed standard data model and access protocol(s)*.

6. TECHNOLOGIES AND STANDARDS

6.1 Standard Background Technologies

Euro-VO has aimed to limit its technical development to astronomically specific software, using standard basic technologies as a background. This means that we do not undertake any original computer science or informatics research. However, we do undertake *assessment research*. Informatics technologies cover a broad landscape of choices, and it is not always clear which technologies are appropriate for our purpose. VOTECH has therefore put considerable effort into investigating and testing different technologies, and building software prototypes to undertake these tests. Standard technologies however continue to evolve. For example, SOAP-based web services seem to be giving way to REST-based web services, and our assessment is that REST-based services are in fact better suited to the VO. *We therefore strongly recommend that Euro-VOTC continue a significant level of technology assessment research*.

At an early stage of our technology assessment we decided that technologies emerging from the academic Grid world (Globus, Condor, etc) were not yet mature or stable enough to rely on. We have therefore built Euro-VO on commercial technologies (standard internet and web technologies; XML; SOAP/REST; X509/DNS, etc.). This will continue to be true, but we also expect to continually re-assess Grid-based technologies. In particular, the suitability of g-Lite, the middleware developed within EGEE and to be consolidated within EGI, will be investigated.

6.2 Emerging Technologies

Some technology areas are not yet stable enough to rely on, but are crucially important. These are the key areas for Euro-VOTC to continue technology assessment.

The first of these is the area of expressing and applying identity and authorisation credentials etc. The underlying standards (X.509 and DNS) are clear, but the technologies used to produce, authenticate, and transport this information are in active debate in the web and educational communities (e.g. OpenID, Shibboleth, Kerberos; hierarchical certificate authorities versus web of trust approach, etc.). Euro-VO has made trials and prototypes in this area, but cannot proceed with implementation until there is a clear IVOA agreement on which technological horse to back. Furthermore, compatibility with other authentication and authorisation initiatives at the European level (e.g. in the network and grid worlds) will be pursued

The second such area is semantics. In the VO, the term "semantics" covers a range from the simplest metadata definitions, to ontologies which can express relationships between terms, and build towards automated reasoning. The biggest success has been the international agreement on "Unified Content Descriptors (UCDs)", essentially a dictionary of standard names for table columns and related entities. This is technologically stable and requires only updating of the content. Going beyond dictionaries to ontologies holds great promise, but has not yet gelled into pragmatically deployable services and tools. VOTECH put considerable effort into assessing this area of technology and building prototype tools (see Section 5) and Euro-VOTC certainly must continue to do so.

6.3 IVOA standards

International standards are at the root of the VO, and the formation of the International Virtual Observatory Alliance (IVOA) is one of the key successes of the global e-Science scene in general. The implementation of the VO is however not necessarily identical in all parts of the world, and different national projects naturally proceed at different speeds in different areas. For these reasons, it is in the interest of Euro-VO to take a *planned approach* to developing IVOA standards, making sure that new ideas are brought quickly to the table, and that we put effort into IVOA working groups to accelerate the standards that we need. It is also clear that Euro-VO as a whole has more weight in IVOA than any one national project. *We therefore recommend that Euro-VO takes a co-ordinated approach to developing standards through IVOA*. End-users represented by Euro-VOFC, and data centres in Euro-VO DCA, are the "customers" for these standards, but Euro-VOTC should take the lead in developing them.

To date there has been significant progress in first generation data access protocols, registry and resource metadata definitions, early semantics standards, and software interface standards of various kinds. The next key steps are in definitions of data models, second generation data access protocols, and standards related to single sign on ability. These are the areas where we recommend that Euro-VOTC concentrate its efforts.

We also recommend that Euro-VOTC and Euro-VO DCA work closely together to examine and analyse the level of standards uptake and compliance in European data centres.

