

Cover Sheet for Proposals <i>(All sections must be completed)</i>		JISC Capital Programme	
Name of Capital Programme: e-Research : e-Infrastructure			
Name of Lead Institution: University of Leicester			
Name of Proposed Project: Advanced Grid Authorisation through Semantic Technologies (AGAST)			
Name(s) of Project Partner(s): NeSC, University of Glasgow			
Full Contact Details for Primary Contact:			
Name: Dr Norman Gray			
Position: Researcher			
Email: norman@astro.gla.ac.uk			
Address: Department of Physics and Astronomy, University of Leicester, University Road, Leicester, LH1 7RH			
Tel: 0116 252 3494			
Fax: 0116 252 3311			
Length of Project: 12 months			
Project Start Date: 1 Feb 2008		Project End Date: 31 Jan 2009	
Total Funding Requested from JISC: £132637			
Funding Broken Down over Financial Years (Apr–Mar):			
Apr07 – Mar08		Apr08 – Mar09	
£26273		£106364	
Total Institutional Contributions: £16544			
Outline Project Description			
The AGAST project will investigate the extent to which semantic technologies will provide a flexible mechanism for easily-delegated access control. We will describe a number of challenging and realistic use-cases, drawn from the PIs' engagement with current projects. Building on this, we will produce a prototype, and confront it with these use-cases in a number of technology demonstrators.			
I have looked at the example FOI form at Appendix A and included an FOI form in the attached bid (Tick Box)	YES (see below)		
I have read the Circular and associated Terms and Conditions of Grant at Appendix B (Tick Box)	YES		

FOI Withheld Information Form

We do not require the JISC to withhold any information

Advanced Grid Authorisation through Semantic Technologies (AGAST) Case for support

Norman Gray (Leicester) and Richard Sinnott (NeSC, Glasgow)

2 October 2007

1 Executive summary

- 1 The technical and infrastructural achievements of the Grid security world are to a large extent vitiating by substantial usability problems. These problems affect both the users who wish to gain access to resources and the resource owners who wish to permit them. Whilst the Grid community has broadly adopted approaches based upon X.509 digital certificates to support Public Key Infrastructures for authenticating a user, expressing and enforcing more detailed access and control policies (authorisation) remains an area where numerous technologies and standards are available, without any having become clearly dominant.
- 2 It is our contention that semantic technologies have the potential to support a lighter touch way in which access control policies can be expressed, and importantly extend the way in which access control decisions can ultimately be made. For many application domains, and in a multitude of current scenarios, the information needed to make a local access control decision by a given resource provider needs to come from a variety of sources. Examples include fair sharing of resources between a number of cooperating sites, or quota management of resources distributed or shared across a virtual organisation (VOrg). Similarly, we want to identify security policy conflicts in the case where an individual holds roles in two different VOrgs.
- 3 The AGAST project will investigate the extent to which semantic technologies will indeed provide a flexible mechanism for easily-delegated access control. We will confront our existing prototype with a wide variety of challenging and realistic use-cases, drawn from the PIs' engagement with current projects.

2 Background

- 4 The heterogeneity of the Grid, in terms of both software systems and authorisation goals, means that it is probably futile to hope that the world wide Grid will converge on a small set of directly compatible authorisation systems; it is probably also unnecessary, since in many cases the authorisation policies are internal, and only the attributes they depend on are shared. This is fundamentally the same interoperability problem that the Semantic Web (SemWeb) community is addressing, and we contend that the focus on lightweight but robust shared semantics can improve upon existing (heavyweight) authorisation infrastructure solutions. This is not a claim that these technologies are redundant or misconceived – they are of course not! – but that they will be a poor 'impedance match' for certain important applications, and that in these cases an SemWeb-style solution will allow them to be used and composed at lower cost. That is, we are forced to develop federated

access control within an ecology of policy-rich but cooperating resources and actors.

- 5 To achieve this we propose an innovative model which uses standardised ontology languages such as OWL to articulate policy and marshal authentication reasoning, offering standardised interfaces to multiple service-ready reasoning systems building on RDF as an integration language. Through this we expect to demonstrate novel ways in which a variety of security policies can be defined and enforced utilizing a variety of distributed information across multiple institutions. The emphasis here is to show the benefits of semantic technologies for the expression and enforcement of security policies. We are not the first to talk of ontology-based access control: the KAoS framework [KAoS] is a framework for articulating authorisation policies using OWL. Our innovation is to use this to integrate secure assertions from multiple sources using different (emerging standard) technologies (see architecture below, Fig. 1), in such a way that integration with an existing Policy Enforcement Point (PEP) is as simple as possible, using a RESTful interface which is integrable with any system which can GET from, and POST to, URLs.

