Proposal full title	***The European Virtual Observatory - VO Technology Centre***
Proposal acronym	***VO-TECH***

List of contents

OVERVIEW

Table 1 - List of participants of the Design Study

Participant number (co-ordinator as participant N°1)	**Organisation (name, city, country)**	**Short name** (as specified on form A2)	**Short description (i.e. fields of excellence) and specific roles in the consortium**
1	University of Edinburgh, representing UK AstroGrid Consortium.	**UEDIN**	AstroGrid consists of ten UK universities and public laboratories collaborating to provide one of the three major VObs projects worldwide. It was one of the founding members of the International Virtual Observatory Alliance (IVOA). AstroGrid has lead responsibility for the VO Technology Centre (VOTC), and so for leading the VO-TECH project. The University of Edinburgh hosts several internationally leading eScience activities : the UK's National e-Science Centre, the National Digital Curation Centre, and the Edinburgh Parallel Computing Centre (EPCC). It also hosts what will soon be the world's largest astronomical database - the WFCAM Science Archive. Edinburgh acts as the co-ordinating organisation for the VO-TECH project.
2	European Southern Observatory	**ESO**	The European Southern Observatory is a European intergovernmental research organisation developing and operating observatories in Chile on behalf of astronomers in 10 member states of the EU. ESO operates the world's largest astronomical data archive with more than 3000 registered users and provides access to 50 Terabytes of data from ESO telescopes. ESO is the co-ordinating organisation of the FP5 AVO RTD project and founding member of the IVOA. ESO will lead the Euro-VO Facility Centre (VOFC), participate in the Data Centre Allinace (DCA), and contribute to several VOTC tasks.
3	University of Leicester, representing UK AstroGrid Consortium	**LU**	AstroGrid consists of ten UK universities and public laboratories collaborating to provide one of the three major VObs projects worldwide. It was one of the founding members of the International Virtual Observatory Alliance (IVOA). AstroGrid has lead responsibility for the VO Technology Centre (VOTC), and so for leading the VO-TECH project. The University of Leicester is the centre of the technical management of AstroGrid, and has a key involvement in the UK national e-Science programme as an e-Science Centre of Excellence. The University is the home to international data centres for ESA's *XMM-Newton* observatory and NASA's *SWIFT* mission, and also hosts LEDAS, the primary European archival centre for high energy astrophysics.

4	University of Cambridge, representing UK AstroGrid Consortium	**UCAM**	AstroGrid consists of ten UK universities and public laboratories collaborating to provide one of the three major VObs projects worldwide. It was one of the founding members of the International Virtual Observatory Alliance (IVOA). AstroGrid has lead responsibility for the VO Technology Centre (VOTC), and so for leading the VO-TECH project. The University of Cambridge hosts several internationally renowned leading eScience and Computer Science activities; the Cambridge e-Science Centre, the Microsoft European Research Centre. It also hosts CASU, a leading astronomical data archive centre.
5	Centre National de la Recherche Scientifique, representing French VO	**CNRS DR10**	The Centre de Données Astronomiques de Strasbourg (CDS) will act on behalf of the French VO. CDS is the world leading organization for astronomical catalogues. It is in charge of the Interoperability Work Area of the FP5 AVO RTD project and started the international actions for interoperability in the VObs in the FP5 OPTICON Network. French VO-CDS will participate in all tasks of VO-TECH project and take special responsibility in DS5, Intelligent Resource Discovery.
6	Istituto Nazionale di Astrofisica, Roma, Italy	**INAF**	INAF is the Italian institution co-ordinating research in astrophysics through the network, consisting of Astronomical Observatories and Institutes geographically distributed over the national territory and of the "Galileo" observing facility located in La Palma, Canary Islands. INAF is active in the fields of grid technologies (deployment of infrastructure and integration of domain-specific applications) and archives of astronomical data (from both ground-based and space-borne facilities).

Table 2 - List of Tasks for the *Design Study (DS)*

Task No	Descriptive Title	Lead participant	Short description and specific objectives of the task
DS1	*Management of DS*	**UEDIN**	Co-ordination, financial oversight, general support, and external presence
DS2	Technical Project Management	**LU**	Project planning, technical co-ordination, integration, external liaison
DS3	New Infrastructure	**LU**	Design and testing of new components, architectural design
DS4	New User Tools	**ESO**	Specification, design and testing of new user tools
DS5	Intelligent Resource Discovery	**CNRS DR10**	Feasibility study, and subsequent design and testing of new components
DS6	Data Exploration	**UEDIN**	Feasibility study, and subsequent design and testing of new components

Table 3 - Summary table of expected budget and of the requested Community contribution

	Participant number													
Task number	**1**		**2**		**3**		**4**		**5**		**6**		**Total expected**	**Max Community**
Amounts (€)	exp. Budget	req. contrib.	exp. budget	req. contrib	exp. budget	req. contrib	exp. budget	req. contrib	exp. budget	req. contrib	exp. budget	req. contrib	**budget (€)**	**contribution requested (€)**
DS1	203514	203514	3840	1920	3840	3840	3840	3840	3840	1920	3840	1920	222714	216954
DS2	267600	267600	13200	6600	31200	31200	31200	312000	13200	6600	13200	6600	369600	349800
DS3	18000	18000	207420	103710	272400	272400	254400	254400	200400	100200	0	0	952620	748710
DS4	0	0	414840	207420	0	0	200400	200400	400800	200400	254400	127200	1270440	735420
DS5	18000	18000	207420	103710	200400	200400	0	0	454800	227400	200400	100200	1081020	649710
DS6	254400	254400	0	0	18000	18000	18000	18000	200400	100200	400800	200400	891600	591000
Total expected budget (€)	761514		846720		525840		507840		1273440		872640		4787994	
Max Community contribution requested (€)		761514		423360		525840		507840		636720		436320		3291594

1. EUROPEAN ADDED VALUE OF THE NEW INFRASTRUCTURE

1.1 European Astronomical Infrastructure. Astronomy is a European strength. It plays a significant role in our culture, our education, and our entertainment, is a demanding customer for technology, and has often driven fundamental developments in science. The integrated quality and volume of European astronomy is second to none, but this is spread over many nations, and any one nation finds it hard to compete with the US. A combined approach to provision of world class facilities through the European Space Agency (ESA) and the European Southern Observatory (ESO) has made an enormous difference. A further step towards coherence has been taken by the recent funding of the OPTICON and RADIONET Integrated Infrastructure Initiatives, providing a European framework for strategic coherence, joint planning and facility pooling. In this proposal we describe a radical new step, the provision of seamless unified access to European data holdings, putting all European astronomers on an equal footing : the European Virtual Observatory. The precursor AVO project grew out of the OPTICON initiative, and both OPTICON and RADIONET strongly endorse our proposal. The Euro-VO vision is of course quite similar to concepts in other scientific and commercial areas, and here too we build on the excellent background from recent EU initiatives - especially EGEE, who also endorse our proposal..

1.2 The Virtual Observatory Vision. The power of the World Wide Web is its *transparency* - it feels as if all the documents in the world are inside your PC. The idea of the *Virtual Observatory* (VObs) is to achieve the same transparency for astronomical data. All the world's data on your desk - all archives speaking the same language, accessed through a uniform interface, and analysable by the same tools. A central goal is **democratisation** : the power the scientist has at her fingertips should be *independent of location*. Such an infrastructure will also encourage *collaboratories* - informal distributed research teams sharing data, workflows, and analysis results in a transparent virtual storage system.

Transparency is also a goal of *computational grids*, where a set of distributed computers feels like one supercomputer on your desktop. The VObs concept can be seen as a domain-specific example of a *data grid*. However it goes one step further, as what is offered is not just access to the data, but operations on the data and returned results which are essential for their full exploitation - for example the ability to stack and mosaic images, to query catalogues and create subsets, to integrate data from different origins, or to calculate a correlation function. Today such analysis is done by end-users after downloading data. In the future we expect that the normal mode will be for such calculations (many of which are quite standard) to be *data services* offered by the expert data centres holding the data. These operations then also need to be standardised to be compatible across many archives. The result is a *service grid*. The VObs will not be a monolithic system, but, like the Web, a set of standards which make all the components of the system *interoperable* - data and metadata standards, agreed protocols and methods, and standardised mix-and-match software components. These standards and software modules constitute the *VObs Framework*. To achieve the whole vision, however, data centres, tools writers, and facility builders all need to accept the new framework and work within it. Five strands of work are needed :

1) Development of standards and protocols, and their international agreement.
2) Construction of "glue" software components - portal, registry, workflow, user authentication, virtual storage.
3) Uptake by data centres, who need to "publish" to the system, i.e. to write VObs compliant data services connected to their holdings.
4) Construction of tools to do science with the data.
5) Establishing and maintaining resource registries and user support systems.

1.3 Priority of the VObs vision. The VObs concept has a high priority in most national astronomy programmes, leading to fourteen projects around the world. Many other nations have recognised the issue, but have not yet been able to afford new projects. Large organisations such as ESA, NASA, ESO, and NSF have all recognised the strategic importance of the VObs, and it is one of the main issues under discussion in the current OECD Global Science Forum study of future large scale facilities. The community itself, as well as its political leaders, has made its interest clear over the last two years, through a number of dedicated VObs conferences and workshops, and special sessions at large general meetings, for example at the last General Assembly of the International Astronomical Union, in Sydney. Some of this drive comes from the widespread interest in grid middleware and e-science more generally, but mostly it comes from awareness of the imminent data flood, and the constantly raising expectations of astronomers concerning the quality and power of web-based tools. The general feeling of most astronomers is that something like this simply has to happen : they are very keen that it happen in an organised and professional fashion.

