

# Security Needs and Usability in e-Science Projects

*Final Report, 16 June 2005*

M. Angela Sasse  
Brock Craft  
Department of Computer Science  
University College London



# CONTENTS

<b>1. EXECUTIVE SUMMARY</b>	<b>1</b>
<b>2. METHOD</b>	<b>2</b>
2.1 Data Collection	2
2.2 Data Analysis	2
<b>3. FINDINGS</b>	<b>3</b>
3.1 Who is a GRID User?	3
3.2 Areas of Success	5
3.3 Documentation	5
3.3.1 General Findings	5
3.3.2 Specific Recommendations	6
3.4 Authentication	7
3.4.1 General findings	7
3.4.2 Subversion of authentication	7
3.4.3 Passwords	8
3.4.4 The “Price of Admission”	8
3.4.5 Digital Certificates	9
3.4.5.1 General findings	9
3.4.5.2 Conceptually Difficult	10
3.4.5.3 Poorly Documented	11
3.4.5.4 Difficult to obtain	11
3.4.5.5 Difficult to import and export	13
3.4.5.7 Break easily and are hard to fix	14
3.4.6 Specific Recommendations	15
3.5 Frameworks and Toolkits	15
3.5.1 Web Services	15
3.5.2 GLOBUS Toolkits	17
3.5.3 Specific Recommendations	20
3.6 Authorisation	20
3.6.1 General Findings	20
3.6.2 Delegation	21
3.6.3 Specific Recommendations	22
3.7 Confidentiality and Privacy	22
3.7.1 General Findings	22
3.7.2 Specific Recommendations	24
3.8 Provenance, Auditing, and Versioning	24
3.8.1 General Findings	24
3.8.2 Specific Recommendations	26
3.9 Availability	26

# CONTENTS

<b>3.10 Security in the Development Lifecycle</b>	<b>27</b>
3.10.1 Security Not a Functional Requirement	27
3.10.2 Security vs. Usability	27
3.10.3 Stewardship	28
<b>4. SUMMARY</b>	<b>28</b>
<b>4.1 General Recommendations</b>	<b>28</b>
<b>Table 1: A Framework for Reasoning about Security Information to Provide to UK National GRID User Groups</b>	7

# Security Needs and Usability in e-Science Projects

*“Ensure that in the quest for security that the progress of uptake of e-Science is not damaged. Don’t allow yourselves to get obsessed with security over and above usability and outreach.”*

*-Study Respondent*

## 1. Executive Summary

The UK e-Science Security Task Force (STF) aims to ensure that UK National e-Science resources are adequately secured. This can only be achieved if all e-Science projects secure their assets adequately, and promote secure practices among all stakeholders (researchers, developers, administrators). As the opening quote of the report illustrates, security is currently often perceived as an obstacle to, rather than an enabler of, e-Science activity. The roots of this perception lie in failure of some stakeholders to recognise their own security needs, and negative experiences with current security policies and mechanisms. Recognising this, EPSRC and DTI commissioned a small survey to identify the security needs of e-Science projects, and usability issues with current security mechanisms. This document reports and interprets the findings from the survey, and identifies a number of measures to increase security awareness among e-Science stakeholders and facilitate the development of secure e-Science tools and practices.

The data have been drawn from interviews with members of the e-Science community, which can be roughly segregated into three categories: researchers, developers, and administrators. Interviews were conducted with members of these communities to gain understanding of the usability issues that people have experienced in their everyday interactions with security in e-Science technologies. Their opinions, when aggregated, yield emergent themes. Whilst much of the data indicate areas for improvement, there are security successes UK in e-Science initiatives. The findings lead to the recommendations in Section 3. Section 2 provides details about the method used for gathering the information that guided the formulation of the recommendations. The specific findings of the interviews, including comments that participants made about the various aspects of e-Science security are provided in Section 4.

## 2. Method

### 2.1 Data Collection

The data were collected by personal interviews conducted during a two-month period during March and April 2005. A cross section of the UK e-Science user community was identified, based upon available data at the National e-Science Centre about ongoing projects, and from information publicly available on e-Science Project websites. These candidates were contacted by email and phone with a request to give feedback to the investigators about their experiences using security in e-Science. Calls for participation were also posted to e-Science community mailing lists. Approximately 50 candidates were identified. This population was not formally randomized, but was determined to represent a reasonably broad spectrum of the members of the e-Science community.

From those initial candidates, 34 respondents were interviewed (n=34). The majority of these interviews were conducted one-on-one, at the respondent's location. Two focus-group style interviews, consisting of three respondents each, were also conducted. Additional data were collected by observation of a GRID security planning session that was held at a major UK University Computing Services Department. At this meeting, remarks pertinent to usability of security were noted in the data record.

Respondents were advised that their participation was voluntary and that they would receive no compensation. The interviews ranged from approximately 25 to 110 minutes in duration. As an aide to the data analysis, the interview sessions were recorded. Respondents gave written consent to be interviewed and to have their remarks recorded. The recordings were later transcribed and anonymised.

The interviews were performed according to a semi-structured or "guided interviewing" method. Prior to data-gathering, a selection of topics was prepared into a formal set of questions known to be relevant to e-Science security (based on previous studies of e-Science security conducted by Sacha Brostoff and Ivan Flechais at University College London). A researcher conducted the interviews by using the prepared questions to guide the flow and order of the conversation, but allowed discussion to diverge to areas of interest that were uncovered as the dialogue progressed.

The interviews were allowed to continue until the all of the prepared questions had been discussed or until the natural end of the conversation. At this point a final, open-ended final question was posed. To elicit general feedback, users were asked, "If there is anything we haven't covered that you would like to get back to the people who make decisions, what would that be?" This was devised to return opinions that might not have been covered during the course of the conversation, and to provide of voice for respondents to address the STF directly.

### 2.2 Data Analysis

Approximately 45 hours of recorded interviews were transcribed and anonymised. The duration and scope of this project required a rapid analysis method, using modified techniques adopted from Grounded Theory<sup>1</sup>. The transcribed conversations were reviewed and responses were coded by topic. The coding allowed responses similar in domain or subject matter to be grouped together. As the guided interviews provided a structure for the discourse during data gathering, data for particular areas of inquiry were generally obtained from several respondents. Resultant themes were identified and ordered according to the frequency of occurrence. Analysis techniques for statistical semantic clustering were not used.

---

<sup>1</sup> Anselm Strauss & Juliet Corbin: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage 1999.

In addition, heuristic knowledge was accumulated by the interviewer (Brock Craft) during the data collection process. This knowledge represents an aggregated impression of the broad, recurrent themes. Together with findings from previous studies ( by Sacha Brostoff<sup>2</sup> and Ivan Flechais<sup>3</sup> under Angela Sasse's supervision), this played a role in deciding which subjects were most important to participants and most relevant for the purposes of this report. However, the frequency with which issues were mentioned was the main criterion for selecting issues for report.

In many cases, the opinions of the participants speak for themselves, requiring no interpretation. Responding to issues that directly relate to their work and interaction from with the e-Science community, participants tended to very clearly articulate the security issues they have encountered.

### 3. Findings

This section documents the research findings. Of course, not all of the findings are specific to e-Science projects. Where only one or two respondents made an observation about a particular problem, only very significant issues that present profound problems for usability were included. When appropriate, recommendations for the STF have been included. These *Specific Recommendations* are relevant to problems identified in the several topic areas and are intended to alleviate specific usability issues that might not be addressed by the more general recommendations in Section 2.

#### 3.1 Who is a GRID User?

Due to the distributed, varied nature of e-Science technology and activities, e-Science means different things to different participants. Their understanding depends on their level of experience of and interaction with people and technologies that could be considered e-Science. Thus, members of the e-Science community have very different levels of knowledge about security, and different experiences in their interactions with security policies and technologies. Research *researchers* are less likely to be familiar with tools such as the GLOBUS toolkit than *developers* of middleware. System *administrators* have different tasks and issues from researchers even when interacting with the same security mechanisms – for instance, digital certificates. Project and site *managers* also play an important role in security that is often overlooked

The e-Science community has a number of stakeholders groups who interact with security:

- **Researchers**

The *raison d'être* of e-Science are the researchers that use e-Science tools and services to discover new knowledge or improve research methods. They tend to have a limited understanding of the e-Science technologies they use and are focused on completing their scientific research. Security is acknowledged to be important, but past experiences of security "getting in the way" of research tasks leads to a certain amount of apprehension and weariness (as exemplified by the opening quote). These stakeholders are often are not aware of the risks that their assets face, do not know why they have to use a particular security mechanism, and often have inaccurate understanding of what security terms actually mean. The data provide evidence of incorrect or imprecise use of security terminology (e.g. "Digital Certificate", "key", "password", and "passphrase" are often used interchangeably). Researchers often have a vague notion of what e-Science or "Grid" actually means; it is very common for them to conceptualise the GRID based upon a custom-authored tool that they use to access e-Science resources. For many, their

---

<sup>2</sup> Sacha Brostoff, M. Angela Sasse, David Chadwick, James Cunningham, Uche Mbanaso, Sassa Otenko: "R-What?" Development of a Role-Based Access Control (RBAC) Policy-Writing Tool for e-Scientists. To appear in Software - Practice and Experience.

