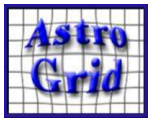
AstroGrid Oversight Committee
Paper AOC(05)2
Oct 20th 2005



AstroGrid-2 Lifetime Plan and Roadmap

Version-1 : May 2005
Version 2 for AGOC : Oct 12th 2005
Author : Andy Lawrence, AstroGrid Project Leader

(A) Introduction.

This is the top-level **Lifetime Plan** for the AstroGrid-2 project. It supersedes the plan in the original proposal (April 2003) and the draft Project Plan (June 2004) on the AstroGrid-2 Twiki. Guided by the long term roadmap, a firmer but still top-level **Cycle Plan** is made at the beginning of each six month cycle. Within each cycle, the Project Manager produces a much more detailed **Project Plan.** All of these documents can be found online on the AstroGrid-2 Planning Pages of the AstroGrid wiki. The online versions should be considered the currently definitive versions.

The Lifetime Plan consists of several parts :

Part A : Introduction
Part B : Vision
Part C : Goals
Part D : Methodology
Part E : Work Areas
Part F : RoadMap
Part G : Governance, Management, and Oversight
Part H : International obligations
Part I : Staff Plan
Part J : Budget Plan Summary

(B) Vision

The Virtual Observatory (VO) 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 (VO) 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. The world-wide super archive becomes the sky, and software the instrument with which we collect data from the sky - hence the metaphor of the Virtual Observatory. The VO concept can be seen as an 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 visualise results, to stack and mosaic images, to query catalogues and create subsets, to integrate data from different origins, or to calculate a correlation function. Such calculations will be data services offered by the expert data centres holding the data. They will be standardised to be compatible across many archives. The result is a service grid. The VO 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 VO Framework. To achieve the whole vision, however, data centres, tools writers, and facility builders will all work within this framework.

The AstroGrid vision

Our vision is twofold : to provide a generic infrastructure for the world; and to build a specific deployment for UK astronomy.

AstroGrid is committed to the open framework approach for the VO. The first part of the AstroGrid vision is to build the *infrastructure* that will make this possible. As well as open standards, this means specific software components which implement those standards. We aim to make our infrastructure available worldwide, by constructing software components that are mix-and-match, and as far as possible plug-and-play, so that they can easily be used with other products from around the world. We therefore aim to build to a high standard of software engineering and documentation.

The second part of the AstroGrid vision is to deliver a working system of daily use to UK astronomers. To achieve this, we will deploy our own infrastructure components at data centres in the UK, work with our colleagues at those data centres, and will integrate tools and applications emerging from the UK and elsewhere.

AstroGrid-2

The original AstroGrid project, as well as sharing the work of developing the key international standards, has succeeded in developing most of the core infrastructure components that make the AstroGrid vision possible - Registry, Virtual Storage, Workflow, Common Execution Architecture, Community, Portal, and Query Builder. However, the deployment was only experimental; tools and applications provided were limited; and developing technologies mean that the infrastructure itself could become much more powerful and flexible. The AstroGrid-2 vision is therefore to **complete and extend** the original AstroGrid vision. The key components are to complete a full deployment, to deploy a wide range of tools and applications, to develop a framework for ambitious datamining work, and to research and deploy new infrastructural technologies, such as ontology and agents.

(C) Goals

Our overall goals remain those of the original AstroGrid project, but AstroGrid-2 also has specific additional goals within this setting.

Overall VO goals

The VO is intended to be a system that will

- allow users to interrogate multiple data centres in a seamless and transparent way
- provide powerful new analysis and visualisation tools
- give data centres a standard framework for publishing and delivering services using their data

AstroGrid scientific aims

- improve the quality, efficiency, ease, speed, and cost-effectiveness of on-line astronomical research
- make comparison and integration of data from diverse sources seamless and transparent
- remove data analysis barriers to interdisciplinary research
- make science involving manipulation of large datasets as easy and as powerful as possible

AstroGrid practical goals

- develop, with our IVOA partners, internationally agreed standards for data, metadata, data exchange and provenance
- develop a software infrastructure for data resources
- deploy a physical grid of resources across key data centres
- develop a searchable, self-replicating and easily manageable resource registry
- implement a working Virtual Observatory system based around key UK databases and of real scientific use to astronomers
- provide a user interface to that VO system
- provide, either by construction or by adaptation, a set of science user tools to work with that VO system
- establish a leading position for the UK in VO work

AstroGrid-2 goals

- expand, revise and improve the core infrastructure
- provide a suite of science user tools, and set up a system for responding to personalised user needs
- research and deploy new automated resource discovery techniques - AstroOntology and Intelligent Agents
- research and deploy techniques in Grid technology, visualisation and datamining, creating an AstroGrid Data Exploration Framework (ADEF)

The original AstroGrid-2 proposal had two additional goals. We were not funded to carry out these activities, but if additional funds were found we would aim to :

- create a UK Data Centre Alliance, and provide co-ordinated effort to take up VO and Grid technology
- develop dedicated outreach software to allow schools and the public to access the Virtual Observatory

(D) Methodology

Planning and Development Philosophy. AstroGrid uses an agile but goal driven development philosophy. Iterative flexibility in both goal setting and development is becoming standard industry practice. It is particularly important in the VO world because the external environment (W3, industry, and IVOA) is evolving fast and somewhat unpredictably, and because detailed goals and requirements only sharpen as customers see the possibilities - a kind of "requirements uncertainty principle". During the first AstroGrid project (2003-2005) we used three month iterations. This had many strengths, but long term oversight (both by PPARC and by the lead investigators) was difficult, and it was very hard to keep to a formal three month delivery cycle. On the other hand, detailed development was found to work better when much faster and informal. For AstroGrid-2 we therefore refined the method. We publish a *Lifetime Plan* (this document); we use six month *Planning Cycles* at which we set public goals and deliverables; and within a Cycle the Project Manager develops a *Detailed Project Plan* which can evolve rapidly and flexibly, often setting a series of two week "story" goals.

