

The AstroGrid Project and the STFC Programmatic Review

Submission to consultation by AstroGrid Board

2008-03-12

Andy Lawrence (Project Leader); Keith Noddle (Project Manager); Nic Walton (Project Scientist)
Mike Watson, Richard McMahon, Peter Allan, Ian McCrea, Mark Birkinshaw, Len Culhane, Simon
Garrington (Co-Investigators)

Note : documents referred to, and other guidance concerning AstroGrid's place in the STFC Programmatic Review, can be found on a wiki page at : <http://wiki.astrogrid.org/bin/view/Astrogrid/STFCProgRev>

(1) BRIEF RESPONSE TO SPECIFIC PPAN FEEDBACK

AstroGrid has been placed in the "lower priority" group, and we have been told that concerns were (i) delayed delivery (ii) limited impact on users. These concerns are understandable but easily addressed.

(i) We are delivering a public release of our product, and an operational system, in April 2008, at the National Astronomy Meeting. This is precisely what we have agreed with our Oversight Committee, and fairly closely what we predicted in our original 2001 proposal. However, unlike building a new spectrograph, we have not had the luxury of holding back judgement until we were finished; users and committees expect and demand useful intermediate prototypes. Such working prototypes have in fact been planned and delivered on schedule. The feedback from the review of our Nov 2006 proposal noted "*..timely progress against the milestones of the previous project.*"

(ii) The prototype system of 2006-7 (the AstroGrid Workbench) has had several hundred registered users, mostly in the UK but also worldwide. This system has provided access to real datasets and has led to peer reviewed science publications. We track use of the software and present this to our Oversight Committee. For example during the month of December 2007 there were 589 user logins and 1,118 remote queries and workflows run. We have also run "science calls" where members of the community submitted competitive proposals to work with AstroGrid to build new tools and/or deploy new datasets.

For this prototype, the standard and correct perception has been that while everybody who tried it could see the potential, it was somehow not quite yet the "daily tool of choice". Four things have radically transformed that situation in the last year, two of which were in our direct control, and two of which were international issues, but where we have been responsible for pushing progress through at global, European, and UK levels. These things were (a) agreement on the necessary international standards; (b) publication of datasets worldwide in VO format; (c) the quality of our user interface; and (d) the robustness of our technical services. Demonstrating the pre-release new system over recent months, the typical reaction is "wow ! when is it coming out ?"

The STFC Oversight Committee, and our Science Advisory Group, participated on March 3rd in a "Critical Service Review" which confirmed our readiness for release, and reacted very favourably. We understand that the chair of the Oversight Committee will respond separately.

We will see by the end of the year whether AstroGrid is a success.

(2) STRATEGIC IMPORTANCE OF THE VIRTUAL OBSERVATORY (VO)

The VO is not a luxury. It is a necessity, and an inevitability, driven by data growth, user demand, science from multiple datasets, and technological bottlenecks, all of which push us towards a service economy, where data access, reduction, and scientific analysis are provided as online services by professional data centres. The aim is to provide a data management infrastructure for UK facilities, and to make it easier to

exploit the data and do the science. Compared to the billions spent building our facilities, the investment needed in this generic infrastructure is tiny, and so very cost effective.

These issues are recognised worldwide; in the USA the need for a transparent data infrastructure for Astronomy was highlighted by the last decadal survey as the highest priority small initiative, and the call for the creation of an operational service has just been released. UK facilities will have no choice but to be VO compliant, and our astronomers will have no choice but to use VO tools. One might argue that we should therefore leave it up to other countries to pursue. The result of this would be

- increased cost of compliance for our facilities
- loss of world leadership in a key area
- loss of influence in ESO and ESA
- loss of influence at the standards body (IVOA)
- loss of income to the UK from generic EC streams
- loss of a significant exploitation lead for UK astronomers
- loss of competitiveness for UK astronomers on international facilities
- slower future science exploitation

Note that loss of influence at IVOA is not a dry matter. For example, we have been able to align the IVOA agenda to produce standards needed by UKIDSS, VISTA, and GAIA, on a timescale compatible with their workplans.

STFC clearly do recognise the need for cross-cutting IT support and a data management infrastructure; this is why it is funding GridPP, and the RAL e-Science Centre, a large part of whose remit is support of Diamond, ISIS, and CLF. AstroGrid has had the hardest problem to solve, producing generic technology and infrastructure for a very wide range of astronomy and solar system facilities and archives. There is a balance to be found between centralised provision, simple market forces, and one-off technology development. However, given that the background technology and the international standards are constantly evolving, it cannot be the right solution to leave everything to chance. We suggest that STFC needs to develop a policy on data management across its whole range of activities.