<sup>3</sup> Ivan Flechais: Appropriate and Effective Guidance for Information Security. Unpublished PhD thesis, Department of Computer Science, UCL, 2005.

application is the "Grid", and they are not aware of the underlying infrastructure and resources, and organisational framework on which the functioning of that application depends.

- **Developers**

Developers build e-Science applications and application suites, such as AstroGRID, CancerGRID, and RealityGRID, or middleware developers who write the tools such as OGSA-DAI, OMII, and Web Services. Developers tend to be very knowledgeable about the specific technologies they use, and have a detailed understanding of the requirements of their customers – the scientists. Indeed, many developers are former scientists; they may be self-taught programmers or have some degree of formal training in software development. Many projects employ trained software developers, and some even security experts. Most developers, however, do not have particular expertise in security, or access to it. Nevertheless, they are the ones who decide on, and implement, security in e-Science projects. To accomplish this task, they rely on documentation and a network of other developers to solve security problems. At some sites, there is a good line of communication between developers and administrators, whereas at others, security issues are only identified once early versions of an e-Science application are deployed.

- **Administrators**

Administrators are the individuals responsible for managing the computer systems and networks on which e-Science projects rely in their day-to-day operation. Some administrators are involved in management of Campus grids or Condor Pools and may have cultivated relationships with developers in the course of their work. They are used to dealing with user problems and are typically and researchers' first point of contact if they encounter problems; they are thus very aware of problems that scientists have with security. Whilst they are typically eager to help scientists, they can become overwhelmed. With the exception of the administrators running the National GRID Service (NGS), most have additional responsibilities for management and maintenance of information technology assets in their institutions. Administrators are generally very knowledgeable about security issues and intimately involved in installing and configuring e-Science technologies. They are very hands on, but often find they do not have the time to plan ahead or pro-actively seek contact with scientist and developers and identify future security needs.

- **Managers**

This group consists largely of principal investigators, Project Managers, and development team leaders. Managers were often the originators of the grants that fund the projects that they are currently responsible for and therefore, have a very personal stake invested in a project's outcome. They come from a variety of backgrounds – many are (former) scientists - and may have little knowledge of e-Science technology and limited awareness of security, and often see it as a "technical" issue, to be handled by technical experts (developers). Managers are often involved in the administration and management at their institutions, looking at the "big picture" and thus would have opportunities to address security issues across individual projects. Some managers have more detailed knowledge of the technology that supports e-Science, but have widely varying levels of experience with security mechanisms. They are concerned about usability and economics of security, since these could affect the uptake and continuing use of e-Science projects.

## 3.2 Areas of Success

The opportunity that e-Science offers, and the intense efforts to by developers to produce useful, useable e-Science tools and services has yielded a number of success stories. Many e-Science projects have begun to fulfil the promise that distributed parallel computation, distributed storage, and virtual organisations have held out. Participants frequently described these accomplishments in their feedback:

Regarding AccessGRID:

*“They love it! Users absolutely love it. It enables them, certainly for very large collaborations to meet very easily and simply without having to do massive amounts of travelling.”*

Regarding the NGS, in general:

*“For me it was very, very useful to have access to NGS, so far.”*

Regarding computational clusters:

*“Access to the cluster...was Eden, for me.”*

Other e-Science tools such as Condor Pools, GRID-FTP, and GSI-SSH have received praise from all four stakeholder groups. Also, respondents generally rated the availability of GRID infrastructure as “very good”, and reported that support for scientists from local administrators and the Grid Operations Support Centre has, in general, been very helpful.

Although this research was not intended to identify such successes, the participants offered many positive examples of projects and services that are actively being used by researchers and researchers in their work. On the whole, they are hopeful and enthusiastic about the progress of e-Science. Notwithstanding the numerous instances up unsolicited positive feedback, there remain some important areas for improvement in the usability of security in e-Science applications, middleware, and services. The following sections detail the significant areas for improvement that were identified in the course of the research.

## 3.3 Documentation

### 3.3.1 General Findings

All participants, regardless of their role, indicated a need for improved documentation. The existing documentation is reported to be “hard to find”, “poorly written” and “difficult to understand”. Many of the comments regarding documentation are pertinent to specific technologies such as Digital Certificates, Web Services, and the GLOBUS Toolkits. Those comments are vital to understanding frustrations with the particular tools. In order to keep the comments in-context, findings about the documentation of specific tools and services are given in the sections of this report that describe those topics. Respondents at all levels were consistently frustrated by a lack of information they needed to complete their tasks. Typical remarks ranged from the mildly frustrated, such as:

*“...good documentation is essential and I couldn't find it.”*

*“So, I didn't have a good mental model of that and I was just sort of skimming web pages and it never really popped out and I was hoping it would because I'm sure it's all quite straightforward.”*

*“I think there's a general lack of information about security on the GRID. That is a general perception I have. Maybe it's just I haven't looked hard enough, or whatever.”*

to the highly critical:

*“The last thing I read from the Security Task Force was a document and it basically concluded that we can't make any conclusions, it's all a mess. And I didn't find that very helpful.”*

It is important to note that not all documentation is appropriate for all readers. Certainly, a technical specification for a developer must be written at a different level of specificity than a general guide intended for researchers new to e-Science. However, the findings of this study indicate that documentation was generally considered to be inadequate.

For developers, there are many websites that provide information about how to use specific tools such as Web Services, GLOBUS, OMII, etc., but there is a lack of information that describes the options for security that are generally available. As one manager of developers described the situation:

*“...we’ve got quite a big group of people on the project with quite a wide range of expertise, but in general, none of us know what’s out there. And that’s really been what’s held us back because we just haven’t even had any kind of basis on how to make decisions about how we should progress.”*

In this particular case, the respondent gave no particular example of what they were looking for, and what information was missing; in security, difficulties cited by users can sometimes serve as an excuse for not bothering with something one doesn't really want to engage with. But many comments described in subsequent sections of this report, identify specific problems encountered by developers using web services and GLOBUS, with the additional criticism that documentation was not thorough enough to be useful and was widely scattered.

Whilst developers complained that they could not find enough detailed information, researchers were overwhelmed by the assortment of unfamiliar concepts and terminology. These comments are typical:

*“Well I think it’s very, very important to have good documentation. And documentation that is written for people who are not used to these things. Because when I even started to use these things I didn’t even know what the GRID was.”*

*“I think that documentation is really important, or is it worth perhaps broadening that documentation and instruction and things pitched at a level that a non-specialist can manage to use without needing to spend several weeks or even several days reading.”*

### 3.3.2 Specific Recommendations

Recommendations for particular technologies are provided in the relevant sections of this document. General recommendations for improving documentation of security in GRID systems are that:

- **A unified source of security documentation should be available in the form of a website or other distribution mechanism. The information contained therein should be designed specifically for the target populations. It would act both as a centralised knowledge resource and means for community development.**
- **The availability of this resource to the e-Science community must be promoted, through advertising or other means.**
- **Security documentation must be targeted for each target audience (stakeholder group), and clearly state specific “do’s and don’ts”.**
- **All e-Science projects should allocate the time and a budget to author jargon-free security documentation for researchers and administrators.**

## 3.4 Authentication

### 3.4.1 General findings

In terms of the researchers' experience of security, authentication is the most significant mechanism. Whilst authentication methods for computing systems have been in place for several decades, few scientists have had to address the problems brought about by the complexity and sensitivity of e-Science technologies.

At present, the authentication schemes used by almost all UK e-Science projects can be classed into two major areas: those that are Web Services based, and those that use PKI systems, such as Digital Certificates. Specific usability problems with those two technologies are addressed in separate sections. But there are some significant usability issues with authentication, regardless of the underlying technology.

### 3.4.2 Subversion of authentication

The security offered by both Digital Certificates and Web Services can be subverted if the user knowingly or unknowingly engages in behaviour that compromises those methods. Common examples are sharing passwords and certificates, sending information via a non-secure channel such as email, or remotely accessing secured systems via a back door.