Collaborative Open Project working. AstroGrid maintains as much collaborative working as possible, through the use of a *Twiki*, through *News* and *Forum* web sites, through the use of Jabber, and through physical meetings. The Twiki, News, and Forum also turned out to very useful tools for collaboration with other projects. We are not an *open source* project in the sense of allowing any outsiders to contribute code; but we are an *open project* in that everything we do is visible from the outside, all source code is available, and we exchange code and ideas freely with related projects.

Engineering Standards. AstroGrid intends its software components to be used by external projects. They therefore need to be constructed and documented to product standards. The software is not a monolithic system, but a series of components. We intend these components to operate with components from other VO projects, and other e-Science projects. This requires strict compliance with IVOA and W3C standards as well as engineering standards of robustness and reliability. Our components are intended to be plug-and-play and also mix-and-match. These aspirations are gradually becoming reality.

(E) Work Areas

The project recognises several broad areas for planning work. These are not "workpackages" in the traditional sense, as they do not have fixed deliverables or resource allocation - deliverables are fixed for each Cycle Plan, and workgroups assigned at that time.

The original proposal had seven work areas : W1 = management and co-ordination ; W2 = new software infrastructure; W3 = research and prototyping; W4 = standards development; W5 = public outreach; W6 = uptake and physical grid; and W7 = data centre support. These have been revised, partly because some activities were not funded, and partly to optimise the match with the areas within the FP6 funded VOTECH design study. The new work areas are as follows. The first few match explicitly onto a corresponding VOTECH DS area.

A1 = DS1 : Programme Management.
This provides top level policy development amongst the AGLI, overall project planning by the management team, co-ordination with external e-science colleagues and the IVOA, financial planning and decision making, and establishing an external presence through web pages, workshops, and so on.

A2 = DS2 : Scientific and Technical Management
This provides leadership and technical integration for the project and is responsible for all external technical deliverables. It is responsible for common processes and standards, the software repository, versioned software releases etc, cand collaboration mechanisms such as Forum, Wiki, Maven, and CVS. Detailed plans are evolved each cycle, and the work of developers and scientists planned and monitored by the Project Scientist and Project Manager.

A3 = DS3 : Infrastructure development.
This is the core of AstroGrid work. It is responsible for delivering fully engineered working software components that can be used in deployed systems. This includes the AstroGrid user interface such as the portal and the workbench. It also includes documenting and releasing the software suite so that other projects worldwide can use it. Much of the basic infrastructure is in place, but new components are envisaged, and much will need replacing and refreshing. Requirements for new and refreshed components will emerge from the R&D work areas below.

A4 = DS4 : New User Tools.
We expect most tools development to happen outside AstroGrid. However, some PPARC funded AstroGrid-2 effort will be matched with the VOTECH funded resource to develop new user tools that work with the infrastructure. (Note that developing general methods for easily integrating third party applications is seen as part of A3, infrastructure development).

A5 = DS5 : Intelligent Resource Discovery.
This work area aims at studying and assessing new technologies that will increase the power of resource discovery - e.g. ontology, the semantic web, intelligent agents. The work includes design studies and prototype implementations before proposing a new component for A3.

A6 = DS6 Data Exploration.
This work area aims at studying 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 AstroGrid-compliant, and how they can be extended to extremely large datasets. The intention is to deploy specific examples of such packages, but also to develop a generic "AstroGrid Data Exploration Framework (ADEF)".

A7 : Standards programme.

Development of standards is at the core of the whole VO enterprise. This work area sets aside effort for contribution to IVOA workgroups. AstroGrid intends to be pro-active in leading several areas of international standards work.

A8 : Deployment
The aim is to deploy AstroGrid software at UK data centres, publishing key datasets within the framework, and releasing successive versions of a working system for actually doing science. This cannot be done by the project alone, requiring us to work with colleagues at UK data centres. The combination of core AstroGrid staff and the key individuals at the data centres is referred to as the "Greater AstroGrid"

(F) Roadmap

Here is a summary of the key milestones on the Roadmap, followed by brief explanations.