2.1 Present Grid infrastructures – multiple attribute sources

- 6 To understand why higher levels of abstraction are needed consider the situation with today's Grid infrastructures, and the web more generally. These are characterised by an explosion of security related solutions, where users will arrive bearing a heterogeneous variety of credentials including Shibboleth/UK Federation credentials, domain-specific certificates, or other supra-JISC warrants such as NVO 'Extensible, Scalable, Secure Service Infrastructure' (NESSI) credentials. Rather than attempting to provide high-level support for all of these, we should instead approach the problem from the explicitly practical angle of asking what authentication/authorisation-related information is potentially available to a web service's PEP, and only then asking how to retrieve and make use of it. This implies that we must avoid technology commitments, or delay the choice of technologies, as much as we can, and instead develop an access-control approach which is as flexible and as technology-neutral as possible. While SAML2 will inevitably be a major component of such an approach, we believe we would significantly and importantly reduce our options by depending implicitly on it. If a dependence on SAML requires important Identity Providers (IdP) and Service Providers (SP) to invest resources in creating SAML-aware endpoints, there is a danger that they will simply not bother.
- 7 For example, we might decide to allow access to a particular resource on the practical grounds that a requestor's home university library would allow them a certain local level of access. Such a criterion might be deemed realistic and secure enough for our purposes, and the relevant information might be already cheap to retrieve in some form (perhaps via a Secure LDAP lookup, or an existing partnership agreement). If, however, our system required the remote library or libraries to invest the time and resources required to install and configure a SAML responder, then they might reasonably decide that such an effort is insufficiently substantially in their or their users' interests as to make the marginal investment worth while.
- 8 In the same vein, a service might rationally invest most of its scarce resources into supporting the 'power users' it is funded for, perhaps with elaborate and non-standard access controls. However the service might at the same time wish, or find it expedient, to offer an alternative service to Grid or VO users, who might arrive with a heterogeneous variety of credentials, and whose access must be managed without elaborate and expensive formalities on either side. This is not perhaps the future envisaged by initiatives such as UKFederation, but by working alongside bespoke authorisation systems (with what one might call 'legacy approaches') we can open the door to more standard approaches, facilitating the use and reuse of SAML (and other) services, thus enhancing the impact and broadening the use of e-infrastructure initiatives.
- 9 Other relevant current sources of pseudo-authorisation information include X.509 end-entity certificates, Secure LDAP and the eduPerson schema, community-specific registries or VO management systems (eg, IVOA registries, VOMS), Shibboleth services, OpenID providers and attributes,

and even project-specific authentication systems (NESSI). Forthcoming sources include RFC3820 proxy certificates, the authorisation capabilities of the iRODS system attached to SMB servers, and potentially Grouper and UKFederation servers. Not all of these will provide SAML2 services in the medium term. However, simply using SAML2 as one important source amongst many lets a PDP take advantage of those investments, while still being able to use those sources for whom such an investment is out of reach. The National e-Science Centre at Glasgow have already successfully shown how a variety of sources of attributes and models of VOs can be used to enforce authorisation decisions by service providers. These include centralized VO models using VOMS and PERMIS to protect Globus toolkit version 4 services as part of the JISC funded VPman project. Decentralised VO models where the attributes needed for authorisation exist across a range of federated sites have also been demonstrated as part of the recently completed JISC funded DyVOSE project. We are thus well aware of the current ways in which VOs exploiting authorisation infrastructures can be established, used and managed – and their associated limitations!

- 10 A further advantage of this approach is that, because it is focused on, and starts with, the aim of exploiting existing heterogeneous sources of authorisation-relevant (attribute) information, federation and delegation capabilities come essentially for free.