Many stakeholders are aware of ways to subvert authentication. The following remarks from different researchers, developers and administrators and illustrate this:

*"Yeah, I'm sure [password sharing] happens all the time for most of these systems. Even if it's not password sharing, it's very easy to take a quick abstract of the data and stick it in an Excel spreadsheet."*

*"And the only way that they could get this off the ground in time was: one person got a Certificate and they used that Certificate everywhere, basically."*

*"Well, just once, I got to do it. Because we had to do something in that moment and we had to share [a certificate]. That's it."*

*"So they said to one of their colleagues – this is someone within the University, both people worked in the University – they said, "You have a Certificate that has been authorized on this particular resource that I need. I need to do this thing for my boss by Friday. Today is Wednesday. I don't know what's gone wrong with my Certificate. There's no way I'm going to get it sorted out in two days. Please, can I use your Certificate?" And their colleague said, "Yeah, sure. Here it is." In fact, what he did was he copied it over and installed it for them in the correct place and removed the password from the copy."*

*"But yeah, maybe he could log into your machine remotely and then change user to be you and then he would have your certificate and he would go."*

*"But you can also do thing, you could set up like a VNC and give them a virtual desktop into your machine or any of the number of technologies that do this."*

Fortunately, they also tend to be aware that this is not a good idea:

*"I think they're fairly specific about the fact that it's personal certificate and shouldn't be shared, so unless you've explicitly got it on the basis of it being a group certificate I wouldn't really want to share it."*

*"I would not share my password."*

Whilst stakeholders are typically aware that subverting authentication is undesirable, they usually do it because they are faced with a lack of time or access to administrative resources that would allow them to complete their tasks. Their decisions are based on the weighing the perceived risk of bypassing authentication against the immediate need to get on with their work. Unfortunately, their decisions may be based upon inadequate understanding or inaccurate perceptions of those risks. Subverting security tends to be less a problem for users of Web Services because stakeholders are

generally more familiar with username and password authentication schemes. Digital Certificates are more prone to authentication subversion because of the usability issues surrounding their acquisition and use.

### 3.4.3 Passwords

Simple knowledge-based authentication is familiar to all stakeholders, but suffers from usability problems that have been documented in other domains: the computer monitor covered with sticky notes with passwords is a cliché that has entered the popular discourse. E-Science applications add further to an existing problem, and as elsewhere, users adopt approaches that emphasize memorability versus security, even though they are aware of the importance of secure passwords. Administrators, who are the point of contact for assistance when a user has forgotten a password, receive many requests due to forgotten passwords. These comments are a representative sample from researchers and administrators:

*“It has to have some relation to something I can remember. Random sequences of letters and numbers, I don’t think I could remember.”*

*“[I emphasise] memorability, I’m afraid.”*

*“Memorability. Because I have many passwords.”*

*“But of course, a lot of people in computer science aren’t [using secure passwords]. I wouldn’t expect the average biologist to.”*

*“They next biggest category [for support] is the typical kind of thing, ‘Okay, I’ve got my account, I can’t login, I’ve forgotten my password.’”*

As a workaround to this problem, many users adopt a reliable method for remembering something. They write it down:

*“The problem comes if there’s something that needs to be secure but you don’t use it very frequently and essentially, I end up writing it down.”*

*“So, if you insist on a password that is complex and changes frequently, people simply write it down on a Post-it and stick it on their monitor, which kind of defeats the purpose.”*

Increasing the burden on memory, people tend to accumulate numerous username and password pairs for the abundance of accounts they must access. Not all e-Science systems need and use authentication mechanisms. However, those systems that do require them may contribute to the burden on memory. The result is that people may get their accounts and associated passwords out of sync. Several respondents identified this as being a problem. When this happens, users may attempt to login to accounts with the username and password pairs that they can remember. If the system subsequently lock-outs their account due to several failed attempts, ostensibly to prevent brute-force cracking, users will have to turn to their administrators to rectify the problem.

### 3.4.4 The “Price of Admission”

Some respondents felt resigned to the usability problems they have encountered with authentication, particularly with Digital Certificates. Both researchers and administrators expressed the attitude that such problems are the “price of admission” for using the e-Science:

*“It’s worth the pain, because the resources are good and you can get some science done on the computers. But getting access is not a fun experience.” [researcher]*

*“So, the price of the admission really is getting your hand on the UK e-Science certificate and learning to use some GRID technologies in your work, I mean, that is the price of admission.” [administrator]*

This kind of perception leads to negative feedback by word-of-mouth networks that it is difficult to use the grid. It is not difficult to imagine that bad feedback about experiences with an authentication scheme will lead to more diffuse negative impressions of e-Science in general, when

stakeholders relate those experiences are to friends and colleagues and these stories are told over many successive iterations.

### 3.4.5 Digital Certificates

#### 3.4.5.1 General findings

Whilst they offer promising potential for providing authentication and secured access, Digital Certificates present a profoundly challenging usability problem for security of e-Science systems. Of all of the issues identified in the course of this research, a large proportion of them stem from difficulties that stakeholders experience with obtaining, using, and managing Digital Certificates.

Before examining these in detail, it should be observed that when they work, Digital Certificates provide a reasonably secure level of authentication. They also offer the promise of the much discussed and pursued "single sign-on" capability, which has a positive effect on usability by reducing the burden on memory. Also, the emailed responses from support in the UK are generally reported as courteous and fast:

*"Sometimes they respond very quickly in the same day or something. If not, maybe the day later."*

*"They respond quickly."*

though this swiftness of response is not a factor in how quickly or effectively stakeholders' problems are resolved.

Once they have been successfully obtained and installed, participants appear to have relatively few problems with Digital Certificates - as long as they do not need to use a different system. It could be hoped that once researchers become more familiar with the technology that these problems will begin to subside. However, every one of the administrator respondents raised concerns that the process of granting certificates to researchers, as it is currently practiced, is not a scalable solution in the long term.

Our participants relayed a number of anecdotal "horror stories" from first-hand experiences. As the process of obtaining a Digital Certificate is procedural, respondents tended to encounter difficulties in similar areas over and over again. Recounting these experiences in a procedural manner lends itself to narratives. These stories vary in details and severity, depending on which part of the Certificate authentication mechanism failed. Highlights from the data include the stories of:

- **The Researcher** who found that documentation on how to obtain a certificate was so poor she had to write instructions for her staff.
- **The European Researcher** who waited for three weeks to receive his certificate from the Certification Authority because there was no CA in his home country.
- **The Researchers** who reinstalled the operating system on their laptops and forgot to export their certificates before doing so. They had to revoke their old certificates and apply for a new one, resulting in a time-consuming delay in their work.
- **The Foreign Certification Authority** that took two months to reply to an emailed request for assistance with obtaining a certificate: "She didn't answer my email and she sent me, two months later: 'Oh, sorry, I found an email in my mailbox.'"
- **The Certification Authority** that waits for batches of requests to arrive before processing them, at which point a member of staff physically walks the certification requests across the campus to complete the certification process.

These are typical of the kinds of stories that participants described during interviews. They point up a few of the key problems that surround acquisition and use of certificates. The following sections describe these and other factors in detail.

### 3.4.5.2 Conceptually Difficult

Respondents do not have a good understanding of what certificates are and why they are useful. Whilst it could be reasonably argued that it is better not to burden users with technical details of how PKI systems work, this has not been the problem. Users are lack of understanding even at much more general, conceptual level. This means they cannot leverage familiar conceptual knowledge to help them to use them, manage them, or fix them when they are "broken".

In particular, the two-part "public and private key" metaphor is unhelpful to all but a small segment of stakeholders. The typical stakeholder is familiar with common systems of authentication from the real world, such as a passport or identity card, where a single, discrete token authenticates them. It is difficult to understate that the notion of a two-part key system is not part of common understanding and therefore, is hard for users to grasp. As one administrator put it:

*"I think that conceptually, public key cryptography is a difficult thing for people to grasp conceptually. Obviously, mathematically it's complex, but conceptually it's not at all intuitively obvious or familiar to people that you can have a sort of secret that is only half-secret."*

Another said:

*"They basically learn a set of procedures they have to execute in order to get the stuff they want but they're not really understanding what they're doing."*

Responses from participants in the study repeat the findings from the classic study that first demonstrated the lack of usability of public key encryption<sup>4</sup>. This issue received the highest number of mentions in the comments on certificates. Some examples are:

*"I think that I needed a much more basic introduction - what a certificate is, how it works."*

*"I'm not sure how it works, exactly."*

*"Basically, I'm just following a recipe of how to apply for this and get a certificate and I'm not quite sure what to do.... It basically said, this certificate will get installed in Internet Explorer for whatever or you can save it out as this or spread it around and all this policy about what to do with the certificate."*

*"And then there was a second stage, which was actually applying for resources on the National GRID Service, but I don't think that was really security related but it was something that I didn't realise at the time."*

*Interviewer: "So do you have a different password for your Private Key? Do you understand what I mean when I say Private Key?"*

*Respondent: "No."*

The last example illustrates another observation: that administrators and security experts tend to use technical terminology with the belief that users will understand what they mean, but this is usually not the case. At the same time, many e-Science stakeholders, particularly those in academia, are not comfortable admitting their ignorance of the technical details, or betray it by asking "stupid questions". Many of our participants stated that this problem could be overcome by education and training, but previous efforts to teach users to use public key encryption<sup>5</sup> have resulted in a tutorial that takes 1.5 working days to complete, and even this is ineffective unless the tools are used several times a week. The key problem is that the underlying metaphor is "broken", i.e. it does not provide the mapping to existing conceptual structures that make

---

<sup>4</sup> [10] Alma Whitten & Doug D. Tygar, Why Johnny can't encrypt: A usability evaluation of PGP 5.0, 8th USENIX Security Composium. Washington 1999.