1.4 The need for a European solution : Euro-VO. The vision is global, but there are many reasons to develop a specifically European infrastructure. (a) We intend to build a concrete implementation, a working system of daily use to scientists, and wish to do this in a manner that uniformly benefits astronomers Europe-wide, regardless of location. (b) Europe owns, leads, or shares a stunning array of astronomical facilities, with more to come during the next two decades. Euro-VO is aimed at getting maximum return from these expensive investments, and at giving European astronomers a competitive edge. (c) Key skills and leading edge technical teams are located at a number of European institutions, who have begun collaborating during the FP5 AVO programme. We aim to co-ordinate these efforts to create a critical mass rivalled only by the US. (d) We are in an excellent position to exploit EU investment in GEANT and now EGEE, giving us the potential for mass-scale compute applications in astronomy . (e) Elements of the vision are best tackled at a level larger than individual nations, but a truly global organisation is impractical.

A partnership of ten organisations has come together to create the Euro-VO as European structure. This enables us to combine national contributions in a way that gives us a value larger than the parts, and to take a lead on the global stage. We will create three interlinked structures.

- **The EURO-VO Data Centre Alliance (DCA)**: a network of European data centres who will populate the system with data, provide the physical storage and compute fabric, and using VObs technologies, will publish data, metadata and services to the EURO-VO .
- **The EURO-VO Facility Centre (VOFC):** an organization that provides the EURO-VO with a centralized registry for resources, standards and certification mechanisms as well as community support for VObs technology take-up and dissemination and scientific program support using VObs technologies and resources.
- **The EURO-VO Technology Centre (VOTC):** a distributed organization that coordinates a set of research and development projects on the advancement of VObs technology, systems and tools in response to scientific and community programme needs.

The VOTC is the main focus of this proposal. An initial implementation of Euro-VO will begin as early as 2005, but we foresee the completion of the vision several years downstream, with significant technical challenges along the way. The job of the VOTC is to carry out a sequence of design studies, leading to implementation in a working Euro-VO in a rolling programme.

1.4 New Science, Enhanced Science. We expect Euro-VO to improve the volume, quality, and cost-effectiveness of astronomical research across an extremely wide range of problems. The first driver is to accelerate the quality of **on-line research**. Astronomers already do much of their research on-line through data centres. The idea is to step up the quality of service beyond simple access to archives by downloading subsets. This will mean the ability to make *complex queries* of catalogues of objects or catalogues of observations, and the ability to *analyse* the data in situ - for example to transform or pan across an image, or to draw a colour-colour-colour plot for selected objects and rotate it. Such improved service can be seen as part of a long trend in astronomy to develop communally agreed *standard tools* so that the astronomer can concentrate on doing the science rather than wiring their own instruments, or hacking their own data reduction software. However we are also driven to this solution by the expected **data explosion** in astronomy. For very large datasets, such as the optical-IR sky survey which VISTA will accumulate at hundreds of TB per year, users can't afford to store their own version, or have time to download it. Data centres are therefore driven to provide analysis services as well as data access. The next driver is the ability to make **multi-archive science** easy. The study of quasars requires data at all wavelengths; finding rare objects such as brown dwarfs involves simultaneous searching in optical and IR data. The idea is to transform this kind of science from slow and painful hand-driven work to push-button easy, so that through a single interface one can make *joint queries* such as "give me all the objects redder than so-and-so in UKIDSS that have an XMM ID but don't have an SDSS spectrum", or ask higher-level questions, such as "construct the spectral energy distribution of the object at this position". Sometimes the tasks will involve predetermined lists of data services, but often they will involve the Euro-VO system making a trawl and deciding what is relevant, using some kind of *registry of services*. We also wish to facilitate **data intensive science**. Some of the most interesting science comes from manipulation of huge numbers of objects. This can mean looking for rare objects, for example those with strange colours or proper motions, or constructing a correlation function, or fitting gaussian mixtures to N-D parameter sets, and so on. At the moment such projects are the province of specialist "power users", but the vision is to make such analysis easy, as a service through data centres. This will require data centres to provide not just storage but also high-powered search and analysis engines and *standard tools* .

1.5 Euro-VO stakeholders. The Euro-VO will be to the benefit of a very large community of end-users - all the individual astronomers in University departments and laboratories across Europe. There are also three clear groups of organisations who have a direct interest in implementing the vision, and seeing that it happens in a well organised and intelligent fashion. (i) The first group is that of established European astronomical data archive centres. They wish their data collections to be put to maximum use doing the best science. They will need to take up the VObs infrastructure, and write compliant services. (ii) The second group is that of data creators - European observatories and mission centres, as well as large consortium science projects. As well as constructing new facilities with output to the VObs in mind, they wish to close the loop between analysis and data collection. (iii) The third group is that of astronomical software specialists - both tools writers such as the Starlink team, and those with a track record in infrastructure development, such as CDS, and existing national VObs projects, such as AstroGrid and GAVO (German Astrophysical Virtual Observatory). Euro-VO is bringing together stakeholders across this range. The list of data centres is potentially rather large, but we have initiated the Data Centre Alliance with national representatives, and have supplemented this with representatives from key organisations such as ESO and ESA, as well as key software groups at CDS and AstroGrid. We are in the process of agreeing an MOU that establishes Euro-VO regardless of the precise funding status. The VO Technology Centre (VOTC) is a more limited consortium of those with a serious interest in and capability for delivering the new technological infrastructure, centred around the original AVO partners. It is this latter partnership that is making the current proposal.

2. SCIENTIFIC AND TECHNOLOGICAL EXCELLENCE

2.1 Quality of the new infrastructure

2.1.1 Current status

(a) Data Centres. Public astronomical data collections in Europe arise three different ways - through primary data creation facilities, through data warehouse specialists, and through added-value project centres. (i) Space mission centres provide on-line access to the accumulating data products, sometimes through ESA entities such as VILSPA and ESTEC, and sometimes through the PI groups responsible for particular hardware. Traditionally, ground-based observatories archived only simple tape copies of raw data, with anything further left up to individual observers. However, ESO has now developed a space-mission quality online archive for data from all ESO telescopes. Such facility centres normally provide only raw data. (ii) A small number of groups - such as CDS (Strasbourg) and LEDAS (Leicester) in Europe, and HEASARC in the US - have specialised in building extensive and documented *collections* of data from various sources. These groups have also been active in developing *standards* for storing and accessing such data, making the first steps towards the VObs. (iii) Often the most popular datasets are not heterogeneous collections, but uniform products from coherent enterprises, such as survey projects, with access provided by the organisations running the scientific projects. European examples are TeraPix (Paris), WFAU (Edinburgh) and CASU (Cambridge), and the most prominent recent US example is the Sloan Digital Sky Survey (SDSS). Such groups make available both raw data and *added value* derived products such as calibrated image sets and source catalogues. Increasingly such groups are starting to offer more than simple download, such as simple SQL queries on the source catalogues.

(b) Virtual Observatory Projects and the IVOA. During 2000-2002, several groups around the world began to develop the *Virtual Observatory* vision described in section 1, leading to three major projects - the US National Virtual Observatory project (US-NVO), the European Astrophysical Virtual Observatory project (AVO), and the UK AstroGrid project. It was quickly apparent that international collaboration was crucial in this area, so the three projects began having regular joint telecons, which led during summer 2002 to the formation of the International Virtual Observatory Alliance (IVOA : see http://www.ivoa.net). Meanwhile other national projects were beginning around the world, and now the IVOA contains fourteen projects. The IVOA has a remit to develop and issue standards relevant to the global VObs initiatives, through a set of working groups. It uses a well defined multi-stage process based on that of the W3C. These are then passed for final endorsement to a working group of the IAU - the top-level professional and standards organisation for the worldwide astronomical community. The IVOA also hosts a discussion forum, cultivates best practice, and holds a series of meetings and workshops.

(c) AVO and AstroGrid progress. The AVO project (http://www.euro-vo.org) was funded by FP5 as an RTD action [HPRI-CT-2001-500030]. It began in November 2001 and will complete in October 2004. It is a Phase A study, undertaking R&D leading towards a concrete Phase B implementation of the Virtual Observatory in Europe. It concentrates on three areas - definition of the science case, developing interoperability standards, and assessing the relevant technologies. The science work area has been led by ESO. The project set up a large *Science Working Group* with interested scientists from across the world, and through it has produced science-led demonstration events, which in turn have driven development of necessary pilot software. The interoperability area has been led by CDS at Strasbourg, who have been trailblazing this area internationally for some years. This work has been international from the beginning, but especially so since the development of the IVOA. AVO has had a leading role in new standards for table data, images, spectra, resource identifiers, and semantics content descriptors. Technology development has been led by the AstroGrid consortium. AstroGrid (http://www.astrogrid.org) is a UK project largely funded by the UK e-science programme, and partly by the AVO contract. It aims at building a preliminary working implementation of the VObs vision. In order to do this, the project assesses new technologies, such as grid middleware and web services, and deploys them in an iterative sequence of trial implementations. The R&D component is a direct deliverable to the AVO programme, but in fact the whole of the AstroGrid project works effectively towards the idea of the Euro-VO.