2.2 Policy frameworks and sustainability

- 11 Existing policy languages (for example PERMIS, XACML and others) are based fundamentally on logic and rule systems. NeSC have considerable experience with the user interfaces (UI) to the use and administration of these systems, and have in the past provided usability feedback for end users, Grid developers, VOrg-administrators and local resource provider system administrators. The OWL language, which is what we propose to have at the core of our exemplar systems, is fundamentally based on the ideas of set theory, and we hope to discover whether, as we anticipate, users find UIs based on this approach easier to manage.
- 12 We believe that, in the context in which relevant projects are operating, the integration of services and authorisation can best be achieved by an approach which composes reasonably orthogonal technologies. Our aim therefore is to develop validated development patterns, and the software we develop will be designed to support these patterns, rather than act as another separately integrable service. This in turn implies that usability and security validation of these will be an important aspect of our project, and so propose to work closely with existing projects with outstanding and clearly articulated access-control requirements. In particular, we will work closely with the grid security group of the International Virtual Observatory Alliance (IVOA, <http://www.ivoa.net>), of which the STFC-supported AstroGrid project is a founder member. If successful, the project outcomes will naturally be embedded in, and sustained by, the Virtual Observatory (VObs) projects; the IVOA is conceived as a long-term standards-setting body for Astronomy, which is a field with a history of managing standards over timescales ranging from decades to millennia. Furthermore, through the direct involvement of NeSC in a range of large scale security-oriented projects we will explore the advantages and disadvantages of this approach through other non SemWeb-based approaches. For example, the VOTES project (Sect. 3.4) requires fine-grained authorisation in a clinical context, and the nanoCMOS project (Sect. 3.5) requires IP protection in a commercial and industrial context; these projects thus represent real test cases for the application of the SemWeb-based approach.

2.3 Architecture

- 13 The basic architecture we envisage is illustrated in Fig. 1. The Policy Decision Point (PDP) is a generic service, which would be little more than a reasoner, acting on an OWL policy provided by the PEP, and queried using the standard SPARQL RDF query language [SPARQL]. The policy defines the logical classes of individuals who are allowed access, in terms of the semantics produced by the translators. The Access Control Information (ACI) would be of various types, from various sources,

retrieved by the PEP, and converted to RDF by a layer of per-format translators which we would develop. These translators articulate the semantics of the underlying ACI (for example SAML), as well as generic information using the ubiquitous 'Friend-of-a-Friend' ontology (known as FOAF, and containing concepts such as names, addresses, and other similar information), or other reused ontologies, since these might be all that the OWL policy requires.

- 14 It is the PEP that is responsible for assuring the integrity of the remote assertions and the secure retrieval of attributes, perhaps using a certificate attached to the access request, but it is the translators and the Policy that are responsible for giving them meaning, within the PDP. As an example, the PEP's SPARQL query might be as simple as: 'is the person with email address foo@example.org in the class of individuals allowed access to the resource?' (placing the responsibility for the access decision on the entity which defines the class), or alternatively the SPARQL query might be sufficiently rich that it can itself be regarded as an expression of the desired policy (making the caller more clearly responsible for the policy). It is the clear orthogonality between the PEP, the translators, the Policy, and the remote data sources that gives the approach its robustness and generality. The Policy is obviously specific to the access control situation, and the PEP is also (because it needs to be at least minimally aware what sort of information is, or is likely to be, relevant to the Policy), but the other components will be generic. All the communications in the architecture above are either standardised (for example with LDAP and SAML) or else simple RESTful interfaces to new services.