<sup>5</sup> Alma Whitten: Making security usable. Unpublished PhD thesis, Carnegie-Mellon University, 2004.

successful metaphors work<sup>6</sup>; the language of the metaphor needs to be re-engineered to allow users to build a working conceptual model.

#### 3.4.5.3 Poorly Documented

Confounding the problem of a complex security mechanism, the documentation which could support correct use of certificates is currently poor. Moreover, even basic instructions were found to be lacking. Of respondents commenting on Digital Certificates, a large number reported that documentation was nonexistent, difficult to find, or hard to understand. Typical remarks were:

*“The documentation is usually very poor.”*

*“It should be more standardised and better documented.”*

*“I seemed to remember when I got my certificate I got this little piece of paper, which was a printout of a document on the Certification Authority website, which was called something like, ‘Caring for your e-Science certificate’. I found it didn’t really explain anything to me in a way that I understood. I feel that the way that the Certification Authority website documentation is organised, it presupposes a level of knowledge of certificates, which most of us probably don’t have.”*

*“You have to go the same browser that you originally used. This is not made clear anywhere, but you have to do this.”*

*“...nobody told me in the beginning where the Certificates are supposed to be [stored].”*

*“Well, there were [instructions], but it was confusing. All the instructions – there were pieces of instruction in different places, on different websites to do different steps.”*

*“Well, I wrote - because I have students – I knew that after me, lots of people would be getting a Certificate. So the easiest way to do it so I don’t forget all the steps and who I had to talk to is to write a document.”*

*“But for instance, you have to know who is the RA that is going to authorize your certificate and their name and their phone number? These things are not online. You understand what I mean? The practical matters?”*

These responses indicate that people are not getting the information they are looking for. This appears to be the result of several problems. Most researchers are not supplied with or made aware of documentation that is hosted at their home institution or which may be available from bodies such as the Certificate Authorities. Compounding this problem, the way that they are introduced to Digital Certificates is not standardised in any way. For e-Science stakeholders, there is no obvious resource for information about Digital Certificates. The lack of effective documentation contributes the problems people have in forming conceptualisations about what Certificates are and how they work, as described above. It also increases the difficulty of obtaining and using them and in resolving problems when they fail to work as expected.

#### 3.4.5.4 Difficult to obtain

With only a couple of exceptions, all participants reported that the process of obtaining and setting up a certificate was difficult, confusing, and time-consuming. One respondent summed up his experience in three words: *“It was hell.”*

It is troubling to note that administrators, who tend to be very familiar with PKI and Digital Certificates, often discount or forget the difficulties that other users encounter in obtaining them. This is perhaps counterintuitive, as administrators are the persons that users are most likely to

---

<sup>6</sup> M. Angela Sasse: Usability and Trust in Information Systems. In: "Trust and Crime in Information Societies", edited by Robin Mansell and Brian S Collins, pp. 319-348. Edward Elgar 2005.

contact in the case of trouble. Yet paradoxically, responses from administrators such as the following were not uncommon:

*“The main UK eScience CA certificate takes about a day.”*

*“[Regarding authentication] Everything seems to have gone fine.”*

*“The actual turn-around for the Certificate will typically take one to two days.”*

It is difficult to account for why some administrators fail to realise that many researchers encounter problems when obtaining a certificate. Although it is true that the Certificate Authority may only require one or two days to issue a Certificate after receipt of a request, from the researcher's point of view, this process typically takes much longer. This is because the process consists of several steps where a combination of user ignorance and poor feedback result in a lack of knowledge about what is going on.

For the researcher, obtaining a Digital Certificate involves learning about the relevant concepts or – at a minimum – the steps they must perform, using a web browser to request the certificate, returning some days later to get the certificate, exporting it (in most cases) to the correct place in their local file system, and sometimes configuring their e-Science application to use the exported keys. A failure is possible at any part of this procedure and feedback about the cause of the failure is usually poor or nonexistent. Though there are a few exceptions, most respondents tend to regard this process as one that is very challenging:

*“It was a very painful process and it went horribly wrong for some reason, nothing to do with me, and is still mysterious to me.”*

*“The guy who was supposed to actually implement the certificate, or set it in process, took a long time to do it, so it took me some months to get it done.”*

*“...depending on what their authorization procedure is it might be really straightforward, it might be really complex, it might do it for you instantly, or it might take days. It might take weeks.”*

*“Probably about three weeks, four weeks.... It did seem a little bizarre that it took me three weeks to get a key or whatever but I think probably my experience was slightly atypical.”*

*“But the second one, the one that I've just had done, came through faster – although the [Certification Authority's] system crashed and I had to reapply.”*

*“One of the guys from [an industrial partner] couldn't get a Digital Certificate by using the systems at [his company]. I can't remember whether it was a firewall problem or what it was. He had to actually request a certificate from his home Internet account.”*

*“If you can manage to do it in an hour, that's okay. If you have two pages to read and if you follow the steps and you can get it done in half an hour to an hour, that's okay. If you have 50 pages to read and it takes 2 weeks it's over the top. That's too much.”*

Obtaining a certificate has been further complicated by the fact that to date, only one particular browser could be used to complete the transaction. Also, the same browser also had to be used for both requesting the key and obtaining it, upon issue, because only the requesting browser retains the proper mate to receive the authenticated Certificate. However, to users who were unaware of this requirement, due to poor or no documentation, this made certificate acquisition more difficult, often forcing users to restart the process with a browser that they would be sure to have access to for both parts of the procedure. Many respondents experienced difficulty:

*“I think that what I hadn't appreciated was exactly how important the order of doing things was and, as it were, using the right browser.”*

*“The browser dependency is a big hang up.”*

*“And then we had to switch between one browser to another to download the Certificate and the Explorer from Windows would work, but if you are in Linux using Netscape – I don’t remember exactly, but some of the users – like Mozilla wouldn’t work. So you had to apply for a certificate – it was really, really not very easy because of these little details.”*

*“There’ve been browser restrictions in the past, which have made it even harder.”*

*“And you have to use exactly the browser they told you.”*

*“It was something to do with the browser and the fact that [our industrial partners] have a very fixed software setup.”*

The browser dependency problem should have been corrected by the time this report is issued. However, this is the type of problem that could have been mitigated or averted entirely by with proper documentation and guidance. It is also indicative of the developmental nature and immaturity of e-Science technology. Presumably, the stakeholders who chose a single browser did so to overcome limitations of the numerous web clients, which follow few consistent standards for mark-up, to say nothing of key encryption. At the early stages, it might not have been easy to foresee that such simple decisions would present very perplexing complications for the user base. However, the experiences of these users can serve as an object lesson for developers of future e-Science applications and services. Quick workarounds have a tendency to become ossified into inflexible and burdensome problems and demo applications are commonly adopted as a production solution, rather than remaining merely a proof of concept.

Unfortunately for many respondents, getting their certificate from the CA was only half of the battle. Having obtained their certificate, many users encountered further difficulties in actually getting them to work.

#### 3.4.5.5 Difficult to import and export

Several participants reported problems exporting their Certificates to GRID applications. For most people, a web browser has a primary function: web browsing. Respondents often failed to understand that their web browser was merely being used as a transport mechanism for their key, because this is not a task people commonly perform with web browsers. It is possible to hammer a nail with a screwdriver, but it is not the model for the use of a screwdriver that people are used to.