7. THE VOTC

Our experience in VOTECH has confirmed that the three-armed structure of Euro-VO makes good sense. The Euro-VO Facility Centre (VOFC) concentrates on liaison with and support for the end-user community, and channels their views on the ultimate scientific goals of Euro-VO; the Data Centre Alliance (DCA) co-ordinates the deployment of VO infrastructure, shares best practice, and makes sure that technical development meets the practical requirements of data centres, facilities, and projects; the Euro-VO Technology Centre (VOTC), working closely with the VOFC and DCA, undertakes shared technology development. In more detail, the key roles of Euro-VOTC are as follows :

- To support the work of European astronomical facilities and data centres
- To support the work of VO tools writers
- To build robust and reliable software components for deployment in the Euro-VO infrastructure
- To carry out continuing technology assessment, including building prototype tools and infrastructure components to test those tools
- To liaise closely with Euro-VOFC and Euro-VO DCA, especially in setting requirements for continued technical work
- To lead the planning and co-ordination of international standards work within Euro-VO
- To directly participate in key IVOA working groups

Euro-VOTC is intended as a virtual organisation co-ordinating the work of many different European partners. A separate document (the Euro-VOTC charter) sets out the organisation that will be used to achieve these aims.

8. RECOMMENDATIONS

Here we collect our recommendations. They are numbered sequentially, but also labelled with the section in which they arose.

(1/3.1) We recommend that operational versions of data publication toolkits be developed, expanded, and simplified, for both professional data providers and individual end users.

(2/3.2) We recommend that Euro-VOTC vigorously pursue the UWS standard in order to standardise the way remote applications can be included in the VO framework.

(3/4.1) We recommend that Euro-VO place a high priority on the interoperability of registries.

(4/4.2) We recommend that Euro-VOFC, and members of Euro-VO DCA move to deploy VOSpace services as soon as possible, and design new user tools that use the power of VOSpace.

(5/4.3) We recommend that Euro-VOTC work with international partners to complete standardisation on Identity and Community expression as rapidly as possible, so that implementations can be deployed.

(6/4.4) We recommend further exploration and implementation of methods of joining the VO and Grid worlds.

(7/4.4) We recommend the development and deployment of server side middleware.

(8/4.5) We recommend that the server-side approach to workflow execution be re-investigated in due course.

(9/4.6) We recommend that investigations of ontology by Euro-VOTC continue, with a view to eventual deployment of ontology services by Euro-VOFC and members of Euro-VO DCA.

(10/4.7) We recommend that Euro-VOFC continue seeking community input through workshops and through a combined Euro-VO Science Advisory Committee.

(11/5.2) We recommend a combined Euro-VO campaign to address the cross-matching problem.

(12/5.4) We recommend that Euro-VOTC continues constructing prototype semantics tools alongside its continuing exploration of semantics technologies.

(13/5.6) We recommend that Euro-VOTC continues activities related to the inclusion of theory data in the VO, actively pursuing the definition of an internationally-agreed standard data model and access protocol(s)

(14/6.1) We strongly recommend that Euro-VOTC continue a significant level of technology assessment research.

(15/6.3) We recommend that Euro-VO takes a co-ordinated approach to developing standards through IVOA.

(16/6.3) We recommend that Euro-VOTC and Euro-VO DCA work closely together to examine and analyse the level of standards uptake and compliance in European data centres

APPENDIX A : VOTECH PARTNERS

Project Co-ordinator :

Professor Andy Lawrence (al@roe.ac.uk)
The University of Edinburgh, United Kingdom

Co-investigators :

Dr Richard McMahon (rgm@ast.cam.ac.uk)
The University of Cambridge, United Kingdom

Professor Mike Watson (mgw@star.le.ac.uk)
The University of Leicester, United Kingdom

Dr Francoise Genova (genova@cluster.u-strasbg.fr)
CDS, Observatoire Astronomique de Strasbourg, France

Dr Fabio Pasian (pasian@ts.astro.it)
INAF, Osservatorio Astronomico di Trieste , Italy

Dr Peter Quinn (pjq@eso.org)
European Southern Observatory, Germany
Note : during the project Dr Quinn was replaced by Dr Fernando Comeron

In addition, although not a formal partner, staff at the European Space Agency (represented by Dr Christophe Arviset) worked closely with the VOTECH project.