(3) CURRENT STATUS

We were not asked to provide any input to the Programmatic Review. Specifically we were not asked to provide a questionnaire, because, as explained in the email sent to us on October 29th 2007 :

*"However, for some projects and services, STFC considered that it was not appropriate to request completion of the generic questionnaire. These are services, projects or facilities that are **still under construction** or where STFC requires a broader strategic review. For these projects, the Review will utilise other sources of information (for example, STFC Council papers detailing the level of agreed support together with any updates available from recent reviews within STFC)"*

We were indeed still under construction, with an announced target completion of April 2008. Although we have had no official information, we understand informally that the material used to review us was our proposal of November 2006 and associated paperwork, and oversight committee minutes up to the middle of 2007. This is very unfortunate timing, in that the situation for AstroGrid has improved radically over the last year, as discussed in section-1. AstroGrid is poised to be a great success this year; but this will not have been apparent from the paperwork used.

Our November 2006 proposal requested funding for five years to begin an operational facility. Following

review, PPARC Council wrote in April 2007 to approve funding of £2.7M for two years only, 2008-9. The complete feedback is available on [this temporary wiki page](#). Here are some key extracts :

"There was a growing demand from astronomers worldwide to develop a Virtual Observatory system, which would not only give on-line access to the growing number of data archives but also allow interoperability by having a transparent data infrastructure"

"The UK group were world leaders in this field and played a significant role in the International Virtual Observatory Alliance (IVOA)"

"The applicants had made timely progress against the milestones of the previously funded project and now had a working beta product, which they had started to roll out."

Looking towards implementation, the key text is :

"Funding was recommended for a 2-year rather than a 5-year programme. After 2 years the project should be reviewed as an operating facility (rather than a project) and a decision for continuation or otherwise made based on user uptake and scientific productivity"

This seemed eminently sensible, limiting PPARC (and now STFC's) exposure, but not taking a final decision until after release. We believe that in fact this is consistent with the spirit and sense of the STFC Programmatic Review, and hope that this means the decision should be to simply announce a firm timescale for review and decision.

AstroGrid is run through a co-ordinated portfolio of grants. Following the April 2007 approval, and subsequent detailed discussions with STFC officials, we had expected new grants to begin in January. None of these grants have been announced, and so our staff are being paid without us holding funding. We have had informal promises of minimum six months funding, but the grants could be released without loss of flexibility, as it is STFC's power to withdraw them anyway.

(4) ASTROGRID ACHIEVEMENTS

- A prototype system has been in use for two years, delivering peer reviewed science
- We are about to release a full featured real VO system, ready to be released to the community at the upcoming National Astronomy Meeting
- We are the founder members, along with the US, ESO, and CDS, of the International Virtual Observatory Alliance (IVOA) - the standards body which is the root of the future astronomical infrastructure - and have throughout been a leading force, setting standards to the advantage of UK facilities and data centres
- We are the only VO project worldwide providing a complete technical and service infrastructure
- We have produced a flexible and powerful user interface
- We have produced key technical solutions which make the VO possible - especially in Virtual Storage, Workflow, an API for tools writers, and application messaging system. AstroGrid is the clear world leader in these areas

- We have delivered working operational services.
- We have written components and toolkits for facilities and data centres to publish their data in the VO, saving effort across a wide range.
- We have published and deployed key UK datasets, as well as providing access to astronomical data worldwide
- We have established strong relationships with the key international players, and in many areas a clear leadership position; we are the acknowledged leaders of VO technology development in Europe.
- We have engaged with the UK University community via an extensive programme of multi-day workshops that introduced the VO to scientists, and provided vital feedback, changing the final design of our system.
- We are probably the only project that unites all of astronomy, solar physics, STP, and particle astrophysics
- We have interacted and collaborated with academic and commercial computer science, and other scientific disciplines, especially Bio-informatics, with whom we have exchanged concepts and software.

(5) COMMITMENTS TO PARTNERS

(5.1) EU project partners

We have EU resources which were obtained on a promise of matching STFC resource - currently, with VOTECH, DCA, and EuroVO-AIDA, approximately 7 FTES in 2008, reducing to 2.5 FTES in 2009. (A new proposal is planned which would bring us back to 5 Euro-funded FTES in 2009, but this cannot go ahead if we have no matching STFC funding.) The project contracts make explicit the partner resource in the workpackage plans, which we based on the resource promised to us by PPARC Council in April 2007. This matched funding is not a binding legal commitment, but is an unambiguous agreement with ESO, ESA, and CDS, and effectively sets a minimum for STFC funding in 2008-9.

As well as this formal commitment, we have been in a strong bidding position with respect to other partners because of the overall size of our project (15.1 FTES funded by STFC), and because of our demonstrable track record and lead in technology development, all giving us credibility to lead workpackages. Closing AstroGrid, or even drastically reducing it, would then lose potential future income to the UK.