2005 : Cycles 1-2

2005 Feb	VOTECH Project Plan
2005 Apr	working system demo at NAM
2005 Apr	Initial AstroGrid workbench
2005 May	V1.0 release of portal and software suite
2005 Jul	V1.1 s/w release ; full documentation
2005 Sep	VOTECH Science Framework Document
2005 Oct	Multiple third party applications in workbench
2005 Oct	Registry Exploration Browser
2005 Nov	Release AstroGrid Common Runtime
2005 Dec	Fully specified Registry with Xquery interface
2005 Dec	Universal Worker Service Working Draft V0.9
2005 Dec	International storage sharing : VOStore
2005 Dec	Initial server side applications

2006 : Cycles 3-4

2006 Feb	Personalised tools Call-1
2006 Apr	Query Builder and Workflow Builder for Workbench
2006 Apr	Deployment release : demo at NAM
2006 Apr	Query Builder and Workflow Builder for Workbench
2006 Apr	VOTECH Infrastructure Study Report
2006 May	Complex search interface
2006 Jun	Major package release
2006 Jun	Demonstration of large data volume problem
2006 Jun	Single Sign-on for VO services
2006 Jul	V1.5 s/w release
2006 Jul	VOTECH Revised Project Plan
2006 Aug	Prototype ontology use
2006 Sep	VOTECH Data Exploration Study Report
2006 Oct	Universal Worker Service Recommendation V1.0
2006 Oct	Client side scripting for workflow
2006 Dec	ADEF Prototype release
2006 Dec	Federated access : VOSpace
2006 Dec	Secure authorised access to VO services
2006 Dec	New Look Portal

2007 : Cycles 5-6

2007 Feb	Personalised tools Call-2
2007 Apr	Deployment release : demo at NAM
2007 Apr	VOTECH Tools Study Report
2007 Jun	Full access control
2007 Jun	Major package release
2007 Jun	First Resource Discovery Prototype
2007 Jul	V1.8 s/w release
2007 Jul	Full client side integration
2007 Sep	VOTECH Resource Discovery Study Report

2007 Oct Integrated ontology services
2007 Dec V2.0 final s/w and deployment release
2007 Dec Second Resource Discovery Prototype
2007 Dec Working ADEF

2008 : post AG

2008 Dec Euro-VO Reference Architecture

Release Plan

AstroGrid will be making releases both of its *software suite* and of its deployed working system. At first these are simultaneous but will gradually become distinct, especially as multiple deployments emerge, and as AstroGrid deployments will increasingly use modules from other projects. The deployments cannot be produced by the AstroGrid project alone, requiring a commitment of effort from UK data centres. However, AstroGrid leads this "Greater AstroGrid" effort. The best arrangements for *long term provision* of VO deployment in the UK is a separate matter that we won't address here.

The critical milestone is the V1 release of spring/summer 2005. This is the first point at which a working system for end-users is released, together with a simultaneous release of the software suite, including documentation, configuration scripts, and so on. Following this, we will release new versions of components separately as they emerge, so that releases will be in almost continuous drips . Because of the mix and match nature of AstroGrid software, and its intended use by external parties, new components will need to be compatible with multiple versions of other components. However this cannot be indefinitely sustained, and so we will also need to occasionally announce the *deprecation* of older versions of components.

There will still be a planning and PR value for annual packaged up announcements which can constitute public releases. Likewise, although science functionality in deployments will continuously improve, we will aim at annual announcements tied to the UK National Astronomy Meetings.

Release Plan Milestones

2005 Apr working system demo at NAM
2005 May V1.0 release of portal and software suite
2005 Jul V1.1 s/w release ; full documentation
2006 Apr deployment release : demo at NAM
2006 Jul V1.5 s/w release
2007 Apr deployment release : demo at NAM
2007 Jul V1.8 s/w release
2007 Dec V2.0 final s/w and deployment release

External Science context

AstroGrid is intended to be completely generic but there are some key facilities we specifically target which are of importance to UK astronomers. The first two are WFCAM and VISTA, which will produce huge new surveys, and for which data processing and archiving have been paid for through the e-science line. The next is SWIFT, which produces unique multi-wavelength capabilities. In the longer term, the key facilities are seen as Planck, Alma, and Gaia.

Infrastructure Components

AstroGrid has some kind of working component in all the key areas - Registry, User Interface, Dataset Access, Workflow, Virtual Storage, Community. We do not yet have the full original intended functionality, and expect this to emerge gradually.

Identity, Authentication, and Authorisation. This is the area where we are furthest behind our original intentions, partly because of the difficulty of international standardisation, and acceptance of new methods by data centres. We expect to achieve internationally consistent community based *single sign-on* by mid-2006, allowing authorised read access; and *secure access* by the end of 2006, allowing full control and write access to relevant services.

2006 Jun Single Sign-on for VO services
2006 Dec Secure authorised access to VO services

Virtual Storage. This is already an AstroGrid hit with Myspace - essentially, remote storage services are provided by the Community that a user logs on to. By the end of 2005 we aim to have international sharing of storage via the VOStore protocol. By the end of 2006 we aim to have federated access to data (merging the MySpaces) together with intelligent data routing (loosely referred to as VOSpace). The user should also then be able to interface to Storage Resource Broker as well as MySpace. Finally we should allow full access control by 2007.

2005 Dec International storage sharing : VOStore
2006 Dec Federated access : VOSpace
2007 Jun Full access control

Applications Integration The key to our success so far has been the definition of a Common Execution Architecture (CEA), and Common Execution Connectors, which define a Web Service interface to applications of all kinds, and allow asynchronous activity. We aim to engineer a new WS-RF based version for standardisation, called the Universal Worker Service (UWS). Another central concept is that of the AstroGrid Client Runtime (ACR), which should complete in 2005 and subsequently allow integration of packages such as IRAF and MIDAS.