AVO is very much on target, and has delivered both key and public software demonstration events.The ’1st light prototype event’ was held at Jodrell Bank in 20 January 2003, whilst the prototype demonstrating ’1st Science’ was held at ESO in 27 January 2004. Full details can be found at http://www.euro-vo.org. Indeed, even with the first preliminary prototypes, new science is resulting from their use. Padovani et al (2004) report on the use of the AVO 1st Science Prototype to discover an enlarged sample of optically faint and obscured qusars from deep multiwavelegth survey data. Meanwhile, US-NVO have used prototype services in the US to mine large optical and infrared survey data sets to discover previously undetected brown dwarfs (Berriman et al 2003).

2.1.2 Components of the Euro-VO vision. The top-level goals of the Euro-VO programme are :

EURO-VO-Objective 1: Technology take-up and full VObs compliant data and resource provision by astronomical data centres in Europe
EURO-VO-Objective 2: Support to the scientific community to utilize the new VObs infrastructure through dissemination, project support, and VO facility-wide resources and services
EURO-VO-Objective 3: Building the VObs infrastructure in working practice, and preparation to meet new scientific challenges, requiring development and refinement of VObs components, assessment of new technologies, design of new components, and their implementation

These objectives will be met primarily by the DCA, the VOFC, and the VOTC respectively. It is the third of these objectives that requires the design study we propose here.

As a result of the initial work by AVO and others we have a clearer idea of the structure that Euro-VO will require. In its widest sense, it requires a healthy community of observing facilities, data centres, a physical grid of resources (storage and compute facilities and high bandwidth links), and software tools writers. But at the heart of Euro-VO is an infrastructural "glue". The key technologies/components are as follows. (i) Standardised data exchange, involving both data and metadata standards, and message exchange protocols (eg SOAP/WSDL, or grid services). (ii) Registries of resources - queryable data collections and other services. (iii) Methods to enable single sign-on - authentication and authorisation. (iv) Methods to compose complex sequences of jobs in the new distributed environment, in a standardised way that other services can understand - workflow technology. (v) Methods for data centres to collaborate in presenting distributed virtual storage to the user - intermediate results, workflow files, log files, etc.

2.1.3 What then must we do ? Current projects will have scoped out the issues very well by the end of 2004. They will also have made significant progress in first implementations of key concepts. But the full realisation of the Euro-VO vision will take several more years. The work needed corresponds to the three interlinked structures described in section 1.

(1) The community of data centres and data creation facilities needs to be brought together in a coherent and organised fashion, creating the *Data Centre Alliance (DCA).* Work is needed to deploy VObs infrastructural software (including standard Grid middleware where appropriate), publish data services, and establish the necessary physical fabric of storage and compute facility. There is also a need simply to act as a forum to exchange knowledge and encourage best practice. (2) A persisting entity - the *VObs Facility Centre (VOFC)* - is needed to co-ordinate the whole programme, to provide and maintain registries for resources, standards and certification mechanisms, and to complete and rollout agreed software components. In addition, there is need to link to the scientific end-user community, and to provide community support, dissemination, and scientific programme support. (3) The above two strands of work concern either human networking or technical construction. However the technical infrastructure is far from ready to deploy in its

final form. The continuing programme of preparatory technical work is the responsibility of the *VObs Technology Centre (VOTC)*. The work programme of the VOTC is precisely the focus of this proposal.

In Section 3.2 we describe how we are trying to set about creating DCA and VOFC from combined national funding and treaty funding. The VOTC is being created from a combination of national funding and the funding requested in this design study proposal. (The possibility of later FP6 proposals to help create VOFC and DCA is under consideration.).

2.1.4 Key technology areas for the VOTC.

(i) Some parts of the technical infrastructure have already made a good start, so the requirement is detailed design study, before being taken forward to implementation. In other areas serious design is only becoming possible now as the relevant technologies and background infrastructures mature. This is most obviously the case for grid middleware, which will now be deployed across Europe by EGEE. Other areas where the technology is stabilising enough for serious design work to begin are in workflow, software agents, authentication and authorisation, and distributed storage.

(ii) New end-user tools need to be developed that can take advantage of the VObs infrastructure. Some of these will be essentially re-worked versions of standard tools - image browser, tuneable source extractor, graph plotter etc - but the new infrastructure also suggests new kinds of tools, such as a tuneable cross-matching tool, a query tool for returned data, a metadata browser, and so on.

(iii) The central issue in the VObs is that of *resource discovery*. Our current work takes us beyond lists-of-links to standardised and queryable registries, of both data collections, and of useable services. The next stage is a system that understands the *semantics*. Much of the connection between software components can then be *automated*, and the user can ask higher-level questions, with a meaningful return from unexpected places. Of course this is very much the agenda of the *semantic web*. One of the most important developments is a prototype Web Ontology Language (OWL). The technology is just arriving at the stage where we can contemplate building it into Euro-VO, but we need first both a feasibility study, and then a proper design phase.

(iv) We have barely started on the ideas of data intensive exploration and analysis offered as services. This includes manipulations, transformations, and statistics of whole data sets, such as Fourier transforms, Gaussian mixture and PCA analysis, outlier location, and then visualisations of such derived products, which are normally multi-dimensional. Considerable algorithmic advances have been made recently in areas such as kd-trees and multi-dimensional visualisation techniques, but they have not been implemented in a distributed environment, and only for thousands of points, as opposed to the billions of points in upcoming astronomical databases. Furthermore, to date one has to be a "power user" - intimately acquainted with the structure and quirks of a particular data set, and nuts and bolts of pieces of home grown software. Our aim is that standardisation will deliver polished and powerful versions of such services that every astronomer can use.

2.1.5 Outputs to the community. Euro-VO will link the core building blocks (ESO, ESA, national observatories, national data centres) into a robust infrastructure which will provide a unified virtual data resource enabling forefront astronomical research. Euro-VO will be a world leading infrastructure that will lead to better, faster, and more cost-effective science across the whole range of astronomical problems. It will give seamless distributed access to all astronomical data resources via common intuitive interfaces; the ability to perform large and complex data discovery and manipulation tasks; the ability to support diverse scientific discovery programmes ranging over cosmology, galaxy formation, exotic objects, and the evolution of stars and planets. It will also encourage the formation of *collaboratories* - distributed on-line research programmes whre groups of researchers, located in multiple European institutions, are to able to conduct analysis in a shared manner, where their processing and workflows, log files, and analysis results are stored in server based systems available to their collaborators in the secure Euro-VO system.

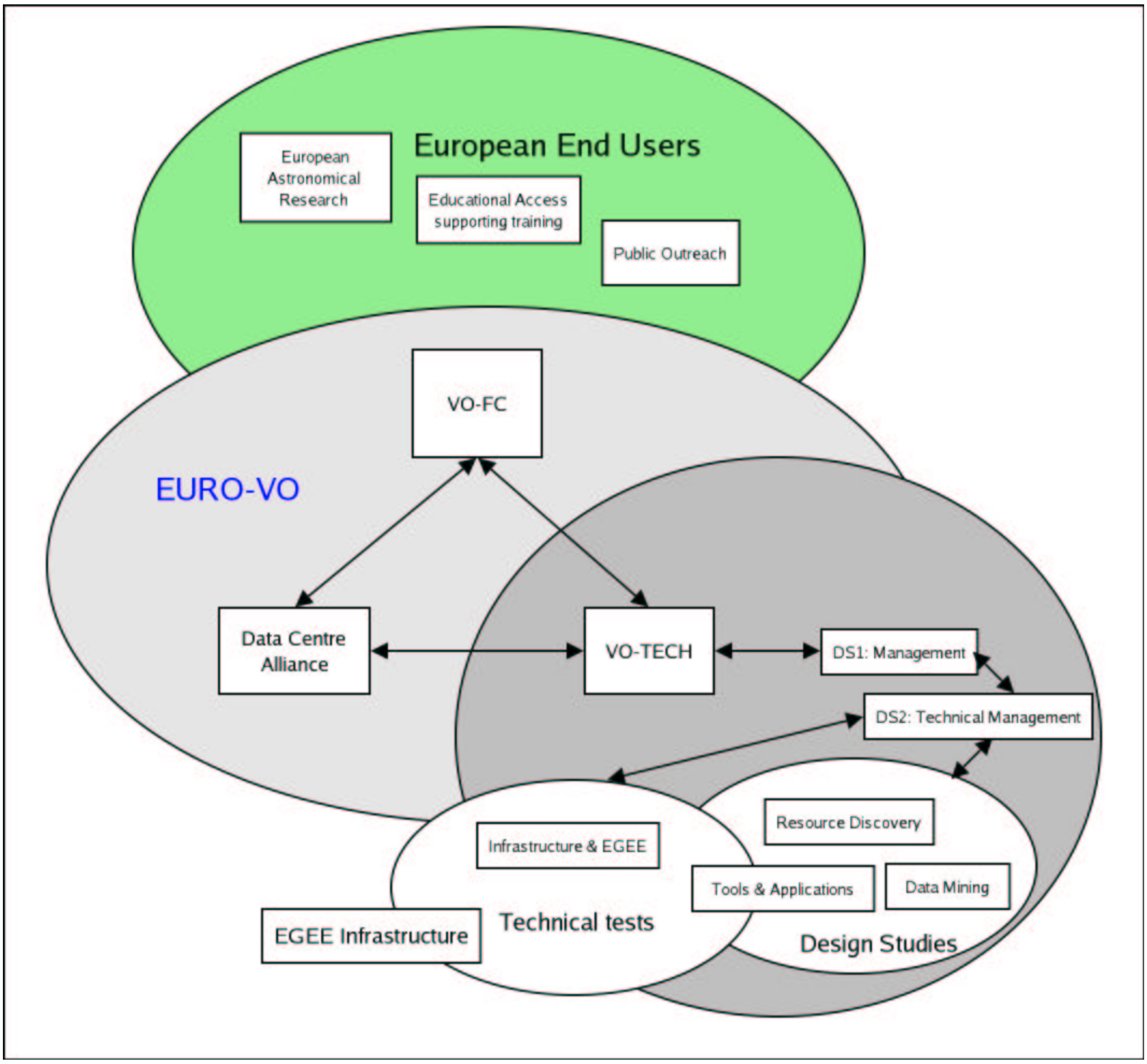
2.2 Quality of the proposed Design Study

2.2.1 Objectives of Design Study.

The **top-level objective** of the VO-TECH proposal is *to complete all technical preparatory work necessary for the construction of the European Virtual Observatory*.