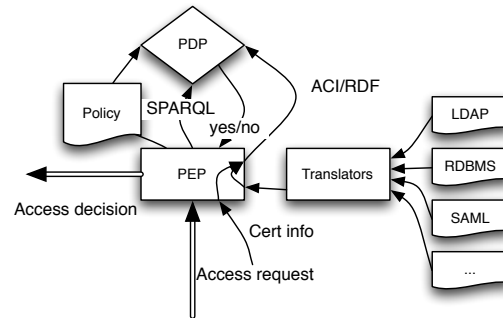


Figure 1: Semantic Technology-based Security Infrastructure

2.4 Novel authentication and attribute technologies

- 15 The most widespread Single Sign-On (SSO) credentials are X.509 certificates (including proxy certificates to RFC3820), with users identified by X.500 distinguished-names. Despite this standardisation, there are acknowledged usability difficulties in managing X.509 certificates, and ongoing development interest in other possibilities, including OpenID, XRI or CardSpace. As well, there are multiple technologies for attribute exchange, including X.509 ACs, SAML, and OpenID Attribute Exchange.
- 16 These various technologies for authentication and attribute exchange have overlapping but distinct use-cases, target markets, and security requirements. Notably, more modest security goals typically result in substantially simpler protocols (at the level of both end users and developers) and a more immediately practical user experience. We believe that the authorisation architecture we propose allows service providers to support services located wherever on this continuum is most appropriate to the provider's goals and audience.
- 17 We believe that the authorisation architecture should be agnostic as regards the authentication mechanisms, and that RDF can in this context come into its own as a neutral integration framework. Thus the authorisation architecture will consist of: one or more clients which will make SPARQL queries [SPARQL] which immediately imply go/no-go decisions; the reasoner which will add semantics to the data in a triple-store; and the triple store which will act as the repository for the possibly heterogeneous sources of relevant assertions. The assertions in the triple store will be added, in a conceptually separate process, by plugins reprocessing X.509 certificates, SAML assertions, LDAP information, or even information obtained from static files, if that is how a particular site can best express it (see Fig. 1).

3 Case studies

- 18 Any novel security approach based upon semantic technologies needs to be tested in realistic domains where realistically challenging policies are required, and needs to demonstrate added value, for example in supporting scenarios that cannot easily be supported by existing authorisation infrastructures. It is our intention to explore semantic technology solutions across a diverse range of representative e-Research applications, and produce a documented evaluation of the overall development and management effort required.

3.1 Astronomy: Virtual Observatory

- 19 Most astronomical data starts off proprietary to an instrument consortium, and becomes public after some longer or shorter interval. VObs projects have so far restricted themselves to the easy case of public data, but although there is increasing progress with SSO authentication, the immediate plans restrict themselves to simple all-or-nothing authorisation policies. A recent survey [UCSRV] of use-cases established that the astronomically relevant cases were generally logically reasonably simple (they did not, for example, require support for negotiation or have elaborate privacy requirements), but nonetheless challenging in being heavily distributed, or in requiring information of multiple types. Relevant use-cases include the following.
- 20 *Case 1:* An archive makes its full contents available to members of a specific collaboration plus researchers based in the UK, and makes available a flagged subset of the archive to researchers based in African countries. Satisfying this requires support for group identity, geographical information about institutions, and being able to reason that, for example, Egypt is in Africa.
- 21 *Case 2:* Users may be subject to complicated quota-type constraints, allowing them to use 100GB for a week, or 1TB for a day, or 10TB while a job is running. This demands flexibility of specification, and the capability of ingesting multiple sources of information.

3.2 Neurological science: GLASS

- 22 The JISC funded GLASS project (Glasgow early adoption of Shibboleth, <http://www.nesc.ac.uk/hub/projects/glass>) began in March 2006 and is exploring the roll-out of Shibboleth across the University of Glasgow, building upon the Novell nSure unified account management system. GLASS is exploring numerous scenarios where centralized authentication and fine grained authorisation are required, and has already demonstrated prototypes showing single sign-on, and fine grained authorisation decisions, via Shibboleth and the use of a variety of attributes. One of these is within the brain trauma area working closely with the Southern General Hospital in Glasgow – specifically to support the European-wide BrainIT network (<http://www.brainit.org>). Existing prototypes showing single sign-on via Shibboleth and use of a variety of attributes for fine grained authorisation decisions have already been demonstrated.
- 23 Currently the GLASS BrainIT demonstrator focuses upon attributes delivered from a single IdP (at Glasgow) to restrict access to federated neurological data sets. This SemWeb case study is focused upon the semantic representation of the existing security policies: as well as supporting the existing policies, we plan to show how further federated security information can be used to decide local access policies, including checking for potential inconsistencies of security policies across sites (scientists should not be both an investigator and an ethical review body member within a given neurological trial); or checking that security policies are interpreted consistently across sites. Logical associations between roles can allow more intelligent authorisation decisions: rather than insisting that every site supports a centrally-defined ‘investigator’ role, logical equivalents or sub-classes of investigator might also be acceptable.

3.3 Bioinformatics: BRIDGES

- 24 The DTI funded BRIDGES project (*Biomedical Research Informatics Delivered by Grid Enabled Services*, <http://www.nesc.ac.uk/hub/projects/bridges>) completed at the end of 2005. BRIDGES

developed both a large scale data Grid infrastructure linking many genomic data resources, and a compute Grid infrastructure allowing simple access to numerous compute resources including all nodes of the NGS, ScotGrid (<http://www.scotgrid.ac.uk>) and local Condor pools. The project specified and enforced fine grained security policies (distinguishing unknown, known, and known-and-trusted user groups) to restrict which computational resources the end users were allowed to access. These solutions did not require end users to have their own X.509 digital certificates, but instead used server certificates for job submission and management.