2005 Nov Release AstroGrid Client Runtime
2005 Dec Universal Worker Service Working Draft V0.9
2006 Oct Universal Worker Service Recommendation V1.0

User Interface. AstroGrid-1 was entirely browser based, with web pages for Registry, MySpace, Query Builder, Workflow Builder, and Job monitoring. Complex functionality was rendered through HTML at the server via Cocoon. This proved to have a heavy development cost. A top priority during Cycle-1 was therefore to develop an additional "Workbench" interface launched through Java Webstart. We will add functionality to this during 2005, whilst maintaining and improving the Portal without significant new functionality. During 2006 we will produce versions of Query Builder and Workflow Builder for the Workbench, whilst developing a new-look portal, which will be released at the end of 2006.

2005 Apr Initial AstroGrid workbench
2005 Oct Multiple third party applications in workbench
2006 Apr Query Builder and Workflow Builder for Workbench
2006 Dec New Look Portal

Workflow and client integration. Workflow depends on the CEA and on the Job Execution Service (JES) which will be gradually enhanced. Building workflows is currently through a browser based tool which is ingenious but has a heavy development cost. We will therefore target a Webstart

workflow builder for mid 2006. Apart from these interface issues, the key advance during 2006 and 2007 will be *client side scripting* for workflow, so that end users can write Python, Perl scripts, or whatever they wish, and put it into their own workflow. Similarly, a user should be able to write a Python script inside which they can call VO services in a standardised way. They should be able to upload algorithms to the server, to run next to the data, and see their own local directories in MySpace. In general your laptop should be part of the VO.

2006 Apr	Query Builder and Workflow Builder for Workbench
2006 Oct	Client side scripting for workflow
2007 Jul	Full client side integration

Registry. The registry is a rich and potentially powerful structure, and we have system for querying it, but there is a feeling that we have yet solved the ergonomic problems of how to unlock its power. The first key goal is therefore to develop an improved query interface and *registry browser*. Meanwhile we are leading international standardisation in the Registry area, which we expect to complete in 2005, with a key change being use of Xquery rather than SQL. By mid-2006 we expect to expose the full power of the Registry to the user, with complex searches. More exciting and speculative improvements in resource discovery will come through the relevant VOTECH research area, with the use of ontologies.

2005 Oct	Registry Exploration Browser
2005 Dec	Fully specified Registry with Xquery interface
2006 May	Complex search interface
2006 Aug	Prototype ontology use
2007 Oct	Integrated ontology services

IVOA standards roadmap

International standardisation is the key to the VO. Agreement between projects is aggressively pursued through the International Virtual Observatory Alliance (IVOA), with mature standards then passed on to Commission 5 of the IAU for formal global endorsement. The IVOA has a clear Technical Roadmap, which AstroGrid staff have been very active in helping to define. This gives both a framework within which to evolve, and a target for new infrastructure - in places where we are moving ahead of international competition, we must plan standards that we would wish the IVOA to adopt.

IVOA milestones :

2005 May	REG	VO-identifiers1.0(R)
2005 May	UCD	UCD1+-V1.0 (R)
2005 Jun	UCD	create vocab and tech boards
2005 Jul	DM	STC-V1.0 (R)
2005 Jul	REG	RM-V1.1 (R)
2005 Jul	REG	VOResource-V1.0 (WD)
2005 Jul	DAL	SSA-V0.9 (WD)
2005 Jul?	VOQ	ADQL-V1.0 (R)
2005 Jul?	VOQ	SkyNode-V1.0 (R)
2005 Jul	UCD	UCD-vocab (R)
2005 Jul	DM	SpecLine-V0.1 (WD)
2005 Aug	VOT	VOT-V1.2 (R)
2005 Aug	GWS	VO-SupportInterface-V1.0 (R)
2005 Sep	VOQ	ADQL-V1.x (WD) (DAL integn)
2005 Oct	UCD	change name/nature of WG

2005 Oct	DAL	SSA-V1.0 (WD)
2005 Oct	VOT	review fuure of WG
2005 Oct	TIG	Theory Use Case Doc
2005 Nov	REG	RegOfReg-V1.0 (R or Note)
2005 Nov	GWS	VO-WS-V1.0 (R)
2005 Dec	VOE	VOEvent-V1.0 (R?)
2005 Dec	GWS	VOStore-V1.0 (R)
2005 Dec	DAL	SIA-V1.1 (WD)
2005 Dec	DM	SED-V1.0 (R)
2005 Dec	DM	Catalog-V0.5 (WD)
2006 Jan	DM	Characterisation-V1.0 (R)
2006 Jan	GWS	SSO-V1.0 (R)
2006 Jan	GWS	UWS-V1.0 (WD)
2006 Jan	GWS	VOSpace 1.0 (R)
2006 Feb	REG	VOResource-V1.0 (R)
2006 Feb	REG	RegInterface-V1.0 (R)
2006 Sep	VOQ	ADQL-1.x (R)
2006	DM	Obs-Provenance (R?)
2006	DM	Quantity-V1.0 ?
2006		new semantics WG ?
2006		Ontology 1.0 ?
2006		Preservation 1.0 ?

Tools and Science Services

The AstroGrid infrastructure is intended to be quite general, and is very powerful, with the ability to compose complex workflows. However this very general power can be confusing and intimidating to beginning users. We therefore intend to pre-package certain workflows corresponding to popular scientific tasks with a more standard “form filling” interface. We refer to the interface to these pre-packaged *parameterised workflows* as *science services*. We also intend to make it simple for third parties to write *applications tools* which use the infrastructure i.e. they can call VO data services, can be included in workflows, can write to MySpace and so on.