(a) Context The VO-TECH preparatory work needs to link closely with the work of the Data Centre Alliance (DCA) and VO Facility Centre (VOFC), with final construction following our design study in mind. The work also takes place in the context of extensive developments - new algorithms, technologies, and protocols - in academic and commercial IT, and especially of course in generic grid middleware. We will not be repeating this work. Our job is to assess these developments and design astronomy-specific modules based on them. Our working links with this external world are excellent. Several of the VO-TECH partners have active working relationships with the academic and commercial IT communities, and we are expecting to work with EGEE as an exemplar application area. It is also worth noting that Astronomy in general and the VObs projects in particular have attracted attention as leading edge but pragmatic exemplars of the new e-science approach - for example some of the VO-TECH have ben invited to talk at Bio-Informatics meetings, as well as general Grid meetings.

Figure 1 shows the conceptual structure of VO-TECH, and how it relates to Euro-VO as a whole, the various classes of user, and the general astronomical infrastructure.



(b) Evolving Design. Construction of the necessary software infrastructure does not fit into the simple "waterfall" model that would describe construction of a hardware facility, the successive stages of which might be summarised as Vision and Concepts, Requirements Analysis, Functional Analysis, Initial Design, Detailed Design, and Construction. These same conceptual divisions occur, but in an iterative or rolling manner, with different components maturing at different times. There are three reasons why this approach is used. First, the technology is evolving very fast, so it is crucial to stay agile. This is true not just for the background technology, but also for the international VObs scene itself, which is constantly evolving. Second, the system is very novel, and user requirements sharpen dramatically as they see trial versions working and get the point - so it is vital to keep end-users in the loop and to iterate. The third reason is that practical experience has shown that software projects succeed much better when run this way - or inversely that "grand design" projects often fail.

(c) Iterative Approach. There are three implications of the above logic for the design study we propose. The first is that it is not a single study, but a connected sequence of studies, in several streams following the key technology areas described in section 2.1.4 - new infrastructure, tools, resource discovery, and data exploration. The second implication is that we take trial implementations very seriously, as an integral part of the design process. Results from such tests feed back into revised designs. The third implication is that we must pay serious attention to international standardisation during the design process - as we recognise what is required, we move these requirements onto the IVOA agenda, and complete final designs around agreed standards.

(d) General Objectives

1. To assess new technologies and study the feasibility of their incorporation in Euro-VO
2. To create designs of new infrastructure components based on those new technologies
3. To create designs of science user tools and datamining services
4. To develop trial versions of new infrastructure components, tools, and datamining services and to test them
5. To decide what new interoperability standards are required, and to define those standards with international partners
6. To liaise with the larger Euro-VO structure, gaining refreshed versions of science functionality and architecture, and feeding back component test results, designs, and trial components for demonstration suites.
7. To liaise with computer science, IT industry, and related applications projects in order to mesh with larger standards and to save work wherever possible

(e) Task Areas. Corresponding to the analysis of section 2.1.4, we separate our study into four substantive task areas, and two more providing co-ordination, integration, and technical leadership.

DS1 : Consortium management. The aim here is to provide the necessary administrative and financial support to the consortium, along with establishing an external presence, through web pages, seminars, workshops and so on. Web presence will be integrated into existing Euro-Vo structures (http://www.euro-vo.org). A training programme will be organised utilising workshops and on-line materials (interactive work throughs, help, FAQs).

DS2 : Technical Project Management. This task will provide technical leadership and planning, through two relatively senior positions - Technical Manager and Project Scientist. Both these will be partner-supplied positions at 50% rate, as they will both have responsibilities in the larger Euro-VO. An informal Design Reference Architecture will be developed to assist design of detailed components. This task also aims at technical co-ordination amongst the partners, developing common coding standards, development processes, software repository and version control, etc.

DS3 : New Infrastructure. This task aims at producing final designs of mature components, as well as assessments, designs, and trials of new components that don't fit into the major categories of DS4-6 below (most obviously grid services, workflow, agents, and distributed storage). In addition it has a responsibility for considering interoperability, integration and testing within the context of the overall Euro-VO architecture, and hence liaising with the VOFC. This will also include full internationalisation of the Euro-VO programme, designing customisation tools for deployment across Europe, and mix-and-match integration with other projects. DS3 will also have the prime responsibility for liaising with the NA4 workpackage of the FP6 programme Enabling Grids for E-Science in Europe (EGEE). We have had initial discussions, and astronomy is seen as an exemplar application area, and is planned for integration from 2005. (See letter of support in annex). The practicability of this depends heavily on input from the VO-TECH work programme.

DS4 : New User Tools. This task will produce designs for new VO-compliant end-user tools, in close collaboration with the VOFC and Science Working Group. We will produce our own list of suggested priority tools, but will also develop cases for original user-specified tools, by working with the VOFC, who will run a programme similar to AstroVirtel, where competitively selected users run real multi-archive science programmes, and effectively perform a gap analysis, requesting new tools.

DS5 : Intelligent Resource Discovery. This task aims at undertaking a feasibility study for developing components based on emergent technologies in the areas of the semantic web and ontologies. On the assumption that this is successful, we will proceed to actual component designs and trial implementations, and standards development.

DS6 : Data Exploration. This task will assess a range of datamining and visualisation algorithms and packages, with a view to assessing how they can be run as distributed services, how they can be made VObs-compliant, and how they can be extended to extremely large datasets. The functionality required for Euro-VO will be developed in parallel with the VOFC, and its feasibility with the DCA, who will need to deploy such services. On the assumption that these studies are successful, we will proceed to actual component designs, trial implementations and standards development.

2.2.2 Implementation plan.

We need to keep a balance between long term planning and clear goals on the one hand, and the need to stay agile and work iteratively on the other hand. We also wish to keep a balance between strong local management structures, and the "redeployable pool" approach that agile development argues for. Our solution is to plan for a sequence of semi-independent six-month long sub-projects, but to keep a strong sense of leadership, planning and co-ordination through DS2. We will establish a *Technical Advisory Panel (TAP)* which will meet every six months, review progress, and agree the sub-projects for the next semester, including named staff, leaders, and goals. These sub-projects will normally map onto the task areas, but for each semester the tasks will be at different stages with regard to feasibility study, trial implementations, standards work, and full design, so that work of each of these kinds is likely to be going on during each semester. In addition to the TAP meetings, we will have annual full meetings, where the Consortium Board will meet followed by an open all-staff meeting and a TAP meeting. These are likely to be co-ordinated with overall Euro-VO meetings, and will monitor overall progress and re-direct as necessary. The initial project plan is tabulated below.

month	meetings	management deliverables	task schedule			task deliverables
0	Kick-off	project plan				
1	TAP-1	task plans				
1-3		sci-fun doc	sci-anal			
3-6			studies-1			study report
7	TAP-2	task plans				
7-12			studies-2	tests-1		test + study reports
12	Full meeting	annual report				
13	TAP-3	task plans				test software
13-18			studies-3	tests-2	standards-1	test+study reports
19	TAP-4	task plans				new standards, test sw
19-24			studies-4	tests-3	designs-1	test+study reports
24	Full meeting	annual report				design docs, test sw
25	TAP-5	task plans				
25-30			arch study	tests-4	standards-2	test reports, arch doc
30	close-out review	close-out plan				
31	TAP-6	closing tasks				
31-36			designs-2		standards-3	new standards
36	final review	final report				final design docs

The plan specifies a series of meetings, and a schedule of deliverables. It deliberately staggers work on initial studies, trial deployments, and standards development work. This is necessary because different areas are at different stages of maturity, but also because the various tasks are actually strongly interlinked. The plan above does not specify which task areas are in which stage at a given point. This is deliberate, because things are evolving fast and we should take that decision during initial project kick-off planning. However, the likely pattern is fairly clear - grid services and workflow material should be ready for testing very quickly, whereas ontology will need significant work on assessment and feasibility study before going any futher. Data exploration is mature in its own terms, but needs initial study of how to implement in a distributed fashion, and is likely to be one of the first areas ready for standards development.

Specific institutions are responsible for taking the lead in particular task areas, but we have selected teams that cut across the institutions. This is necessary to keep broad experience and insight into the system, but it is also a specific aim for us to collaborate and pool skills and experience across Europe. Although the effort per task area will not change substantially, the named individuals on each team may not be fixed. Who works on what will be for the partnership to agree each semester through the TAP. However in this they will be strongly guided by the Technical Manager and Project Scientist, who will bring proposals to each TAP.

If the proposal is funded, then during the contract negotiation stage, and before project kick-off, we will produce a *Project Plan*, which will include a more detailed and concrete set of milestones and deliverables for the first eighteen months. We believe it would be a mistake to produce such a detailed plan now, given the constantly evolving context.