- 25 To access large-scale resources such as NGS and ScotGrid in future, the current Globus-based BRIDGES implementation will require finer-grained policies, depending on (distributed) authorisation and Grid environment information to schedule jobs and allocate resources. This demonstrator will show how these policies can be represented using semantic technologies, and how we can add value by showing how to do load balancing as a function of the resources already consumed by that user, that VOrg, and that institution.

3.4 Clinical Science: VOTES

- 26 The MRC funded project VOTES (*Virtual Organisations for Trials and Epidemiological Studies*, <http://www.nesc.ac.uk/hub/projects/votes>) is a three-year project looking at building a Grid framework for clinical trials and observational studies. The project began in October 2005 and involves the National e-Science Centre at the University of Glasgow and partners at Oxford, Nottingham, Leicester and Imperial College. Clinical trials and clinical systems more generally place many and various demands upon security infrastructures to support the various activities of recruitment, data collection, and overall management of the trial itself.

- 27 Fine grained security is essential in this context to ensure that the right data is made available to the right people for the right purpose. A key aspect of the work is that VOTES is not concerned with developing a single Grid infrastructure for a specific clinical trial or study, but with developing a Grid based framework through which a multitude of clinical trials can be supported. Versions of this framework utilizing GT4, OGSA-DAI, GridSphere and VOMS and PERMIS have already been demonstrated. In this demonstrator we will show not only how the Semantic Web can support the existing policies on access and usage of the services, but importantly, what they can add that cannot easily be supported right now.

3.5 Engineering: nanoCMOS

- 28 The EPSRC pilot project nanoCMOS (*Meeting the Design Challenges of nanoCMOS Electronics*, <http://www.nesc.ac.uk/hub/projects/nanocmos>) began in October 2006. The project is led by the University of Glasgow with the e-Science component led by NeSC Glasgow. It provides a rich vein of scenarios which will stress test the SemWeb-based approaches in this proposal.

- 29 The electronics domain demands infrastructures that support IP protection, be it for designs of transistors, data sets, or quantum-mechanical simulation codes. The nanoCMOS project has developed an OMII-UK-based Grid infrastructure – providing secure access to large scale HPC resources such as the NGS – through which device level designs and simulations can be linked through to higher level circuit and system simulations, to predict the overall behaviour of systems.

- 30 This scenario presents challenging and legally-charged policy decision problems concerning the access to and usage of HPC compute and data services; for example, it is often the case that data and simulations in this domain are not allowed to run on the NGS for IP reasons. We must also be able to support potential changes in such policies, as new resources are added or old resources removed, or else simulations and data are produced ‘on the fly’ and should only be made available to those in possession of newly defined roles. Scenarios such as these have proved difficult to support with existing authorisation systems, and we hope to show how SemWeb-based approaches can do better.

4 Appropriateness to this call

4.1 Fit to programme objectives

- 31 This project addresses a number of key points in the call, since by supporting flexible access control it will encourage **broader and more effective use of the e-infrastructure** and provide feasible **enhanced security**. As a consequence, it will provide **new ways of retrieving and processing (proprietary) data**. Its use-cases will explore **integration with other key initiatives** in astronomy, biosciences and materials science. These use-cases will help us implement **sophisticated access management for data** in real cases which require that we **examine both access roles and access rights** in cases where the rules are variously **complex**, making use of **delegated credentials** and attributes. The project already has a proof-of-concept implementation.
- 32 The project's key outcomes include **better understanding, via use-cases**, by investigating the potential of **new ways of thinking about authorisation**. As such, it will **inform future developments** and provide **guidance to the sector** (both via project final report).

4.2 Value to JISC community

- 33 The core value to the JISC community will come from investigating a novel approach to the expression of security policies, which focuses on the exploitation of **existing distributed attributes** with **semantic services**. This investigation will require our articulation of realistic multi-domain **use-cases**, and producing **demonstrators** to challenge the approach.
- 34 We intend, as part of this project, to undertake extensive **dissemination** activities, describing the approach and its **evaluation**. In addition, we will make presentations and submit papers to the various relevant forums. As well, we will be able to provide **feedback** to JISC on the applicability of semantic technologies in this area, in the context of already existing applications.