VOTECH research and feedthrough

With our European partners in VOTECH, AstroGrid2 will carry out a research programme aimed at development of future components. This is not blue-skies research, or academic computer science research. Rather, it takes new technologies being developed elsewhere, assesses them, builds prototypes, and then designs components to take forward into the mature infrastructure. As well as looking out for new components, this work also aims at identifying as early as possible the need for new standards and protocols, which we would then take forward to the IVOA agenda. The VOTECH work programme is divided into six separate "Design Study" areas.

DS1 Consortium management. Work is co-ordinated on a six monthly cycle, but 180 degrees our of phase with the rest of AstroGrid-2. Board meetings take place annually, and meetings of the Technical Advisory Panel (TAP) every six months

2005 Apr	TAP meeting
2005 Sep	TAP meeting
2005 Nov	Board meeting
2006 Apr	TAP meeting
2006 Sep	TAP meeting

2006 Nov Board meeting
2007 Apr TAP meeting
2007 Sep TAP meeting
2007 Nov Board meeting

DS2 Technical Project Management. This includes the responsibility to produce an overall project plan, a science analysis document, a sequence of software releases, and a final Euro-VO architecture.

2005 Feb Project Plan
2005 Sep Science Framework Document
2006 Jan Baseline software release-1
2006 Jul Revised Project Plan
2007 Jan Baseline software release 2
2008 Jan Baseline software release 3
2008 Dec Euro-VO Reference Architecture
2008 Dec Baseline software release 4

DS3 New Infrastructure AstroGrid will be leading the new infrastructure design study. The substantive improvements anticipated are therefore exactly as discussed above, with the VOTECH component of AstroGrid-2 peforming technology assessment and prototyping, leading on to mature component construction within the PPARC component of AstroGrid-2. The VOTECH project imposes one additional commitment - a study report in early 2006.

2006 Apr VOTECH Infrastructure Study Report

DS4 New User Tools This is understood to mean client side tools, rather than server-side applications which are pursued within DS6. The framework provided by CEA and ACR within AstroGrid makes it relatively easy to adapt tools for the VO, and so we expect a steady stream of these. However, we aim to adapt major packages, such as IRAF, in distinct releases; and to run a community call for new personalised tools, in combination with workshop events.

2006 Feb Personalised tools Call-1
2006 Jun Major package release
2007 Feb Personalised tools Call-2
2007 Apr VOTECH Tools Study Report
2007 Jun Major package release

DS5 Intelligent Resource Discovery. This is the most open-ended but potentially the most exciting part of VOTECH developments. What we have already is standards for resource and service metatadata, registry interfaces for accessing these, and a dictionary of standard terms, the UCDs, which make possible standardised queries. The next steps come in a more structured approach to UCDs building towards ontologies, defining relations between terms, and operations applicable to terms, with the aim of more powerful and flexible queries becoming possible. Once again, we will not be doing any original research here, but rather testing emerging technologies and prototyping. However external developments in ontologies and the semantic web are coming very fast, making this a significant challenge, and leaving the final outcome the most uncertain part of AstroGrid-2.

2007 Jun First Prototype releases
2007 Sep VOTECH Resource Discovery Study Report
2007 Dec Second Prototype releases

DS6 Data Exploration. Provision of server side applications is a key part of delivery of the VO vision, as every user will have access to fast search engines next to the very large databases we expect, and supercomputer analysis engines to produce correlation functions, cluster analyses, and so on. As with client side tools, lots of relevant packages for visualisation and data mining already exist, with the key task being integration into the VO framework. For very heavy duty applications however it is not yet clear the necessary infrastructure is complete, especially in grid-like aspects of bulk data movement and replication, and optimally allocating and re-routing long-lived jobs. We are working towards a concept currently referred to as the "AstroGrid Data Exploration Framework (ADEF)".

2005 Dec	Initial server side applications
2006 Jun	Demonstration of large data volume problem
2006 Sep	VOTECH Data Exploration Study Report
2006 Dec	ADEF Prototype release
2007 Dec	Working ADEF

(G) Governance, Management and Oversight

The AstroGrid Lead Investigators (AGLI)

The AGLI is the ultimate authority - the "Board". It defines top level policy, takes decisions on resource allocation, and is ultimately responsible for the strategic direction of the project. It also provides oversight and guidance. The AGLI meet monthly, usually by telecon. A list of AGLI members and records of meetings are kept online on the AstroGrid wiki. The AGLI are :

Cambridge	Richard McMahon	Bristol	Malcolm Bremer
CLRC	Peter Allan	CLRC	Ian McCrea (solar lead)
Edinburgh	Andy Lawrence	Exeter	Tim Naylor
Leeds	Ken Brodlie	Leicester	Mike Watson
Manchester	Simon Garrington	MSSL	Len Culhane
Portsmouth	Bob Nichol		

The Management Team

Strategic direction and planning is developed by the management team for approval by the AGLI. The management team is responsible for producing the six-monthly Cycle Plan and seeing that it is implemented; for producing material for PPARC; and for liasing with external partners and projects. They report monthly to the AGLI.

Project Leader	Andy Lawrence
Programme Manager	Tony Linde
Project Manager	Keith Noddle
Project Scientist	Nic Walton
AGLI reps	Mike Watson, Richard McMahon

The **Project Leader**, assisted by the AGLI reps, has overall responsibility for project policy and delivery, acting on behalf of the AGLI from day to day.