(5.2) Facilities and Missions

All recent astronomical, solar, and STP facilities and missions have built an assumption of the existence of the VO into their plans. Most recent projects have explicitly recognised that AstroGrid, not just the VO in general, will make compliance easier and cheaper for them. The VISTA, Gaia, Solar Orbiter, and Planck proposals for example explicitly built workplans around this reliance. It is not a coincidence that the AstroGrid consortium is more or less the same as the major data centres in the UK - Cambridge, Edinburgh, Leicester, Manchester, MSSL, RAL.

(5.3) Global IVOA partners

The IVOA includes contributions from almost twenty projects around the world, but in practice progress in

the working groups relies on effort from US-NVO, AstroGrid, CDS, ESO, and ESA. As well as general contributions, we currently provide chairmanship or vice-chairmanship of four out of nine working groups - Data Access Layer, Applications, Data Models, and Grid and Web Services.

(5.4) Euro-VO

As well as the project effort commitments, we are part of a structure known as Euro-VO, established in 2005 through a Memorandum of Understanding agreed by several agencies including PPARC. (The MOU is available at the [temporary wiki page](#)). This does not commit STFC to any specific funding, but rather to an *"agreed co-ordinated programme of work, directed towards the establishment of a persistent Virtual Observatory (VO) research infrastructure for European astronomy."* It also set up key bodies, including a Science Advisory Committee, and the Euro-VO Executive Board. (The AstroGrid Project Leader, A.Lawrence, is a member of this four person Executive Board.)

Euro-VO is comprised of three parts, all of which are virtual distributed entities. (i) The VO Facility Centre (VOFC) is led by ESO and ESA. (ii) The Data Centre Alliance (DCA) is led by CDS-Strasbourg. (iii) The VO Technology Centre (VOTC) is led by AstroGrid.

The MOU was intended to apply through to the end of the VOTECH project, December 2008. Following this the intention of the partnership is to hold a review *"to assess progress and identify requirements for continued activities."*

(6) ECONOMIC and EDUCATIONAL RELEVANCE

Much of the underlying technology in AstroGrid is quite generic. There is potential for development with a wide range of partners.

Early in the development of AstroGrid, there was significant interaction with commercial partners, especially Oracle, IBM, and Microsoft Research. Later, this took a back seat, because the key task was to focus on delivery specifically for astronomy, and because alignment with other astronomers worldwide was much more important for success. However, AstroGrid has always been insistent on using and developing industry standard software, unlike particle physics grid work, which has always been based around specialised academic software from US supercomputer centres. Furthermore, the nature of the astronomical data infrastructure problem, like bio-informatics, but unlike particle physics, is almost identical to the problems faced by large business "enterprise computing". (An interesting example was seen at a Pharmaceutical Industry conference at which talks were given by A.Lawrence for AstroGrid and C.Goble for MyGrid ... a large company like Johnson and Johnson has thousands of scientists working in dozens of quasi-independent labs around the world; they need them to share their data and tools transparently. Johnson and Johnson is like the entire European astronomical community...)

Now that AstroGrid (and the Bio project MyGrid) have matured and delivered working product, it is almost certainly the right time to begin commercial collaborative work. AstroGrid and MyGrid are already collaborating fruitfully, so in fact it would make sense to do this together.

There are also significant outreach and educational possibilities. The release of "Google Sky" last year attracted much attention, and later this year the Microsoft equivalent (World Wide Telescope) will be released. Some even wondered whether such developments removed the need for the VO. This is not the case of course. Google sky is essentially a toy, and its developer, who was recruited by Google out of the US VO effort, refers to Google Sky as the "gateway drug" to the VO, which is the real thing. (A second VO worker, this time from AstroGrid, was recently recruited by Google). However, VO workers in the US, AstroGrid, and China rapidly developed links to real data. Planetarium style software is potentially one of several useful "front ends" to VO services, and an important link to public and educational use of data.

(7) THE ROUTE FORWARD

We recognise STFC's difficult constraints, but suggest that the kind of careful plans it is putting in place for many other projects and facilities - minimising commitment, while setting up a process for deciding the future, with no long term promise - is in fact already in place for AstroGrid. The April 2007 funding approval awarded money only for 2008-9, but allowed time to complete the product and assess its success. A review would be expected in late 2008. We suggest that the correct course of action is therefore simply to implement this plan.

Given the Euro-VO MOU, a review during 2008 should involve European partners, and especially ESO, ESA, and CDS. We would suggest that there should on the same timescale be a wide ranging review of STFC data management covering all facilities.

At the end of 2009, there would be several possible options. (i) Commit to a further period of UK operations, at a similar or reduced level. (ii) Begin proper European VO operations, in which AstroGrid plays a partial role. (iii) Stop completely, hoping for co-operation between facilities and data centres to somehow pick up the pieces, at the risk of overspend in those projects.