3. RELEVANCE TO THE OBJECTIVES OF THE SCHEME

3.1 Justification of the proposed Design Study

3.1.1 Scientific Need. The Design Study proposed here is a crucial step in bringing about the the Virtual Observatory vision outlined in section 1 - all the world's databases on your desk, analysable by the same tools, and with powerful new tools for data intensive exploration and analysis. This new way of doing things is obviously very desirable, but also necessary, for several reasons. (a) Every new development, in both web technology in general and astronomical data archives in particular, is raising the expectations of astronomers with respect to the quality, power and flexibility of data services. There is a danger that European astronomers will fall behind their international colleagues. (b) Even current styles of access and analysis will become problematic as both data volume and heterogeneity increase over the next few years - we will drown in data. Although we assume that Moore's law will continue, data volumes are increasing faster than this, and key measures - disk-CPU bandwidth, and last-mile network bandwidth - are increasing more slowly. This bottleneck demands the data archive centred VObs style of working. (c) Data intensive kinds of analysis, manipulations of millions or billions of data points, has already produced important science, but has so far been limited to occasional "power-users". There is considerable demand from astronomers in general to have easy access to such powerful tools. (d) The Design Study itself is not just a technical fix - it needs to be carried out keeping real science aims and functionality inside the loop, especially given the evolutionary nature of the VObs development. We have built in interactions of the Project Scientist with the VOFC, and maintaining liaison with the general Euro-VO Science Working Group.

3.1.2 Technological Need. Our trial implementations so far have made impressive progress. However we are a long way from achieving the full vision of Euro-VO. (a) In many cases we have a clear idea of the new components we need, but considerable *technical preparatory work* is needed before implementing them - designs and tests. (b) The most exciting possibilities depend closely on quite new technologies. These will not just "plug-in" - there is a clear need for *feasibility study* followed by more *technical preparatory work*. (c) We are not aiming at an old-fashioned scientific lash-up. We aim at building an infrastructure of lasting value and international quality. It must be robust, professional, and flexible, as well as useful. This is particularly true because of the modular nature of the infrastructure, so that for example Chinese astronomers can use components of our infrastructure, along with other components from the US and others they have built themselves. All this requires development to product engineering standards, which in turns mean that a properly planned and executed Design Study is absolutely crucial.

3.1.3 Deliverables. At the end of the Design Study, we should have the following deliverables.

- a series of study reports in the areas of grid services, the semantic web, ontology, datamining, visualisation, agents, workflow, and distributed storage
- a final architecture design for Euro-VO
- a series of design documents for selected tools
- a series of design documents for new infrastructure components
- a series of internationally agreed astronomical interoperability standards in all necessary areas (ontology, workflow, etc)
- trial implementations of new infrastructure components

- interface specifications, to allow external projects to use components, data centres to publish data, and for and external user development of new tools and services

In collaboration with VOFC and DCA, we will also contribute to further deliverables :

- a study assessing the financial implications of Euro-VO in terms of construction and operations
- design document for the technical operation of user support and training for the Euro-VO
- Euro-VO project plan

3.1.4 Expected Users. The immediate users of the Design Study are the other parts of the Euro-VO structure - the VOFC and DCA, with whom we will collaborate in the final construction phase. We also expect significant interest and uptake from several clases of users of the resulting infrastructure. (1) The very large community of astronomical end-users. If we are successful, we expect that the Euro-VO infrastructure, the datasets that populate it, and the tools available through it, will become an integral part of the daily life of almost every astronomer. It will become the normal way science is done. (2) The community of professional data centres. They will need to structure their archives in a Euro-VO compliant manner, install the necessary components, and construct and publish data services. (3) The data creation facilities, who will henceforth build every new telescope, every new instrument, with Euro-VO in mind. (Realisation of the importance of this is already becoming clear in talks at conferences by instrument builders). As well as the major facilities, this includes consortia of astronomers who construct major new scientific data sets using those facilities.(4) Science tools writers. Some of this activity will go on inside data centres, some in specialised groups developing new data mining algorithms and so on, and some will be undertaken by interested individuals. Tools will also include major theory tools - simulations, photo-ionisation codes, etc. They will all henceforth write such tools with Euro-VO in mind. All the above applies across wavelength regimes and all types of astronomical endeavour. (5) The educational system and the general public - easy access to the best data and tools at students's desktops.To maximise the effectiveness of this will require a specially designed portal. The Euro-VO VOFC will develop an outreach component to address this.

3.1.5 Licensing Policy. The Euro-VO, in common with other VObs initiatives, is working in the public domain, publishing their products into the public domain, using open-source principles. Reports resulting from this design study will be published by the Euro-VO. Software elements produced during the design study will be released under the IVOA Public License (which is currently under development).

3.1.6 Enhancement of existing infrastructures. VObs technology is generic. It will enhance almost all existing astronomical infrastructures. In particular use of the archived data in European data centres will be greatly increased and of higher value. This will continue to be true as the new data creation facilities place further data in those archives. Essentially this completes the promise and maximises the return on investment for the very expensive pieces of hardware already constructed or under construction for European astronomy - telescopes, instruments, and spacecraft. One of the most exciting prospects is closing the loop between data analysis and data creation. Astronomical facilities design with Euro-VO in mind, and establish their data in VO-compliant data centres; users query a variety of such datasets, and when this exposes the scientific requirement for new data, they can be linked directly to a telescope time application system, including instrument simulators and so on.

A significant factor in Euro-VO is *democratisation of access*. You don't have to be in Paris or Cambridge to have access to the best data, or to PetaByte scale data storage, or to the CPU power necessary to calculate a galaxy correlation function, and you don't have to write your own correlation function code - its all available, documented and robust. A scientist of equal talent will be equally effective anywhere. This will be a powerful encouragement for the integration of astronomy communities in new countries soon to join the EU. The value will also spread more widely through the science community and the public, through educational access to the best data and tools. Astronomy has a strong track record in development and use of the Internet (ESO hosted one of the very first web servers), and attracts wide interest, encouraging the diffusion of new ideas.

Finally, we envisage that new kinds of human infrastructure will emerge. The transparency and democracy of access, and the ability to share material over a secure distributed system, will encourage the growth of Virtual Organisations and informal *collaboratories* - groups of users can not only see the same data but can store and share derived results, workflow and log files, their own deposited material, working documents and so on, making these elements securely available to a defined list of authenticated individuals.

3.1.7 Measures of Success. We will actively monitor success. The first indicator will be whether or not the new infrastructure is taken up by European data centres. There are around two dozen such key sites, so it will be clear whether this takes place. The second indicator will be the rate of use of datasets published as VObs services, as well as tools, and portals. We will ask the data centres to monitor this for us. Finally, we will encourage end-users wherever possible to make clear in published papers when they have used the new infrastructure to perform their analysis.

3.1.8 Project Risks. The main risks to delivery of the Design Study objectives are tabulated below, together with their relative probability of occurrence, their impact on the project, and our plan for mitigation.

Risk	Prob.	Imp.	Mitigation Plan
International community fail to agree standards	1	3	Series of specialist meetings; agree international roadmap; make standards work explicit part of work programme
Small staff volume per task; sensitivity to illness/loss of key individuals	2	2	Modular approach to minimise single point failures; spread tasks across team.
Requirements drift	2	3	VOTC Project scientist has overall oversight; keep end-user community engaged;
Infrastructure software components and/or tools fail to meet requirements	1	3	VOTC project scientist monitors progress; iterative design and test process.
External technologies and standards fail to mature or stabilise in time	2	2	Modular approach to minimise lock-in; iterative approach to test and design cycle; maintain active links with wider computer science and Grid world
Infrastructure overtaken by external events in other projects	2	2	Modular approach to minimise lock-in; maintain working links with other projects especially US-VO
Programme too ambitious	3	2	Iterative approach to design and test cycle; oversight by Euro-VO Executive Board; engage both scientists and developers in design and test cycle.
Better components and/or tools developed elsewhere	2	1	Participation in international standards programme; keeping active links with other projects, application areas and computer science to identify such possibilities; if other components are better - use them !

3.2 Exploring the feasibility of the infrastructure

3.2.1 Effort for the Design Study. The spirit of the Euro-VO project overall and the VO-TECH project in particular is to utilise existing funding as far as possible, but to add EU funding both to add to the total effort and specifically to foster Europe-wide collaboration and uniform standards. All the partners have therefore committed staff effort from existing funding. This requirement for commitment also guarantees that the partners involved in VO-TECH are serious and experienced players in the field. They all have both track record and commitment. AstroGrid is one of the three major VObs projects worldwide, but sees its role increasingly in European commitment. The consortium as a whole has considerable resource funded by PPARC over 2005-2008. The CDS has been the prototype VObs centre for many years, and has a long term commitment in this area. ESO is the most important centre in European ground-based astronomy, and has set new standards for ground-based archives. It sees Euro-VO as central to the future of international astronomy. Finally INAF represents both Italian VObs and Grid projects, but also a particular interest in Trieste and Napoli for the importance of data mining and the application of new algorithms for astronomy. Thus we have assembled the key team in terms of both skills and motivation for solving the technical issues for Euro-VO.

3.2.2 Project Plan for Euro-VO. The Euro-VO partnership will be constructing an overall plan for implementing the vision, and funded VO-TECH activities are a crucial element in this. As well as building towards well understood science requirements and functionality, the VO-TECH programme will recommend a Technical Architecture design. The full partnership however needs to understand the implications of this for technical operations, component maintenance, and user support, as well as the effort required for the DCA to implement the infrastructure and publish services. We will aim at a full understanding of the financial implications of the construction and operation of Euro-VO, as well as a project timeline.