This is exacerbated by the fact that among browsers there is, of course, no common directory for export of keys and no standardisation among e-Science applications about where the applications expect to find keys. The result was that many respondents had problems using the keys, once they were installed. Combined with the frustrations that many had experienced in obtaining one in the first place, this only added insult to injury:

*“This is not easy for a novice and it’s not easy for even a hardened user.”*

*“Exporting is a problem.”*

*“Probably, if you want to use it on the GRID, you then need to export it from the browser into some other application, at which point all kinds of fun and games ensues as people try and do that.”*

*“Well, it’s not a pain, it’s just it doesn’t flow, if you know what I mean. When I say it’s a problem, it’s not a problem. It’s just tedious I suppose.”*

*Respondent: “Just having one bit doesn’t work, so you need to have both of them. So you do that and then you put the two bits in, of course, different places.”*

*Interviewer: “They can’t be in the same directory?”*

*Respondent: “They could be in the same directory, but most of the GRID software that I’ve come across expects them to be in different places.”*

### 3.4.5.7 Break easily and are hard to fix

Digital Certificates are brittle. Once a certificate has been installed on a given system and configured for a GRID application, changes to the system or the e-Science application frequently lead to authentication failure. Magnifying this problem, e-Science applications and the security services offered by the GLOBUS Toolkits and Web Services give poor feedback to users about what has caused authentication to fail. Messages are terse and technical. They do not generally provide help that will assist the researcher in resolving an authentication failure. This excerpt serves to illustrate the point:

*Interviewer: "So it was looking on your computer for a Certificate and it was just in a different directory?"*

*Respondent: "Yeah."*

*Interviewer: "And so, for that reason, it wouldn't work?"*

*Respondent: "Yeah."*

*Interviewer: "And it didn't tell you that it couldn't find it."*

*Respondent: "No. No, not really."*

*Interviewer: "Not in any meaningful way?"*

*Respondent: "Yeah, that's the idea. So, I'm a user. I know Linux. I can install things, but I'm not an expert on Certificates and things like that."*

Stakeholder involved with good technical knowledge may also experience difficulties. Developers admitted encountering problems getting authentication mechanisms to operate as expected. Many self-described experts interviewed in this study admitted that they often have to turn to outside help when they encounter failures. Some typical responses from more experienced participants were:

*"Well, yeah. Nobody knew very well what was happening, why I couldn't access the GSI-SSH and finally one of the guys who knows about Certificates and stuff found out that it was simply because it was looking for the certificate where it wasn't."*

*"You don't get information that it looks in the wrong place, it just tells you where it looked for it and if you don't know where the Certificates are – it's just a little line that says 'Looking for Certificates in /etc/certificates or /.globus/certificates or something like that.' I saw it by chance."*

When certificates fail to work as expected, some users have learned enough to make efforts to repair the problems on their own. However, if the solution is not immediately obvious, they must turn to experts that they may or may not know. This causes people to lose faith that their applications will work as expected:

*"We have demos and the demos work very well, at home. But only because we were here, and not at home, one of the guys who's really an IT guy, spent three hours to make the demo work. I don't know the reason, but you always have problems."*

*"Well if it is something that I use everyday and it is in my computer I don't spend any time. But if something changes – like you change the computer you're working on – something changes, anything, then it's a problem, always."*

*"Yeah, as long as the access works, it's easy. But if there is a problem, it's very hard to fix. You need somebody who supports you. Usually, I'm quite experienced with using all sorts of computers, but fixing a Certificate problem is not straightforward. So usually I have to ask somebody."*

When certificates work well, they are seamless and transparent. This is generally construed to be a positive outcome. However, difficulties emerge when that very transparency causes users to forget the technology they are relying upon. Three different respondents described having to revoke and request a new certificate because they had lost one, due to rebuilding a system. Describing this scenario, one user was so exasperated with the whole process that he finally gave up:

*“Well, since I got my Certificate, I had reinstalled my laptop and my Private Key was in the browser. I had completely forgotten about it. And so I couldn’t renew my Certificate. And guess what? I gave up. I said, ‘I’m not going to be bothered with this.’”*

### 3.4.6 Specific Recommendations

It is likely that many GRID systems will continue to use Digital Certificates as an authentication mechanism for the foreseeable future. Given that this is the case, the following specific recommendations are offered to mitigate the problems identified by this research:

- **Specifications should be developed which would help Project Planners and Developers to identify exactly under what circumstances a GRID project should consider using Digital Certificates as part of the authorisation scheme.**
- **Documentation and training on the acquisition and use of Digital Certificates should be enhanced. All new e-Science stakeholder should receive some form of coherent, consistent introduction, delivered in a non-technical, jargon-free form, which would introduce them to the important concepts and advise them how to seek help. In simple terms, it would describe what a Digital Certificate is, why they are needed, and how to obtain one. This probably requires re-engineering the standard metaphor and language attached to Digital Certificates.**
- **A unified support strategy for dealing with Digital Certificate problems should exist. The current support methods are disparate, and subject to finger-pointing. Researchers are confused as to whom they should address their issues. A single point of contact would be helpful, someone who would be responsible for resolving users’ problems. Ideally this same person would be an advocate who would take partial ownership of their issues and help them to determine an effective solution, mediating with the necessary parties, if needed. Alternatively, the boundaries of responsibility for support, whether they be those of an Administrator at the local institution or a party at the NGS level, need to be identified and communicated.**
- **The “price of admission” attitude is a threat to the uptake of e-Science. Awareness, education, and training should be designed to dissipate and prevent this attitude, particularly for administrators.**

## 3.5 Frameworks and Toolkits

### 3.5.1 Web Services

Web Services (WS) have been available to developers for over a decade and as such, have had time to mature further than newer e-Science technologies. Developers involved in using Web Services have come to rely on communities of practice that are actively engaged in using and enhancing these tools. Web Services are supported by major commercial players such as Microsoft, IBM, and Sun, and have established communities of Open Source developers as well. This tends to enhance the robustness of the tools and the flexibility with which they can be applied to a variety of security needs.

Web Services have been used successfully to provide basic security, such as authentication, encryption, and transaction processing for some time. Indeed, the commercial sector has used web services to conduct its business for many years. Thus, this resource has been taken on board by many e-Science projects, particularly those that have developers who are already experienced with using some of the available tools. Many e-Science projects boast of successfully using Web Services solutions to meet their GRID security needs. One respondent observed:

*“At the moment, the only standard solution that we have is just plain Web Services over a secure socket because most languages support that, and whilst computer scientists kind of pooh-pooh that whole way of doing things, it actually works, and it works with lots of languages in a fairly interoperable way.”*

However, whilst their utility has been widely recognised, Web Services are only just beginning to develop mechanisms that are useful to the needs of the e-Science community. Like other e-Science technologies, some of these tools are not yet mature and many have yet to be tested in production environments. For some projects that are facing extremely sensitive privacy issues, web solutions are not considered a mature, secure option:

*: “...you can’t [wouldn’t be able to] put up an insecure web service serving up patient data. You’d never get to the point of being able to turn it on.”*

Though it is not true for all e-Science projects, the majority of e-Science systems that rely on Web Services security only use basic services to address such needs as authentication or message-level encryption, rather than the more complex and challenging ones posed by job management of parallel processing, distributed storage, and authorisation. They also tend not to be specifically designed to address security issues such as provenance, auditing, and delegation.

In addition, Web Services suffer somewhat from a lack of effective documentation, though there is a great deal of it available. However, such documentation appears to exist in quantity, rather than quality. It also tends to be aimed at Developers who implement the services, rather than researchers. Although this is not a problem when Web Services stay “behind the scenes”, it presents challenges for researchers when the services break, when they must be configured by End-users, and when they must be integrated into environments running many other services.