The Project Scientist develops science requirements for the project, and works to see that the project meets these requirements, leading the Science Team, and advising the management team, AGLI, and development team as appropriate. He is also the main liaison point with the astronomical community.

The **Programme Manager** has responsibility for the definition of AstroGrid product, the strategic technical development of AstroGrid, its financial control, and overall planning. He is also the VOTECH Project Manager, and is the main liaison point with external bodies and projects, such as PPARC, the UK e-Science programme, JisC, the IVOA, and the other components of Euro-VO.

The **Project Manager** is responsible for detailed delivery of AstroGrid software, planning and directing the work of the development team, and co-ordinating the work of technical tools such as CVS and Maven. He works directly with the development team and liaises with the science team as appropriate.

Advisory Bodies

Advice is taken approximately every six months from meetings of two bodies. The first is the AstroGrid Science Advisory Group. This acts both to comment on project progress, and to act as a pool of beta testers. The second is a body called the AstroGrid Technical Advisory Group

(AGTAG). It contains a mixture of developers from within the project, data centre representatives, and other expert parties such as Starlink staff.

External Oversight

AstroGrid-2 is overseen by a PPARC committee known as the AstroGrid Oversight Committee (AGOC). This contains a mixture of astronomers, computer scientists, and members of other e-science projects. It monitors progress against the Lifetime Plan and each Cycle Plan; oversees the expenditure of the project; receives reports from the AGSAG; and requests any paperwork it sees as necessary to assess the state of the project.

The project maintains a set of OversightPages to make it easy to track progress externally.

(H) International Obligations

The only **formal** obligation concerns the VOTECH project. This is discussed in detail below.

Considerable informal obligation arises with the The International Virtual Observatory Alliance. The IVOA is a voluntary organisation and so entails no formal obligations. In practice though, AstroGrid is committed to being a key player, playing active roles in both top-level policy and strategy, and detailed standards development within workgroups.

Futher formal obligation may emerge within the Euro-VO structure - a Memorandum of Understanding (MOU) is in preparation between European partners that will commit us at least morally to playing a role in the emerging Euro-VO as a working structure. PPARC will be the formal signatory for the UK, although in practice the participation is through the AstroGrid consortium. Euro-VO has three parts - the VO Facility Centre (VOFC); the Data Centre Alliance (DCA); and the VO Technology Centre (VOTC). The UK is involved in all three, but the one we lead is the VOTC. The VOTC is seen as a loose umbrella organisation, that may spawn a variety of projects. So far the only project concerned is VOTECH.

The VOTECH project

The public home page for VOTECH can be found at http://eurovotech.org. The internal project pages are http://wiki.eurovotech.org/bin/view/VOTech/WebHome

(1) OBLIGATIONS

At a formal level, AstroGrid-2 and VOTECH are distinct projects. VOTECH, as an FP6 project, runs as a legal contract with the EC. We are committed to the deliverables described in the project plan which is an appendix to the contract. We also have a moral commitment to our European partners to make the whole project a success, especially because we are leading it. We therefore take VOTECH very seriously on its own terms, and not simply as an augmentation of AstroGrid-2.

We also have an obligation to PPARC. The obvious danger, because we are committed to supplying PPARC funded AstroGrid staff effort to VOTECH, is that the remaining effort is too small to meet those obligations. The solution is that the work plans for VOTECH and for AstroGrid-2 are aligned so that the maximum possible amount of work counts for both.

(2) AIMS of ASTROGRID and VOTECH.

We could crudely divide AstroGrid-2 into AG-P (Prototyping) and AG-E (Engineering). AG-P is exactly the part that we make count twice, for VOTECH and AG2. Then

VOTECH+AG2-P is about

- Discovery (what is available, what can we do)
- Evaluation (how does it work, will it fit)
- Prototyping (trial builds)
- Design

AG-E is about

- Stabilisation (make it work predictably)
- Productisation (package for installation)

- Deployment (make it available)

The VOTECH+AG2-P bit is sometimes referred to loosely as "research", but this does NOT mean blue skies research. It means evaluation, and design, as above, explicitly aimed at feeding into the engineering.

(3) VOTECH MANAGEMENT STRUCTURE

The management structure is described in detail elsewhere. Briefly, the utimate authority is a VOTECH Board which contains the co-Is from each partner. This is equivalent to the AGLI, and takes policy and resource decisions. Day to day management is undertaken by the Project Manager (Linde), the Technical Lead (Noddle) and the Project Scientist (Walton), together with the Project Co-ordinator (Lawrence) - i.e. the same as the core of the AstroGrid Management Team. Programme planning is however the job of the Technical Advisory Panel (TAP), which has a variety of technical representatives from the partners, and spread across the various DS areas.

(4) ADVISORY STRUCTURE

Programme planning comes from the TAP. Formally, this is "advisory" with ultimate authority remaining with the Board. It also acts as the Technical Advisory structure, equivalent to the AGTAG. This means that in the future we can add TAP members from outside the VOTECH consortium, i.e. from ESA, or from data centres.

Science advice comes through a unified Euro-VO science advisory structure. Formally this reports to the Euro-VO Exec, but it also produces whatever advice requirements etc that VOTTC/DCA/VOFC need. VOTC-PS, VOFC-PS and (hopefully) DCA-PS run this as a collaborating triumvirate. It has an independent chair chosen from the science membership.