3.2.3 Euro-VO partnership and funding. The Euro-VO partnership includes the VO-TECH partnership but is considerably wider. Currently it includes the European Space Agency (ESA) as well as national projects and organisations from Germany, Spain, and the Netherlands. We are building towards a Memorandum of Understanding by the middle of 2004, in which all partners commit some level of staff effort to one of VOFC, DCA, or VOTC, using funding from national projects and organisations as well as subscription funding from ESO and ESA. Following this we will be examining further opportunities for EU funding. To date we have a limited amount of funding for interoperability meetings, as part of the OPTICON project. We will also continue to develop possibilities for further members of the Euro-VO partnership. (New national projects have begun since the start of the Euro-VO programme).

3.2.4 International context. To make Euro-VO a success, it must be consistent with all international developments. The mechanism for ensuring this is the International Virtual Observatory Alliance (IVOA) , which acts as both a forum for discussion and exchange of best practice, but also as a formal standards development body. The partners in this proposal have already played a large part in creating and running the IVOA. An integral part of our project plan involves elucidating required international standards, and actively taking them onto the IVOA agenda.

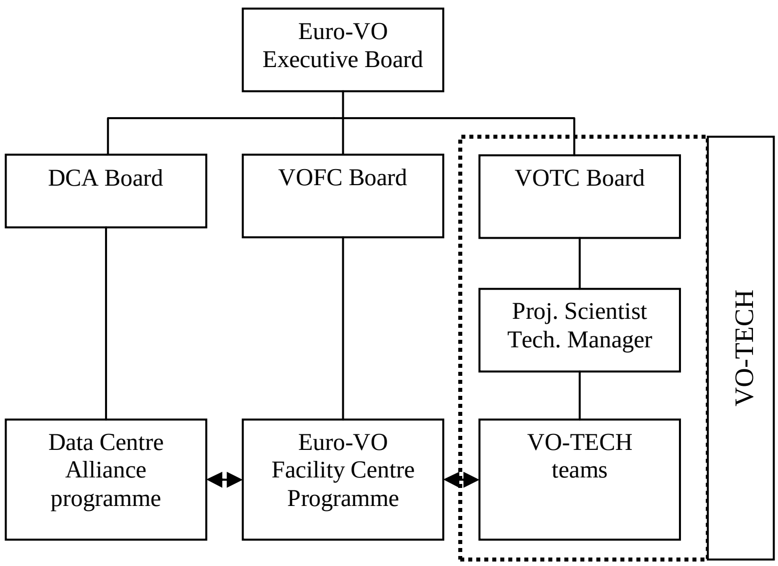
3.2.5 Wider IT and e-science links. Some of the problems we address, and solutions we envisage, are quite specific to astronomy. Others of course are familiar in a wide range of disciplines, or are being addressed within the academic and commercial computing worlds. The rapid evolution in these areas is precisely why our project plan is iterative and evolutionary. We will conform to standards and save effort by deploying know solutions and actual software wherever possible. All the partners have the necessary strong links. AstroGrid is part of a co-ordinated UK e-science programme, with links to GridPP, OGSA-DAI, and MyGrid. Both AstroGrid and French VO cover solar physics and space plasma physics in their remits. CDS participates in projects funded nationally in the frame of Action Concertées Incitatives (ACI) GRID and Masses de Données (Massive Data), in close collaboration with French IT laboratories. The Italian data mining work is explicitly in collaboration with local computer scientists. Through AVO and AstroGrid we have established an Astro Research Group within the Grid Global Forum; and we have made contact with EGEE and are likely to develop astronomy as an "EGEE application area". Members of the partnership also have working relationships with key individuals in CERN, Microsoft, IBM, and Sun. As well as generic grid middleware, and computer science developments, other science application areas are working on similar problems - especially Bio-informatics, Earth Science, and Particle Physics. In AstroGrid, we have made a point of talking to such projects and attending each other's meetings.

3.2.6 Political and funding context. The partners in Euro-VO believe that the VObs infrastructure is central to the future of astronomy. A pulse of funding is needed now to design and create the infrastructure, but it has long term implications once the bar has been raised in this way. An implication is that funding agencies need to explictly consider the cost of supporting a healthy data centre infrastructure, as well as a high level body such as ESO taking some infrastructural long term responsibility. Also, every new mission or observatory needs to plan in the context of the VObs and the new standards it implies. Through AVO and this new proposal, ESO is already making an initial commitment. ESA is also exploring these issues for Space Science more generally through the Space Grid study (see http://www.gridtoday.com/03/00616/101542.html). Some of the partners (A.Lawrence and P.Quinn) are currently involved in an OECD Global Science Forum study looking at the future roadmap for large scale facilities in astronomy world wide, with a specific remit to analyse and promote such issues. Similar issues recur in other science areas of course. P.Quinn is taking these issues forward at the level of European science treaty organisations through a EIROFORUM thematic working group (http://www.eiroforum.org).

4. QUALITY OF THE MANAGEMENT

4.1 Management and competence of the participants.

4.1.1 Euro-VO structure. The VO-TECH project fits within the larger Euro-VO programme, which contains three linked parts - the DCA, VOFC, and VOTC. Each of these activities is managed by its own Board, but the whole programme is overseen by a three person Euro-VO Executive Board (Lawrence, Quinn and Genova : VO-EXEC). The VO-EXEC ensures that close liaison is kept between VOFC, DCA, and VOTC,and in particular the VOTC Project Scientist and Technical Manager also report to the VOFC Board.

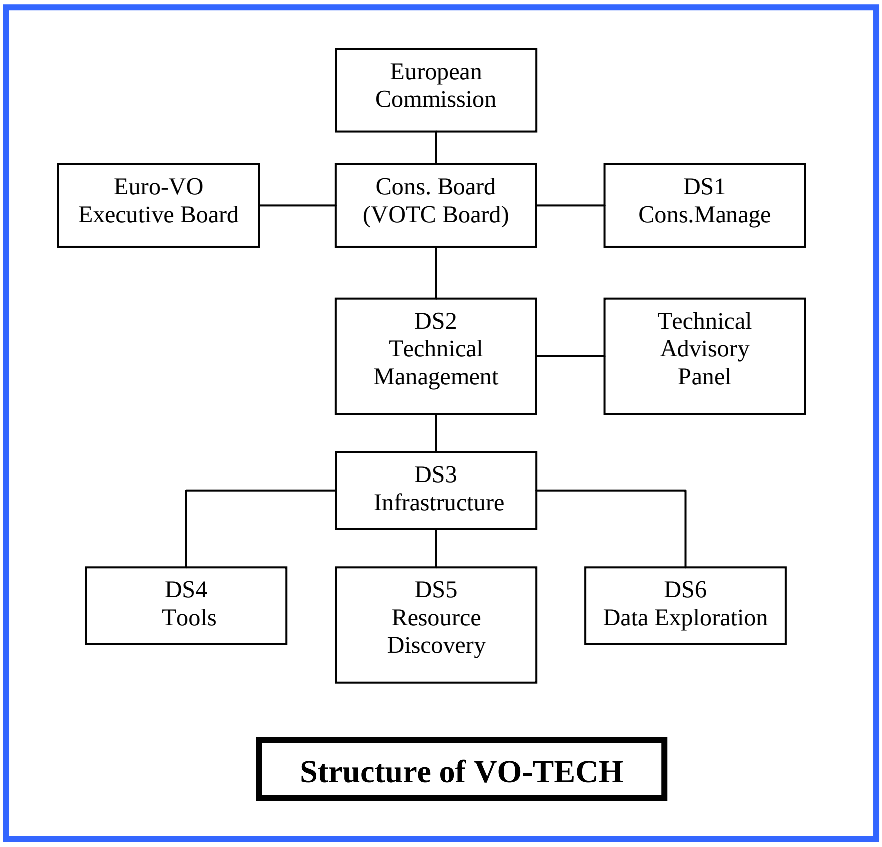


Structure of Euro-VO

4.1.2 VO-TECH Project Co-ordination and Oversight

Project co-ordination will be undertaken by A.Lawrence, at the University of Edinburgh. He will be responsible for all communication with the Commission on contractual matters. He will be supported by a professional administrator and by the Finance Department of the University of Edinburgh who have dedicated teams supporting each University College. They will be responsible for processing payments to partners and will assist in compiling information required for completing cost statements. They have direct experience in the financial management of EU projects. The University of Edinburgh under the Framework 5 Programme collaborated in some 220 projects. The co-ordinator, administrator and web developer will together provide both an external presence and an internal knowledge management system.

We will form a **Consortium Board** composed of the named investigators from each partner, supplemented by an administrative member if and when necessary. For the duration of the VO-TECH project, the Consortium Board is the same as the "VOTC Board" in the Euro-VO structure above. The purpose of the Consortium Board is oversight of the project - its setup, financial monitoring, resolution of issues between partners, and overall scientific and technical policy. It does not run the project from week to week. This is done through the VO-TECH project management team (see below). The full Consortium Board meets annually.