Developers do face problems finding useful information. Documentation for Web Services, particularly those which originate from the Open Source community, is often widely scattered, difficult to find, and of relatively poor quality, leaving developers to scour the web for answers or to turn to their colleagues for help:

*“So, I know there are one or two people in the university who have figured out how to, say, use WS security and get it up and running and so I know who those people are and so they’re probably the people I would go and talk to, to find out more.”*

One Developer suggested that:

*“Examples [of how to implement security using Web Services] would be great.”*

Web Services also offer patchy coverage of the security issues that developers would like to solve. In particular, support for a wide range of platforms is found to be lacking:

*“Although Python has Web Services support, it doesn’t have a very easy to use WS security tool, which is one of the problems we have.”*

*“For example, you can get a decent WS security toolkit for Java and .Net but if you’re using any other languages, it’s a real pain.”*

The balkanized nature of Web Services development also means that middleware Developers must face challenges authoring tools that will be flexible enough to serve the needs of both Web Services and the GLOBUS Toolkits. One respondent said:

*“...it becomes very difficult to write software that is generic enough to provide that security over all the different platforms...Because they’re fundamentally different in how they [Web Services and the Globus Toolkits] do things.”*

Perhaps one of the more significant problems surrounding the use of Web Services for development of security in e-Science applications is that they allow for a “bolt-on” approach to security and usability. Generally, they permit developers to author applications and services which will run entirely without any security mechanism, such as authentication. This makes it possible to treat security as an accessory added almost as an afterthought, after other development stages have been completed. This can contribute to a perception that security is not part of the core functionality of a project. The following observation nicely summarizes this problem:

*“One of the things that I have been told about by developers is: one of the things a lot of them like about the web service paradigm is that it explicitly – and incorrectly, I might add – but it explicitly separates out security from the rest of the transactional framework. And so you can implement your web service just fine and then you worry about the security later, basically. And you try and put it in or not, as you might. Now obviously, any security person will tell you that’s just doomed and going to be absolutely disastrous from a security standpoint. From a developer standpoint it’s great. Because they don’t have to worry about this complicated security thing.”*

Another respondent warned that:

*“Just being able to get a message from A to B and some sort of security involved in that doesn’t tell you much at all about how the larger patterns of [systems] interaction should be done. So it’s really only a small piece.”*

The relative ease with which Web Services and Web Services Security can be implemented and the successful experiences of many developers makes it likely that e-Science projects will continue to view these technologies as a viable option for development of GRID applications and services, and their attendant security implementations. Indeed, one major UK University has determined that it will use Web Services tools exclusively in the development of all of their e-Science technologies, at least for the foreseeable future.

### 3.5.2 GLOBUS Toolkits

The GLOBUS Toolkits (GTK) have been available for several years and are just beginning to reach maturity with the introduction of GTK 4.0. They have offered Developers alternatives for authoring robust e-Science applications and services, including some tools for adding various security features. However, for most respondents, the number of GTK components and their associated learning curve, combined with a lack of adequate documentation has posed serious usability problems.

It must be noted that the respondents surveyed have based their opinions on earlier versions of the GTK. Previous versions were widely believed to be buggy. However, many of these problems have been advertised by the GLOBUS Consortium as being eliminated in GTK4. One middleware developer was very optimistic about the anticipated release GT4, preferring it over OMII. However, even if the problems identified by respondents as plaguing the previous releases of the GTK will have been eliminated in GTK4, these problems have created a bad reputation, predisposing many in the e-Science community to avoid GLOBUS entirely. One research group has already decided not to use GLOBUS components in any way. The responses of participants will serve to shed light on these issues.

Beginning with documentation, several participants identified shortcomings and felt that they were not able to locate enough information to use the GTK. Respondents described difficulties locating documentation and understanding what they found. They also encountered deficiencies in the available documentation. Some representative responses were:

*“I’d say I wouldn’t even know where to start looking.”*

*“Yeah, the documentation of GT3 security took a long time to understand.”*

*“I suspect had I not had that meeting with that woman, I would not have known – I wouldn’t have read the documentation enough to find there had been that change until someone tried to run it and it all fell over”*

*“It just didn’t seem well documented. We were guessing at things”*

One participant characterized the necessity of building personal networks in order to be able to compensate for a perceived lack of useful documentation:

*Interviewer: “So then, is it fair to say that you have to basically build your own personal contact networks for people who know GLOBUS so that you can use it?”*

*Respondent: “Yes.”*

Even when sufficient information was available, many respondents reported problems using GTK to meet their needs, beginning with basic installation of the toolkit components:

*“Not all of it, but a lot of it is very difficult to install and maintain, the GLOBUS Toolkit being the prime culprit in this arena, being one of the most difficult pieces of software I’ve ever had the misfortune of trying to install.”*

*“And I think the thing that struck me was that this was a 2-day class just to learn how to install GLOBUS and they kind of casually said in an offhand way the way the network is set up, if you get your configuration wrong, you can take down the entire country. We did.”*

*“But even then, it took me writing quite a few emails to various people to even work out that there wasn’t a Windows installation of – GSI-SSH is the GSI client for logging on. The information that there wasn’t a Windows installation was quite hard to come across.”*

*“But I can’t point now to a nice web page where it’s all [installation] described on.”*

The perceived lack of information contributed to negative impressions of the GTK, particularly when Developers encountered the shortcomings of the early releases. Immature software is often dogged by instabilities and bugs. As most of the developers that have used GLOBUS have worked with earlier releases, these negative impressions have proliferated. Many respondents reported that the GTK is plagued by poor implementations, instabilities, and platform-dependencies which contributes to the impression that the toolkits are brittle. Some representative remarks were:

*“There didn’t seem to be a toolkit release where everything actually worked consistently.”*

*“The thing we hadn’t thought about – we didn’t switch on the security and when we switched the security on, GLOBUS Java library for GRID FTP – the security part of that was broken. And it took me a long time to get that one fixed.”*

*“And it seems there are three possible problems and each problem is assigned to a different person and the persons are attributing the problem to each other. So it’s a vicious circle.”*

*“I couldn’t get the basic testing of the installations to work properly.”*

*“Anyway, when they changed from one version of OpenSSL to another that changed the way in which Distinguished Names are validated, basically. And so that meant that certain CAs, like the UK e-Science CA, their CA Certificates suddenly, overnight, ceased to be valid, basically, because of a change in the underlying OpenSSL toolkit.... So suddenly around about January, February, March 2004, there were people up and down the country whose GLOBUS installations suddenly stopped working overnight when they upgraded.”*

*“Well, it had previously worked and I installed it in a different way and I couldn’t get it to work and I had no idea you needed to pick up that file and it was looking everywhere else for it.”*

*“And to commit to SRB using GSI technology, I needed to do an installation on Linux. There was no client available for Windows at the time. So, basically, we just didn’t use it because all of our users are Windows-based.”*

The last example highlights the particular problem of platform dependency. Security solutions that are tied to a specific operating system or platform present real problems for developers and researchers who cannot or will not change platform just to get the security tools. For most respondents, who may face institutional barriers to platform change, keeping the existing, familiar systems is more important than increased security.

Adding to the difficulties that developers have using the GLOBUS tools, the evolutionary nature of the toolkits has meant that the releases are not always compatible. In many cases this has meant that respondents had to allocate substantial project time to rewriting software to be able to accommodate for changes between GTK releases.

*“If you did it in GT3, you will have to re-author for GT4. If you do it for GT4, when GT4.2 comes along, you might get away with a rebuild, you might not. I don’t know.”*

*“We are re-writing our stuff for – we’ll to a re-write for GT4 anyway.”*

*“What we’ve actually had to do is remove all the GT3 code, build a layer above it and then put GT3, GT4 and OMII in underneath.”*

*“One of the problems with GLOBUS is that every different version is completely different and completely incompatible with the previous one and every version is absolutely monstrous in terms of scope and usability, so these are the big problems with GLOBUS.”*

*“Yes, how you do delegation in GT4 is quite different from how you did delegation in GT3 and it’s quite different from OMII has proposed security.”*

Also, some developers feel that the GTK will not be able to meet their security needs. One respondent said:

*“What I discovered is that there are gaps in the security model used by these GRID type technologies that would concern me if we wanted to roll the system out wider than just the systems for which we’re directly responsible.”*

These experiences have led developers to be wary of GLOBUS and to anticipate that the new release of the GTK 4 will not offer the promised solutions, particularly regarding e-Science security:

*Interviewer: “So you suspect that the new GLOBUS Toolkit will have similar kind of problems?”*

*Respondent: “I wouldn’t be surprised. I know they’ve done a security review and it’s getting better and better. But if somebody suddenly came out tomorrow and said there’s a big security loophole in GT4 security, it would not be a shocker.”*

*“It makes us slightly wary of it [GTK4], but also the way that every version of GLOBUS is completely different and completely incompatible with the previous.”*

*“It’s certainly no rumour that GT4 is about to be released. There are rumours that it’s actually good.”*

It remains to be seen whether the new release of GTK 4.0 will begin to address some of the usability problems that the respondents identified. Whether or not this is true will depend on the robustness of the new software and the individual experiences that Developers have trying to use it. Positive experiences will change attitudes. For example, GRID-FTP as a single component, is generally regarded as being one of the most successful and reliable applications for data transfer. If GTK 4.0 and subsequent releases offer the same quality in terms of usability and reliability, it is likely that the general opinion of GLOBUS among Developers will be improved.