(5) WORK STRUCTURE

The VOTECH project is divided into six work areas, which are known as "Design Studies". DS1-6 exactly match AG work areas A1-6. There is a named leader for each DS who will remain fixed. There is also an initial list of which of the new EU-funded staff work in which area, but this is not considered fixed.

DS1 = Project Co-ordination (Andy Lawrence)
DS2 = Technical Management (Tony Linde)
DS3 = Infrastructure (Keith Noddle)
DS4 = User Tools (Markus Dolensky)
DS5 = Automated Resource Discovery (Sebastien Derriere)
DS6 = Data Exploration (Bob Mann)

(6) PLANNING CYCLE

The planning cycle is very similar to AstroGrid-2

- The Project Plan acts as a long term Roadmap, and lists key deliverables and milestones for the complete project lifetime.
- There are six month planning Stages. The TAP agrees the goals for each DS , and the responsibilities for each partner. We derive a simple agreed top-level document.
- The phasing of VOTECH Stages compared to AstroGrid Cycles is currently under discussion.

During each cycle, the top-level management from Tony Linde will be relatively light touch. The

DS leaders build more detailed plans, but much is left to local management.

(7) RESOURCE TRACKING

Legally, only the EU funded staff are visible. All the rest is to do with healthy running of the project, not to do with any formal obligations. Its useful to divide this in two parts.

(i) EU funded VOTECH staff. Each partner returns annual cost statements. We have considerable flexibility on expenditure, as long as it is "actual, economic, and efficient". (The last part means clearly spent to the benefit of the project). As a minimum then we certainly have to track exactly who is paid on the project as well as travel and equipment.

(Detailed Note : The project plan does state however how many staff months we expect to expend by partner and by DS. It is not clear whether we will be held to this. We believe this is not a legal obligation, but our contract officer will ask questions if we are doing something very different. We would be wise therefore to record how many months each funded staff member is deemed to have performed in each area. Note the careful use of "deemed". It is not clear whether this will be audited in any way, so it is probably up to us to declare suitably. Probably what we should do is ask each partner to send the co-ordinator a statement at the end of each cycle.)

(ii) Partner funded VOTECH staff effort. This is an informal agreement between the partners. The EC do not care about it at all. Partners are under no obligation to specify exactly what partner funded effort has been used, but in practice they are in fact quite open about this.

(I) Staffing Plan

CURRENT STAFF COMPLEMENT

Leadership

Andy Lawrence (Edin)	Project Leader and VOTECH Co-ordinator
Tony Linde (Leic)	Programme Manager and VOTECH Project Manager
Nic Walton (Cam)	Project Scientist for AstroGrid and VOTC
Keith Noddle (Leic)	Project Manager and VOTECH Technical Manager
Mike Watson (Leic)	VOTECH Co-I and management team AGLI rep
Richard McMahon (Cam)	VOTECH Co-I and management team AGLI rep

Co-ordination

Bob Mann (Edin)	VOTECH DS6 Leader
Keith Noddle (Leic)	VOTECH DS3 Leader
John Pye (0.17 FTE) (Leic)	AstroGrid Finance Assistant
Peredur Williams (0.2 FTE) (Edin)	VOTECH Administrator
TBA (0.5 FTE) (Leic)	AstroGrid Web and System Manager
TBA (0.5 FTE) (Edin)	VOTECH Web and System Manager

Science Team

Anita Richards (JBO)	Radio specialist
Silvia Dalla (0.5 FTE) (UMIST)	Solar and STP specialist
Eduardo Gonzalez (0.5 FTE) (Cam)	Extragalactic specialist
Jonathan Tedds (0.5 FTE) (Leic)	Galactic specialist

Development and Deployment Team

Elizabeth Auden (0.5 FTE) (MSSL)	Solar and STP infrastructure
Catherine Quin (Leic)	
Kona Andrews (Cam)	
Noel Winstanley (JBO)	
Dave Morris (Cam)	
Phil Nicholson (Leic)	
Jeff Lusted (Leic)	
Guy Rixon (Cam)	
Martin Hill (Edin)	
Kevin Benson (MSSL)	
Paul Harrison	seconded to ESO

VOTECH developer/researchers

John Taylor (Edin)	
Garry Smith (Ports)	
Richard Holbrey (Leeds)	
Mark Taylor (Cam)	
Alasdair Allan (Exe)	starting January 2006)
TBA (Leic)	

STAFF ROLES

Below are short explanations of the above staff types. These apply to the roles required by the project, and are not "job descriptions" for the current individuals carrying out those roles, which are the responsibility of the various organisations employing the staff. These roles were first set out in the staff plan of July 2004, as part of negotiations between the various consortium organisations and

PPARC. The main changes are in the clarification of Programme Manager / Project Manager roles, and in the addition of several VOTECH developer/researcher posts.

Over the lifetime of AstroGrid, the requirements of the project will evolve, and so the required staff roles, and even required skills, may likewise change. (In particular we have in mind the changing balance between software product development, deployment and operations, and research into new technologies). Staff are made aware of the changing needs of the project, and so of the requirement to be flexible. We will use existing staff wherever possible and sensible. If staff changes are needed to deliver our goals, the project will work sensitively with the employing organisations.

Leadership and Co-ordination Team.

The **Project Leader** has overall responsibility for project policy and delivery, acting on behalf of the AGLI from day to day. The **VOTECH Co-ordinator** leads the whole VOTECH project, but also the UK component. This work is assisted by the two other **VOTECH co-Is** in the UK, who also act as further representatives of the AGLI in day to day management of the combined AstroGrid and VOTECH programme.