4.1.3. VO-TECH Project Management.

Two key senior staff are responsible for planning and co-ordinating the VO-TECH programme : the *Project Scientist (PS)* and the *Technical Manager (TM).* Both of these are AstroGrid-supplied staff positions devoting half their time to VO-TECH. They will prepare long term work plans, Gantt charts, and budgets for VO-TECH, which will be revised every six months, with the assistance of the Technical Advisory Panel (TAP - see below). They report to the Consortium Board. The PS additionally has responsibility for gathering science requirements and analysing required functionality, and for liaising with the user community, the Data Centre Alliance, and external projects, e.g. EGEE, Planck, etc. The TM, assisted by a senior software developer, will also lead coding standards and development processes for the whole VOTC programme, maintaining a code repository, code integration, and technical aspects of the project knowledge management system in conjunction with the co-ordinator. Finally the TM will oversee and co-ordinate the sub-projects (see below). The PS and TM will work closely with the other components of Euro-VO and the user community, elucidating requirements and feeding back designs and trial components for integration into the final Euro-VO infrastructure.

.

4.1.4 VO-TECH Programme Planning.

Programme planning is undertaken by the TM and PS, with the assistance of the *Technical Advisory Panel (TAP).* The TAP is a small panel containing a mixture of academic investigators and technical staff. It's members are appointed by the Consortium Board, with advice from VO-EXEC. It meets every six months, receives and discusses reports from the TM and PS, and agrees the VO-TECH work plan for the following six month period.

Our aim is to achieve a balance between strong local teams and Europe-wide coherence, and between sound long term planning and the agile adaptability needed when working with a background technology that is continuously evolving. The overall long term plan is presented in section 2. Within this plan we will work on the basis of six month iterations. The TAP will meet twice a year and debate the work programme for the next six months, based around proposals brought forward by the TM and PS. A small set of discrete sub-projects will be agreed, each with named workers and a workgroup leader, and distinct goals to be achieved by the next TAP. The sub-projects will have relative autonomy over the following six months, agreeing their own working practices and so on, but the PS and TM will monitor progress, and all organisations will make use of the common coding standards and development processes we have evolved. The sub-projects will normally divide amongst the defined Design Study Tasks. Given these tasks, and given the experience of the partners, we have an expected distribution of effort versus task and partner - see below.

4.1.5 Competence of the partners.

The VO-TECH programme is relatively specialised, requiring key skills and experience and specific motivation, and so does not involve all the partners involved in Euro-VO overall. We have also limited the partnership to partners ready to commit additional staff effort from existing funding sources. Letters confirming this additional staff effort are available on request.

- The co-ordinating partner is the **University of Edinburgh**. The co-ordinator (A.Lawrence) is AstroGrid Project Leader, directs data centre activities through the Wide Field Astronomy Unit, and is a member of the Euro-VO Exec Board. The University hosts the UK National e-Science Centre (NeSC), which has a global profile, and there are strong interdisciplinary links between astronomy, particle physics, and informatics. Edinburgh therefore has responsibility for DS1, consortium management.
- Technical leadership of the VO-TECH programme will be provided by the **AstroGrid Consortium**, represented here by the Universities of Cambridge, Edinburgh and Leicester. The consortium runs one of the main three VObs projects worldwide and currently has a team of more than a dozen experienced software developers. AstroGrid is an existing partner of the AVO project, with lead responsibility for the Technology work area. AstroGrid has secured substantial continuing funding from the UK e-science programme and is able to commit 6 FTEs of effort to VO-TECH, and in particular, will provide the functions of Project Scientist and Technical Manager. The competence of Edinburgh is discussed above. Leicester is the Technical Management centre of the AstroGrid project. **Cambridge and Leicester** are both well known astronomical centres, run high profile data centres, and have significant cross-disciplinary e-science activities. AstroGrid overall has lead responsibility for DS2, Technical Project Management, and DS3, new infrastructure components.
- The **European Southern Observatory** (Munich) is an intergovernmental treaty organisation and is Europe’s leading organisation for ground based astronomy. It is the key link to data creation facilities and their requirements. ESO's Data Management Division (DMD) has set new international standards for ground-based archives. Peter Quinn is the Director of DMD, AVO Project Leader, currently chair of the International Virtual Observatory Alliance, is a member of the Euro-VO Exec Board, and leads the Euro-VO project overall. ESO also has experience in developing end-user tool requirements, and delivering such tools, through the FP5 AstroVirtel programme. ESO is therefore expected to take an especial interest in DS4, new tools.
- The **Centre National de la Recherche Scientifique**, which manages the newly created Action Specifique Observatoire Virtuel France, acts on behalf of the French VO, which is

rrepresented here by CDS, led by F. Genova. CDS is the world leading centre for astronomical catalogue data, with a long track record, and a strong team of software engineers and specialized astronomers. CDS got the international VO interoperability standard definition effort off the ground, first as a Working Group of the FP5 OPTICON network. It is a partner in the AVO project, with lead responsibility for the Interoperability Work Area. F.Genova is a member of the Euro-VO Exec Board. Other participants in the French VO also have a high expertise, e.g. Terapix in data exploration methods. CDS will commit 3 FTEs of effort from their team. CDS is a Joint Research Unit of CNRS and Université Louis Pasteur. CNRS is the project partner and Université Louis Pasteur will act as a third party under Clause 23. CDS will take a special role in DS5, automated resource discovery.

- The **Istituto Nazionale di AstroFisica** is an institution located in Rome which co-ordinates Italian research in astrophysics. Its activities related to the development of a national astronomical grid and virtual observatory (the DRACO project) are co-ordinated by Fabio Pasian. The work for this project will take place in the Trieste and Naples areas; INAF will commit 2 FTEs of effort, and is expected to take an especial interest in DS6, data exploration.

4.1.6 Distribution of Effort.

Consortium Management will be undertaken by the co-ordinator. Technical Project Management will be shared by the partners through the Technical Advisory Panel, but run day to day by AstroGrid staff. The substantive tasks, DS3-6 will be on average distributed evenly across the partners, but with special interests as indicated above. The table below summarises the staff effort distribution, in staff months integrated over the project life. The bulk of the table shows total staff effort, including both partner provided effort, and the effort that will be funded by the contract requested here. The final line shows the effort requested in this proposal.

STAFF EFFORT MONTHS	Coordinator	AstroGrid	ESO	French-VO	INAF	task total
DS1 Consortium Management	Sectry 18 Admin 18 WebDev 18	PI time 9	PI time 3	PI time 3	PI time 3	72
DS2 Technical Project Management		ProjSci 18 TechMan 18 SenDev 36				72
DS3 Infrastsructure Components		Senior 36 Junior 108	Junior 36	Junior 36		216
DS4 New User Tools		Junior 36	Senior 36 Junior 36	Junior 72	Senior 36	216
DS5 Intelligent Resource Discovery		Junior 72	Junior 36	Senior 36 Junior 36	Junior 36	216
DS6 Data Exploration		Senior 36 Junior 72		Junior 36	Junior 72	216
Partner total	54	441	147	219	147	1008
Partner requested	**54**	**216 (AC)**	**144 (FCF)**	**108 (FCF)**	**72 (AC)**	**594**

4.2 Justification of financing requested

The funds requested cover three items : staff salaries, personal computing equipment, and travel. Each of the partners has committed effort funded from other sources, so that salary costs are those needed for additional effort to complete the team. Travel and personal equipment is however requested to cover all team members. Note that all the costs detailed below are before applying a standard 20% overhead.

4.2.1 Types of Staff. Most of the staff will be professional software developers, although some will be scientific Post Doctoral Research Assistants (PDRAs). Some of these can be relatively junior, but in each task area, and in each partner institution, we need at least one Senior Developer or Scientist. As well as the fact that such people are usually very productive, we need a pool of such experienced staff to act as Team Leaders for the sub-projects set twice-yearly. The Technical Manager and Project Scientist need to be very senior and experienced staff, but these will be supplied from existing AstroGrid personnel.

4.2.2 Staff cost rates. We have agreed standardised staff rates based on actual typical UK rates on Edinburgh University pay scales. A *Junior PDRA/Developer* is taken to be at scale point AR1A.8, at at a salary of £25,451, plus employer on-costs of 25%, converted to euros at a rate of 1.52, making €48,357. This will have some variation across Europe, but is a reasonable estimate. We therefore use a standard figure of €49,000/year. For a *Senior PDRA/Developer*, we assume someone on scale point AR2.8 at a salary of £33,679. Based on this, we use a standard staff rate of €64,000/year. For UK partners, who enter under the AC cost model, we use these two staff rates directly. ESO enter under the FCF model, and have a standard average organisational staff rate, regardless of grade, which is €101,900. CDS and INAF also enter under the FCF model. They do not have a fixed organisational rate, but estimate their full costs at 100% above additional employer costs. We have therefore calculated the cost of staff at CDS and INAF using the above standard consortium rates. Note that for the UK partners, the committed partner funding comes in as whole FTEs at zero cost, and the EU funding as requests for whole FTEs at 100% of additional cost. For the other partners, the committed partner funding comes in as half-FTEs of all staff, and the EU funding as requests for the other half-FTEs of all staff, at 50% of full cost.

4.2.3 Equipment Costs. Modest equipment is required, for development purposes. (Deployment of the final Euro-VO software by data centres on more substantial equipment is a separate issue). Each new EU funded project team member requires personal computing equipment, which we budget as a one-off purchase at €5000. In addition, each institution requires a dedicated development machine for testing purposes. This is envisaged as a mid-level PC or workstation, budgeted at €5000. (The development machine costs are put under DS2).

4.2.4 Travel Costs. As far as possible, we will be using common information, document, and software management systems to make such a distributed project viable. Nonetheless, at least some travel to each other's institutes is essential, even for developers - to brainstorm ideas, to talk through key issues, to understand each other's working practices, and simply to build a team spirit. In addition, there will be travel to formal full consortium meetings and TAP meetings. We envisage each team member (including both EU-funded and partner-contributed staff) travelling to another partner organisation on average five times per year. We budget these trips at €1000 each, assuming an average airfare of €500, and a five day trip with expenses at €100/day.