Another way to improve developer acceptance is to support learning and provide guidance for the both the new GTK and Web Services. Developers identified a particular need for this kind of support during the research:

*“GLOBUS has had various security models and they’ve changed a little bit. But one’s not going to be able to interoperate them just by reading the documentation.”*

*“I don’t think you can tell people “Thou shalt have security” without telling them a little bit about what kind of security, without giving them a little bit of advice to go along with it.”*

*“WSI basic security profile helps, but when you – how do you use it? How do you know what bit a service is using? It does not come for free. It requires effort. And without that effort, you can get all these services which everyone’s confined to a very small fraction of that space until their tooled up again.”*

*“The most critical thing to the GRID community in the UK is that they [Developers] want strong guidance on what security models to adopt within the applications and services which they themselves build and seek to deploy amongst themselves and their friends.”*

### 3.5.3 Specific Recommendations

Most e-Science projects have identified security or implementation needs that will require the use of Web Services or the GTK. Regarding these, the following recommendations are offered to improve usability:

- **A single documentation resource should be developed (e.g., a website) regarding implementation of e-Science security using WS or GTK. It would provide case studies and examples of successful implementations. Sources of further documentation from the community-at-large would be listed.**
- **Specifications should be developed which would help managers and developers to identify exactly under what circumstances a e-Science project should consider using either Web Services or GTK as part of the security scheme. These people need a flowchart that will help them identify exactly which technologies to use and when.**
- **Training and development programmes should be continued and enhanced so that Developers will be adequately prepared to meet their security implementation requirements.**
- **Relationships with key industrial partners and consortia should be cultivated. This would provide opportunities for these groups to interact with e-Science security Development teams. This is one of the key ways that developers learn to solve their problems.**

## 3.6 Authorisation

### 3.6.1 General Findings

Although there are a few research communities who are not worried about others seeing their data or results, most consider their code and processed data (but not the raw data) to be intellectual property, which is important their research or academic careers. They are therefore very keen that others will not be able to view their data unless it occurs under their careful control. Across the board, respondents said that only owners, and usually administrators, should be able to amend or delete data:

*“Deleting I think is something that I don’t think anyone other than the owner should be able to do. Amends should only be allowed in a change-tracking type of environment...”*

*“I wouldn’t like anyone to be able to delete my data, I don’t think. Certainly not delete bits of code, no. I would be extremely worried. No, I can’t imagine any circumstances in which I’d want to let anyone delete a bit of data.”*

*“Now if any random person in the collaboration – let alone outside – had access to that data and could delete it or amend, that would be catastrophic for us completely.”*

*“You should be able to say, this data is only readable, writeable etc. by the owner or by a selective group of other users.”*

*“You almost certainly care, unless you’re a complete muppet, you almost certainly care if the data’s been tampered with.”*

*“No, I wouldn’t like anybody to modify my data. But I would like people like the students and other colleagues to be able to access my data.”*

Though they recognise the need, researchers also understand that delegation, depending on their particular security requirements, may be hard to control:

*“How we take subsets and abstracts of that data and share it in a legally acceptable way is quite an interesting problem.”*

*“There is no good way of handling authorization. The authorization framework that most GRID systems offer is too coarse-grained. So it basically, tends to come down to, ‘Yes you can use my resource, or No, you can’t.’”*

For administrators, who may need to manage and audit this authorisation beyond a rudimentary level, this is also a non-trivial issue. Administrator feedback indicates that current software to support this task and others, has not reached the level of maturity that can effectively support larger groups of researchers, and entities such as virtual organisations:

*“While there’s no problem setting up a secure channel [of authorisation] between two organisations...when you set up 10 organisations, you can’t get all the IT managers to sit down in a meeting and agree a standard protocol...”*

When they are deployed, many e-Science applications require administrative support in the form of managing authorization of users and resources, and backup. For some of the prototypes, it has been unclear how this administration will be provided and where this support will come from.

Developers, faced with a need to provide such capabilities have encountered challenges, as well. Although many recognize that resources such as the GLOBUS toolkit(s) and Web Services may be able to assist with authoring software that permits this type of authorisation and policy management capability, it has often been the case that development of security has centred mainly on other areas such as authentication.

### 3.6.2 Delegation

Many researchers have found the tools at their disposal to be effective for sharing resources with colleagues. Providing a capability for delegating authentication can be very useful for sharing results with colleagues and stakeholders or for increasing the efficiency of the work. Respondents said:

*“And you want to allow them [colleagues] to go over that data and do things to it all the time.”*

*“The [users] will, in a controlled way, want to make their databases sharable. It will become a requirement increasingly.”*

*“...the context where people want to share data that often tends to pop up is that you’ve often got someone doing a PhD, who may want to be able to share their data with their supervisor. So they only actually want to share data with a small number of people.”*

*“The benefit is that it means that you don’t have to spend the time doing it yourself...”*

*“It would be sort of nice to add [a colleague] to access the NGS as well to see what is going on directly.”*

They also are aware that there can be risks to sharing access:

*“...if they were to run your jobs in any way that was not correct, data would be created by you that would not be correct.”*

Administrators are more acutely aware of such risks. They tend to have very different requirements and to consider delegation of authentication to be very risky:

*“There are no benefits.*

*“I would be very loath for other GRID users to be able to stop my jobs because you always get the scenario of who has the knowledge to say that their job is more important than your job. No, I think administrative stopping and holding of jobs, yeah fine. Non-administrative, I think, is not a good idea.”*

This perception on the part of administrators may extend to a biased belief that users who need delegation capability have simply not planned their work very well:

*“Yeah. As far as users go, I can’t imagine a scenario where it’s so life-and-death that you would need to get someone else to do something. That would be indicative of bad planning from a user point of view.”*

Fortunately, most developers for e-Science projects tend to be aware of user requirements regarding authorisation. For e-Science projects that provide workflows as part of the user interface, this has been an important part of the functionality. Indeed, for some projects, providing authorization whilst maintaining other security requirements, such as privacy, is a main goal of the research. However, developers may fail to take into account that in addition to researchers, administrators have usability needs it comes to managing and controlling authorised access.

### 3.6.3 Specific Recommendations

The findings from the research indicate that:

- **Each GRID project is unique. Thus, across-the-board recommendations would not suit every project. Rather, Developers should continue to engage with all Stakeholders in a project throughout the development process so that their solutions will address usability requirements for Authorization. AEGIS<sup>7</sup> offers a developer-centred method for such a process.**
- **Developers should consider security requirements and workload of system administrators and to gather their requirements for managing authorisation mechanisms.**
- **Projects with complex authorization requirements involving privacy and personal data such as health records should allocate a substantial amount of time and resources to overcome the special usability problems these security requirements present.**

## 3.7 Confidentiality and Privacy

### 3.7.1 General Findings

Confidentiality and privacy of data is key in stakeholders’ experiences of usability of security. If researchers do not believe that sensitive data is secure, they may be unwilling or unable to use e-Science technologies. Among respondents, the need for confidentiality varies with the type of data and the domain of research. Information that is available in the public domain from major databases in fields such as genetics, astronomy, and geology are not considered particularly sensitive. At the more sensitive end of the ‘confidentiality continuum’ are data patient records

---

<sup>7</sup> Reference in footnote 3.

which have not yet been anonymised, and meta-data about who has run what queries in highly competitive research domains such as bioinformatics. An additional complication for e-Science systems is that the access of one or two records alone that do not contain sensitive data could, when aggregated, yield information that could compromise privacy and potentially violate legal restrictions. Respondents clearly identified these problems:

*“There’ll come a point where somebody asks a query that is completely innocent, but we know that if they were happening to combine that with 36 other queries that ten different groups have asked cumulatively in the last two years, they would then be able to identify one person.”*

*“The number of people who get a brain scan in the UK as a whole – but even if you don’t take the UK as a whole, in our institution, this particular institution – is so small that even if you anonymise the data, you’ve offered a very, very limited – it’s not like it could be any one of thousands or hundreds of thousand of people. No. It could be any one of ten people...that’s just not good enough.”*

*“Yes, who accessed what, and when, in terms of confidential data, is itself, confidential data.”*

These are e-Science projects strong security needs. But even for respondents who do not worry about violating legal restrictions, there can be serious consequences for breach of confidentiality, particularly if researchers’ careers depend on first publication of important findings. Several participants identified the importance of keeping such information private or confidential among trusted peers:

*“Because there’s a great deal of competition in the publishing of results, it’s a very active competitive field. So it’s important that we maintain that level of control over access.”*

*“But if other researchers can see that I am looking at a particular piece of a particular genome, they can guess what I’m working on. And they’ll go, “Hmm. He’s working on that. That probably means there’s something interesting there, so I’m going to start looking as well. And then they steal my whatever...”*

*“It’s the results. The results the only things that they particularly want to protect.”*

*“There’s quite a rush to produce papers and be first authors on papers. So I would think that even though the raw data may be public, if people have done very special processing to their data and proven that specific spectral lines are there and certain other spectral lines are not, that they would want to keep close wraps on that until they can get their paper published.”*