5 Workplan

- 35 The project will be managed by Gray, according to JISC standards and procedures. After an initial generic development phase, primarily at Leicester, the focus of the project will move to the application of the work to the various case-studies, primarily at NeSC, before a final joint consolidation and dissemination phase. We assume a start on 1 February 2008.

Work Package WP1 Project management

- 36 The project manager will ensure appropriate coordination between the various project goals, liaising with Prof Sinnott and the various RAs working on WP4, and helping drive the dissemination efforts of WP5. The allocated staff time (totalling 16sm) is rather thinly spread over the 12 month duration of the project, to give adequate flexibility to the scheduling of the various WP4 tasks within their respective projects. This has the side-effect of giving the project a good deal of flexibility in the face of unforeseen events.

5.1 Phase 1: scoping and architectural elaboration (T0-T8)

- 37 The initial scoping work will focus on capturing the requirements from the various application domains. We note that through the PIs' existing knowledge of these on-going projects, we are well placed to identify core case-studies, as well as identifying more advanced and challenging possibilities. This phase will identify the state of the SemWeb art, available repurposable ontologies, and relevant tools. As well, this will take advantage of an existing proof-of-concept implementation by Gray, and elaborate through its overlap with WP4.

WP2 User requirements and scoping (1sm Leicester/1sm NeSC)

Deliverable D2.1 A document describing the case-studies, their background, and why they have been chosen for semantic elaboration.

D2.2 A document describing the state of the art in semantic technologies, relevant ontologies for expressing security policy, and associated tool support.

WP3 Elaboration of architecture into initial prototype (4sm Leicester, 0sm NeSC)

D3.1 Definition of architecture and core components

D3.2 Initial examples of representative policies in OWL, **D3.3** initial service prototype, comprising PDP and immediately useful translators, and **D3.4** initial integration of simple PEP client with service.

5.2 Phase 2: Implementation of case studies (T4-T11)

- 38 In this phase, the prototyping work of WP2 will be confronted with the identified use-cases. This will drive the development and extension of the prototype WP3 deliverables (the project will of necessity use an agile, rapid prototyping, development methodology). All but one of the case-studies are hosted at NeSC, the other at Leicester. This is a ‘long thin’ phase, flexibly scheduled, since in all cases, the work will be done by project-specialist developers (to be identified), at appropriate points in those projects’ timelines.

WP4 Demonstrators to validate prototype (2sm/4x1sm)

Tasks 4.1–5 VObs access control application, and case-study support (Leicester); BrainIT, BRIDGES, VOTES and nanoCMOS demonstrators (NeSC)

D4.1–5 Software demonstrators for each of the noted case-studies.

D4.6 Reports on the practical experiences of SemWeb technologies in building these demonstrators and the consequences thereof for Grid users, developers and resource managers.

5.3 Phase 3: Dissemination (T6-12)

- 39 As well as the deliverable reports, the project will prepare papers for relevant conferences, workshops and other activities. Although these will predominantly appear towards the end of the project, the interaction with the other projects in the flexibly-scheduled phase 2 means that these might occur earlier.

WP5 Dissemination (2sm/2sm)

D5.1 Papers for international grid conferences describing how semantic technologies can benefit Grid security; contributions to discipline-specific publication routes (for example, IVOA Notes).

D5.2 Document describing the overall lessons learned in developing and enforcing semantic security infrastructures and their linkage to Grid infrastructures.

D5.3 We intend to organise an e-Science Institute workshop on the area of semantic authorisation technologies, pulling together the grid, security and semantic communities, to develop shared strengths, facilitate the exchange of ideas, and demo project results.

D5.4 AGAST final report.

5.4 Risk Analysis

- 40 *Risk: technology base and standards change (high probability, low impact).* The authorisation and attribute-exchange technology which this project is targetting is in active development. However, part of the justification for our SemWeb-style approach is precisely that it gives us agility and adaptability in the face of such change, so that we would go so far as to welcome such a change.
- 41 *Risk: multiple project dependencies affect scheduling (low, medium).* We have scheduled 16sm of effort over a 12 month project, distributed over multiple staff, working on existing e-Science projects in parallel. The consequent timetabling risk can be mitigated by identifying relevant staff commitments as early as possible; most of the relevant staff are working on projects already managed by Prof Sinnott.
- 42 *Risk: personnel changes (medium, low).* With the exception of the PIs, we are depending on interactions with projects rather than individuals. None of the relevant projects are single-person efforts, and we are therefore insulated to a large extent from problems due to personnel changes.
- 43 *Risk: distributed teams (high, low).* The key personnel have had long experience of working