4.2.5 Cost Tables for each Task

DS1 : Consortium Management. The main request is for administrative and financial support, which will be provided as partial staff effort from the Research Support and Finance sections of Edinburgh University. We cost this based on standard UK pay scales. For the secretary we assume 50% of a *Secretary* at scale point CN3.1, at a salary of £12,997, plus employer on-costs of 25%, converted to euros at a rate of 1.52. For the *Administrator* we assume someone at scale point AA2.1, on a salary of £21,125. In addition we request 50% of a Web Developer at the same rate as the Administrator, to develop a web presence for the project, and maintain cohesion between partners. We also ask for general costs - cost of publication production, dissemination of information, running workshops etc, which we estimate at €3000/year. All the above funds will be held by the co-ordinator. However we also add a standard cost for at least one audit for each partner, at €3200. We also list here the estimated time of academic investigators at zero salary cost.

DS1 **Organization**	**Partner-funded Personnel (months)**	**EU-funded Personnel (months)**	**Personnel cost (Euro)**	**other cost (kEURO)**
Astrogrid:Edin(AC) PI time Secretary Administrator Web Developer	 3 0 0 0	 0 18 18 18	 0 37041 60177 60177	 9000 gen.costs 3200 audit
AstroGrid:Leic(AC) PI time	 3	 0	 0	 3200 audit
AstroGrid:Cam(AC) PI time	 3	 0	 0	 3200 audit
France-VO (FCF) PI time	 3	 0	 0	 3200 x 50% audit
ESO (FCF) PI time	 3	 0	 0	 3200 x 50% audit
INAF (FCF) PI time	 3	 0	 0	 3200 x 50% audit
TOTALS	**18**	**54**	**157395**	**23400**

DS2 : Technical Project Management. The Technical Manager and Project Scientist will be shared with the AstroGrid project. In addition we request a Senior Developer to assist the top-level co-ordination team, concentrating on integration and testing issues. In this task we also request a development server for each site. We also request travel funds two meetings per year for the academic investigators, to participate in the project planning process.

DS2 **Organization**	**Partner-Personnel (months)**	**EU-funded Personnel (months)**	**Personnel cost (kEuro)**	**Equipment cost (kEURO)**	**Travel cost (kEURO)**
Astrogrid : Edin (AC) Senior staff development server	0	36	192	5 5	21
AstroGrid: Leic (AC) Technical manager development server	18	0	0	0 5	21
AstroGrid : Cam (AC) Project Scientist development server	18	0	0	0 5	21
France-VO (FCF) development server	0	0	0	5 x 50%	6 x 50%
ESO (FCF) development server	0	0	0	5 x 50%	6 x 50%
INAF (FCF) development server	0	0	0	5 x 50%	6 x 50%
TOTALS	**36**	**36**	**192**	**27.5**	**72**

DS3 : New Infrastructure. Each of the substantive task areas are planned to work with a total team of six, with at least one senior developer, which in this case we place in Leicester, who will lead this task. The infrastructure task is central to everything else, so we request an additional senior developer here, to be placed in Cambridge. Additional partner-funded staff will come from AstroGrid.

DS3 **Organization**	**Partner-Personnel (months)**	**EU-funded Personnel (months)**	**Personnel cost (kEuro)**	**Equipment cost (kEURO)**	**Travel cost (kEURO)**
Astrogrid : Edin (AC) Junior staff	36	0	0	0	15
AstroGrid: Leic (AC) Senior Staff Junior Staff	0 36	36 0	192 0	5	15 15
AstroGrid : Cam (AC) Senior staff	0	36	192	5	15
France-VO (FCF) Junior staff	0.5x36	0.5x36	147 x 50%	5 x 50%	15 x 50%
ESO (FCF) Junior staff	0.5 x 36	0.5 x 36	152.85 x 50%	5 x 50%	15 x 50%
INAF (FCF)	0	0	0	0	0
TOTALS	**90**	**126**	**533.925**	**15**	**75**

DS4 : New User Tools. Again, we have a team of six in total, three of which are requested here. ESO will lead this area, and therefore we will expect to place a Senior Developer there, although it makes no difference to the costing as ESO have a standard rate.

DS4 **Organization**	**Partner-Personnel (months)**	**EU-funded Personnel (months)**	**Personnel cost (kEuro)**	**Equipment cost (kEURO)**	**Travel cost (kEURO)**
Astrogrid : Edin (AC)	0	0	0	0	0
AstroGrid: Leic (AC)	0	0	0	0	0
AstroGrid : Cam (AC) Junior staff	0	36	147	5	15
France-VO (FCF) Junior staff	0.5x72	0.5x72	294x50%	0	30 x 50%
ESO (FCF) Senior staff Junior staff	 0.5 x 36 0.5 x 36	 0.5 x 36 0.5 x 36	 152.85 x 50% 152.85 x 50%	 5 x 50% 5 x 50%	 15 x 50% 15 x 50%
INAF (FCF) Senior staff	0.5 x 36	0.5 x 36	192 x 50%	5 x 50%	15 x 50%
TOTALS	**54**	**162**	**542.85**	**12.5**	**72.5**

DS5 : Intelligent Resource Discovery. Once again, we have a team of six in total, three of which are requested here. France-VO will lead this area, and we therefore request a Senior developer there.

DS5 **Organization**	**Partner-Personnel (months)**	**EU-funded Personnel (months)**	**Personnel cost (kEuro)**	**Equipment cost (kEURO)**	**Travel cost (kEURO)**
Astrogrid : Edin (AC) Junior staff	36	0	0	0	15
AstroGrid: Leic (AC) Junior staff	0	36	147	5	15
AstroGrid : Cam (AC)	0	0	0	0	0
France-VO (FCF) Senior staff Junior staff	 0.5x36 0.5x36	 0.5x36 0.5x36	 192 x 50% 147 x 50%	 5 x 50%	 15 x 50% 15 x 50%
ESO (FCF) Junior staff	0.5 x 36	0.5 x 36	152.85 x 50%	5 x 50%	15 x 50%
INAF (FCF) Junior staff	0.5 x 36	0.5 x 36	147 x 50%	5 x 50%	15 x 50%
TOTALS	**90**	**126**	**466.425**	**12.5**	**60**

DS6 : Data Exploration. Once again, we have a team of six in total, three of which are requested here. Edinburgh and INAF will jointly lead this area, and we therefore request a Senior developer at each.

DS6 **Organization**	**Partner-Personnel (months)**	**EU-funded Personnel (months)**	**Personnel cost (kEuro)**	**Equipment cost (kEURO)**	**Travel cost (kEURO)**
Astrogrid : Edin (AC) Senior staff	0	36	192	5	15
AstroGrid: Leic (AC) Junior staff	36	0	0	0	15
AstroGrid : Cam (AC) Junior staff	36	0	0	0	15
France-VO (FCF) Junior staff	0.5x36	0.5x36	147 x 50%	5 x 50%	15 x 50%
ESO (FCF)	0	0	0	0	0
INAF (FCF) Junior staff	0.5 x 72	0.5 x 72	294 x 50%	10 x 50%	30 x 50%
TOTALS	**126**	**90**	**412.5**	**12.5**	**67.5**

5. Other issues

5.1 Gender Issues

A driver in the development of the Euro-VO will be increasing access to high class research facilities. In carrying out the design and technical preparatory work assessing the development and deployment of the infrastructure and components that will form the European level VObs infrastructure, one factor will be considering possible gender issues.

A key area to note is that the VObs design allows access to research tools and data irrespective of location. This coupled with the growth of home broadband networking, enables flexibility of use from home, enhancing options and flexibility when considering work/life balance.

The design studies will consider the effects of gender for instance on user interface to the developed product. This will ensure that the end user system supports the differing needs to the research community.

Appropriate reference will be made to the recommendations resulting from EU FP5 studies into the effect of gender in the development and use of information systems.

5.2 Ethical Issues

No issues requiring consideration by ethical committees are raised by the activities proposed in this programme.

In terms of data protection requirements, any information required to be held by the system, for instance in developing the Euro-VO authorisation system, will be in full compliance with national and EU legislation (Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data).

5.3 Outreach

The Euro-VO is primarily targeted at supporting astronomical research. However, it will have a significant impact in a number of associated areas, not least Outreach.

The use of AstroGrid to provide better access to astronomy data to support the teaching of science and other subjects is currently being explored. Relevant areas here of interest include the evolving metadata descriptions of the data resources which open up the possibility to 'tag' data that is of relevance to various target groups - thus depending on who you were you would automatically be pointed to relevant data resources. An example might be that a 12 year old child searching for information on the moon would be presented with a completely different set of information than the same query performed by a 16 year old.

Through the internationalisation design aspects through VOTECH it is anticipated that the relevance will be increased to cover possible educational access uses across Europe. In terms of inclusion of the interested amateur, and the public, the specific public Euro-VO interface portal would provide access to material especially of relevance to these communities. For the more interested public astronomers, they would be able to provide their observations to the professional community through uploads to the dataset access module of the Euro-VO. This would be especially important for instance in the areas of observations of bright night time objects such as novae, comets and supernovae, where members of the public often play a very important role.

Thus, aspects of the design and technical preparation work proposed here would focus on including a wide range of stakeholders into the system, improving the public understanding of this area of science, and increasing their ability to participate.