The Project Scientist develops science requirements for the project, and works to see that the project meets these requirements, leading the Science Team, and advising the management team, AGLI, and development team as appropriate. He is also the main liaison point with the astronomical community.

The **Programme Manager** has responsibility for the definition of AstroGrid product, the strategic technical development of AstroGrid, its financial control, and overall planning. He is also the VOTECH Project Manager, and is the main liaison point with external bodies and projects, such as PPARC, the UK e-Science programme, JisC, the IVOA, and the other components of Euro-VO.

The **Project Manager** is responsible for detailed delivery of AstroGrid software, planning and directing the work of the development team, and co-ordinating the work of technical tools such as CVS and Maven. He works directly with the development team and liaises with the science team as appropriate.

The **AstroGrid Finance Officer (0.17 FTE)** provides a small amount of specialist assistance in monitoring and tracking AstroGrid2 finances, which is spread across many separate grants at many institutions, and including cross-institution invoices. This requires someone with general project finance control experience, and specific experience of PPARC grant mechanisms and regulations. The **VOTECH Administrator** likewise monitors and tracks expenditure of EU funds, but also provides more general administrative assistance, co-ordinating meetings, liasing with Brussels, and preparing annual cost statements and reports.

The **VOTECH Design Study (DS) Leaders** plan and lead the work of scientists and developers across Europe in the various VOTECH work areas, and report this work to the VOTECH consortium. This is not a formal management task, as the VOTECH partners are essentially independent, but rather uses a “light touch” approach of encouraging collaboration and co-ordination.

The Science Team.

In addition to the Project Scientist, we need scientists to (a) further develop the requirements in specific areas, (b) to test system components and provide feedback, at much more technical and detailed level than the AGSAG, (c) to participate in the international programme of standards development, and (d) to interact with data centre staff in constructing working deployments of

AstroGrid software. This requires individuals who are research active and experienced astronomers, but who also have significant technical skills in key areas such as XML and standard web technologies. These skills are required not just to be able to interact fruitfully with the development team, but also to be able to participate in IVOA and GGF standards development at full technical level, e.g. development of XML schemas.

The balance between the above areas may evolve. Furthermore, in addition to the above demanding range of skills, the project ideally needs to cover a range of astronomical areas - optical, infra-red, radio, and X-ray astronomy; galactic, extragalactic and planetary astronomy; observation and theory; solar and space plasma astronomy. Obviously no-one individual covers all the experience and skills areas, so the project tries to cover as much as possible with a balanced team.

Development and Deployment Team.

Core Team. We require a core team of **software engineers** who will build, test, and integrate code, following the programme set out by the Programme Manager, and under close supervision by the Project Manager. The overall team is expected to cover several types of work.

(1) Coding of core infrastructure components.

(2) Wrapping and integration of existing tools and applications.

(3) Implementation of new components that come out of the research programme.

(4) Contributions to the research programme. (Note this does not mean either astronomical research, or computer science research - rather, it means assessment of emerging technologies and development of prototypes.)

(5) Contributions to the IVOA standards programme.

(6) Deploying, operating, and maintaining AstroGrid components on real machines, and assisting data centre staff and other VO projects worldwide in deploying AstroGrid components.

As with the science team, it will not be feasible that everybody will be an expert in all five areas. The project needs to maintain a balanced team to cover all requirements. However, we do expect that most staff in the team will cover multiple areas, and they are required to be flexible about which type of work they are deployed on. The team members will be building production quality code in an agile and iterative environment and so need to be disciplined professional developers with skills and experience in key technologies such as Java, XML, and web services, with a proven track record in developing working code within a team setting. There also needs to be enough senior staff capable of acting as workgroup leaders when required. However, the project does not define a separate permanent group of workgroup leaders.

Solar physics infrastructure. Most of the infrastructure we will build is generic, but at least one developer from the core team will pay special attention to infrastructure requirements for the solar physics and STP areas. This is an AGLI agreement following the merger of the original AstroGrid-2 and SSVO projects.

Project Infrastructure. We have two **Web Developer/System Manager** posts, each of which is a half post meant to be shared with local system manager requirements. One post is for general AstroGrid support, and the other for VOTECH support. There are three areas of work. (1) Development and maintenance of collaborative technology for the project - News, Forum, and Wiki web sites, and the CVS and Maven systems used for software development. (2) Development

and maintenance of public information web pages. (3) Maintenance of machines hosting AstroGrid components and AstroGrid services.

VOTECH developer/researcher team

VOTECH aims at investigating and assessing relevant technologies, and prototyping new software components using those technologies. The core AstroGrid developers are sometimes expected to carry out work of this kind, but for staff hired specifically for VOTECH this is their main task. Staff in this team will usually be professional developers, but who have the right experience, skills or aptitude for technology research, and able to understand the science requirements. They will produce code rapidly in an experimental environment. Some staff may be astronomical researchers, but who are also capable of technology research and of production of prototype code.

(J) Budget Plan Summary

This is a summary of the operating budget as at May 2005.
More precise accounts are reported to the AGOC.

Total Personnel	3,580,709	
Travel	336,100	
Capital Equipment	10,000	
Miscellaneous	20,000	
Total non-personnel	398,350	
Contingency	20,941	(5% of non-personnel line)
TOTAL	4,000,000	