	Feb08–Mar08	Apr08–Jan09	Total
Directly Incurred Staff			
Leicester: Dr Norman Gray, 25%	XXXX	XXXX	XXXX
NeSC: Prof Richard Sinnott, 8%	XXXX	XXXX	XXXX
Leicester: RA, grade 7, 50%	XXXX	XXXX	XXXX
NeSC: RA, grade 7, 50%	XXXX	XXXX	XXXX
Total directly incurred staff (A)	XXXX	XXXX	XXXX
Non-staff			
Travel and expenses	667	3333	4000
Hardware and software	5000	0	5000
Total directly incurred non-staff (B)	5667	3333	9000
Directly incurred total ($C = A + B$)	XXXX	XXXX	XXXX
Directly allocated			
Staff, Leics: System admin (10%)	931	4656	5587
Staff, Leics: secretarial (10%)	659	3295	3954
Estates & Lab	3893	19467	23360
Directly allocated total (D)	5483	27418	32901
Indirect costs, Leicester	5552	27760	33312
Indirect costs, NeSC	2740	13701	16441
Indirect costs, total (E)	8292	41461	49753
Total project cost ($C + D + E$)	XXXX	XXXX	XXXX
Leicester contribution (DA staff costs, inc. associated indirects)	-XXXX	-XXXX	-XXXX
Amount requested from JISC	26273	106364	132637

Table 1: AGAST budget (all figures in £)

in distributed projects. Dr Gray is a regular visitor to Glasgow working with colleagues in the astrophysics domain. Numerous discussions have taken place with Prof Sinnott at Glasgow on these technologies and their potential benefit for future Grid security models.

- 44 *Risk: benefits do not generalise (low, high).* We have in our case-studies a demandingly broad range of application domains, in different e-Research areas, and hence can directly explore generalisation risks. Further, because it is a goal of this project to minimise technological interdependencies, we would expect even non-generalized demonstrators to be useful. If we discover that the semantic technology *pattern* (in software engineering terms) is more important than any individual implementation, then this will be a valuable outcome of this project.
- 45 *Risk: Sustainability.* One goal of the project is to develop the access-control system recommended by the IVOA to resource owners worldwide. If clear benefits of SemWeb technologies for security have been visibly demonstrated then we fully expect this to will lead to wider take-up of these approaches by the Grid community. Prof Sinnott is on the UK e-Science User and Security Task Forces, the OMII-UK User Group, and is chair of the OGSA-DAI User Group. All of these channels of communication will be used to present experiences and recommendations from the project thereby promoting both visibility and potential sustainability.

6 Budget and justification

- 46 We note that this project represents excellent value for money due to the expertise that the project partners bring, and importantly our ability to capitalise on the variety of on-going and significant e-Science projects upon to which the PIs have access. The primary RAs will be supported by other

researchers at NeSC and Leicester, who will contribute their efforts to ensure that the overall goals of the project are realized. As well, the primary contact is embedded within the international VObs community, addressing a pressing problem, and can therefore anticipate support, both in specification and implementation, at all stages of the project. We believe that this represents a unique and cost effective opportunity to explore how semantic technologies can benefit UK e-Science and wider efforts such as the e-infrastructure, and that it gives JISC-funded research an agenda-setting position in the international VObs.

47 To ensure the project achieves all of its objectives, we seek funding for researchers at Leicester (3sm Dr Gray, plus 6sm other RA) and at NeSC (1sm Prof Sinnott, plus 6sm one or more per-project expert). Given the requirements for a high level of experience in Grid technologies across a range of application domains and in semantic web technologies, we require researchers with a high level of IT competence.

48 We expect to widely disseminate the work both nationally and internationally in Grid/e-Science, Semantic Web *and* the application disciplines (Astronomy, Neuroscience and Bioinformatics). This includes the UK e-Science AHM; the six-monthly IVOA interoperability workshops; the International and European Semantic Web conferences, and respected international conferences such as the IEEE Grid conference and/or Cluster Computing Grid conference (CCGrid 2007). This breadth implies a relatively large travel component to the proposal, and we request £4000 in total.

7 Personnel

49 Dr Norman Gray is a researcher for the European Virtual Observatory's VOTech project, and is based at the University of Leicester. Prior to this, he was part of the Starlink Project, a long-running PPARC-supported distributed astronomical software project, developing and maintaining a large volume of astronomical data-analysis code, and through that he was involved in VO projects almost since their inception. Dr Gray brings to this project long experience of scientific software engineering, knowledge of the VO context, and Semantic Web development experience.

50 Professor Richard Sinnott is the Technical Director for the National e-Science Centre (NeSC) at the University of Glasgow. He has over 70 book chapters, journal and conference publications and has led a range of Grid security projects including DyVOSE*, BRIDGES*, ESP-Grid*, GEMEPS, GEODE, GLASS, VOTES, GHI, nanoCMOS*, SBRN* and UK e-Science ETF* projects (see <http://www.nesc.ac.uk/hub/projects>). The starred projects are concerned particularly with the security issues and solutions associated with HPC-based access to Grid infrastructures. For example, within the BRIDGES project he has led the implementation of solutions showing how an advanced authorisation infrastructure based upon PERMIS can be used to provide both data security and compute security. Drawing on this pool of highly relevant Grid security focused projects makes Prof Sinnott an ideal person to lead the demonstrators of this project. Prior to coming to Glasgow, Professor Sinnott ran his own consultancy company in the area of real time distributed systems especially in the telecommunications domain. Prior to this he was employed as a senior research scientist by the GMD Fokus, Berlin to work on and manage numerous major international projects involving companies and research institutions from around Europe, the United States and Japan. Professor Sinnott has a broad background in open distributed processing systems and acted as editor for numerous international standards in this area.

Appendix: References

[KAOS] KAoS Framework, Florida Institute for Human and Machine Cognition, <http://ontology.ihmc.us/kaos.html>, 2004.

[SPARQL] SPARQL Query Language for RDF. W3C Working Draft, 4 October 2006. <http://www.w3.org/TR/rdf-sparql-query/>, 2006.

[UCSRV] Access-control use-case survey. Euro-VOTech planning wiki, <http://wiki.eurovotech.org/twiki/bin/view/VOTech/AccessControlUseCases>, 2006.



**University of
Leicester**

**Department of
Physics & Astronomy**

University Road
Leicester LE1 7RH · UK
Tel: +44 (0)116 252 3574 (*Head of Dept*)
Tel: +44 (0)116 252 3492 (*Direct Line*)
Tel: +44 (0)116 252 3575 (*Dept Office*)
Fax: +44 (0)116 252 3311
Email: mab@star.le.ac.uk

Head of Department &
Professor of Astrophysics & Space Science
Professor Martin A Barstow
BA, PhD, CSci, CPhys, FInst.P, FRAS

28 September 2007

JISC Executive
Northavon House
Coldharbour Land
Bristol
BS16 1QD

Dear Sir/Madam

The Physics and Astronomy department of the University of Leicester is keen to explore the possibilities of using semantic technology to facilitate access management for archives. Leicester is one of the leading institutions in the AstroGrid and VO-Tech projects, and has a leading role in astronomical data archive services through the LEDAS system, which has been at the forefront of on-line data access in the UK for a number of years.

We also operate the XMM-Newton Science Survey Centre on behalf of the European Space Agency. Consequently, we have a strong interest in cost-effective federation of authorisation, and have colleagues with world-class expertise in the archiving field, who can support the project's astronomical use-case with advice.

We wholeheartedly endorse this proposed research programme and will give it the necessary departmental support if it is successful.

Yours faithfully

PROFESSOR M A BARSTOW
Head of Department



National Grid Service
Grid Operations and Support Centre
CCLRC e-Science Centre
Rutherford Appleton Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QX

Prof. Richard Sinnot
National e-Science Centre
Glasgow
Scotland

I am writing on behalf of the National Grid Service in support of the 'AGAST – Advanced Grid Authorisation through Semantic Technologies'.

The NGS would like to support this project as it has the potential to draw on experience from existing projects and knowledge of developments and current requirements within the UK Grid community and develop services or solutions that may enable the NGS to move further forwards in providing fine grained access control for end users from a variety of different communities. Solutions that can assist the NGS in providing a service that is attractive to a multitude of user communities will be of great benefit in developing and delivering a long term national research infrastructure.

Yours Sincerely

A handwritten signature in black ink that reads 'A. Richards'. The signature is written in a cursive style and is set against a light beige rectangular background.

Dr Andrew Richards
Executive Director, National Grid Service