*“At one end you have raw data needs to be very tightly controlled by only a few people. And as you process it, it becomes freer and freer.”*

Some research communities are so small that a publication based on stolen data would be obvious. Many of these respondents work with data in such a way that its meaning would be much obscured to anyone but the researchers themselves. Thus, compromise of anything but unpublished results is not seen as problematic. These respondents tend to be less concerned with whether their data privacy will be compromised because it would be obvious that results were stolen:

*“...it would be hard to imagine that you would be in a situation whereby somebody would have access to your data and use it in a way that didn’t actually get noticed by other people in the community.”*

*“It’s not a very private issue, I suppose. I wouldn’t want people stealing my results and publishing them before I had a chance to ... But I think it would be difficult for anyone to work out exactly what I was doing just from raw data I suppose. I suppose it’s a remote possibility.”*

*“If some rival, researcher X in institution Y stole my output, I don’t see how they could use them anyway because they don’t have the model and it would take them ages. You could try and make up a story of the results. You just wouldn’t do it.”*

*“...if someone did publish something, they would be very easily able to spot that results had been nicked from them.”*

One of the most troubling problems for confidentiality of e-Science systems is that while secure mechanisms can be devised, it still remains relatively easy for users to knowingly or unknowingly move secured data to an unsecured location or medium. Familiar tools such as cut-and-paste and clear-text communication channels such as email pose profound challenges and are deeply integrated into the work processes of most computer users. Researchers therefore tend to be unaware of the dangers that they can pose. Developers and administrators are less likely to be ignorant of these problems:

*“Yes, because they really do believe their data is secure. They don’t understand that if they put values into a web page interface, that is actually not secure. They just assume that no-one is going to be able to intercept that.”*

*Respondent: “They’re only likely to share it with other people in the same group they’re working.”*

*Interviewer: “How would they do that?”*

*Respondent: “At present, they share it by sending each other emails.”*

*“And I have users who have said to me, ‘Well we were interested in the idea of GRID computing until we realized that that would mean our data would be going to computers all over the place and we’d have no idea what those computers were and no control over them.’”*

*“...what we want is a system which is still secure and doesn’t just leak through people copying things onto CD and emailing things to each other.”*

A positive finding is that most e-Science projects which deal with privacy and confidentiality issues are keenly aware of them, and have allocated substantial resources to develop effective software and services. It is also encouraging that most e-Science developers and administrators interviewed for this research exhibited thorough knowledge of the significant challenges that confidentiality and privacy pose. However, they are more likely to see them as technical hurdles, rather than as usability problems which have consequences for whether users will accept and benefit from e-Science technologies.

### 3.7.2 Specific Recommendations

The findings from the research indicate that:

- **Developers must (continue to) engage with researchers to establish their security needs, particular risks what data is sensitive.**
- **Researchers must change working practices to avoid compromising sensitive data through the use of insecure clear-text channels such as cut-and-paste and email, and copying of data onto unsecured devices such as memory sticks. Awareness of the risks of these practices must be increased. Managers must take responsibility for changing researchers’ working practices.**

## 3.8 Provenance, Auditing, and Versioning

### 3.8.1 General Findings

Like confidentiality and privacy, the extent to which provenance and auditing of data are possible affects stakeholders’ experiences of security. Researchers must trust that their e-Science applications will successfully provide these features, or they may be unwilling to use them. Researchers need to substantiate their research conclusions with reproducible results, which is only possible if they can demonstrate the source of those results. Published results that are not reproducible are regarded with suspicion, and can potentially damage scientific careers. Thus, researchers need to be able to rely on e-Science applications that generate consistent results and which can account for how those results were obtained. A usable way to demonstrate the

provenance of data must therefore be a goal of most e-Science projects, particularly those relying on computational grid resources. Researchers were quick to identify this:

*“Yeah, with any of the scientific experiments, pieces of software we have on here, [provenance is] extremely important.... Papers are being published using this data, these are full scientific results that are being generated, so you cannot risk false data or false results. We’ve seen how that can damage scientific careers.”*

*“...if you conduct an analysis on a given complex GRID-enabled genetic database and carry out some data processing function, [it is important] that enough of an audit trail can be maintained of that process that some 3rd party could inspect it, if appropriate, and verify that it did indeed take place in a given way, and even potentially reproduce the results.”*

Corollary to provenance are auditing and versioning. Auditing provides a record of what has happened to data, when, and by whom changes may have been made. Developers require versioning capabilities to track changes, maintain discrete work units, and assist with troubleshooting. All of these needs were all identified by respondents:

*“Experience has taught me that an untracked change can wreak havoc.”*

*“People make mistakes and the fact that in a versioning system you’re able to revert to previous versions has been crucial.”*

*“For this part of the system...audit is the basis of the relationship.”*

Administrators rely on auditing capabilities to reconstruct what has gone wrong in the case of a system failure and to identify possible security breaches:

*“...they [Administrators] want to know who is using their resources and they want to know what’s going on. And in an environment like the University, it is actually a requirement, it’s a rule of being connected to the network that you can say who is using your machine at a particular time, so that if that person does something bad, we can find out who it was, at least potentially.”*

One particularly daunting difficulty, though not a common one, is that an audit trail can be used to identify individuals. This can be problematic in scenarios where protection of anonymity is important, as this respondent from a e-Science project involving NHS data describes:

*“We want the audit log to be centralised and protected in exactly the same way as the data is protected. And I don’t think the GRID is going anywhere near that space in terms of its architectural design.”*

The challenge for developers is to balance all of these needs in the applications that they author whilst providing a usable interface to these services. Historically, these requirements have gained relatively low priority in traditional software development. Only in exceptional cases have they been envisioned as part of the core functionality. The same has been true for most e-Science applications, the exceptions being in the medical records domain in projects such as CLEF, eDiamond, and CancerGRID. This means that audit features, which have been traditionally designed for technical users such as Administrators have not received the same attention to usability as other application features:

*“...almost no audit facilities have been ever built to be user-friendly to a non-technical expert, because audit logs are traditionally looked at by techies.”*

*“Unfortunately, for lots of GRID services...they do break, they do fall over. It can be a nightmare trying to work out – their auditing capability tends to be appalling...because the software doesn’t seem to be designed to handle that level of audit.”*

Substantial work must be done to investigate the usability requirements of auditing, even within the software development community outside of the e-Science Programme. The requirements for provenance, auditing and versioning tend to be very specific to the projects where they are needed. Thus, for projects that envision audit as an important feature, it is important to identify user needs early in the development lifecycle.

### 3.8.2 Specific Recommendations

The findings from the research indicate that:

- **Developers must (continue to) engage with stakeholders throughout the development process to ensure their solutions address security requirements for provenance, auditing, and versioning.**

## 3.9 Availability

Researcher concerns regarding availability of resources, whether they be computational or storage resources, is very domain-dependent. For some respondents, delays pose no problems whatsoever, whereas others are very sensitive to them. People fall between these two extremes, depending on the amount of time that they must wait to complete their objectives.

Certain research domains, wherein e-Science technology is used in a more exploratory and experimental fashion, are willing to endure reasonable outages of service. These respondents are content to access their stored work and process computational jobs, as time allows. Some of these participants said:

*“Analysis pretty much works on human time scales and jobs often take hours to run. Not having your data available for anything less than an hour is not disastrous.”*

*Interviewer: “What would be the effect of not having access to your data for a period of time, say, a second, a minute, an hour?”*

*Respondent: “Absolutely nothing for those periods of time. A week or two would be awkward.”*

Respondents from other domains may be less flexible about delays in availability if they have contingent processes, work under tight deadlines or have pressing obligations to publish results:

*“It would depend upon the length of time the system was unavailable, but if for a long period it was unavailable, you could end up with no access to that particular set of data. It would end up that the run may be compromised; the result of the run might be compromised.”*

*“Because, imagine, I am writing a paper and I have a story and I have to go back to my data to prove the hypothesis I have, for example. If I cannot access my data, I cannot write the paper and I am stopped. “*

*“...I have an organization and a schedule. So I organize my tasks so if something doesn't allow me to do the task I want to do for that day – that is not very nice. It's very stressful, right?”*

Others among the delay-averse group take advantage of e-Science computational resources because they allow for faster modelling or computation than processing locally. They may also use grid storage resources to gather large quantities of raw data in real time. If the duration of an outage exceeds the time they would be able to complete a job locally, or they are unable to store data during an experiment, this delay is deemed as unacceptable. Typical responses from participants were:

*“And sometimes I have no problem with waiting for a couple of hours to get into the queue. I know I put it there and – but sometimes it happens that a job hangs around for a day in the queue. Then I can do it on a desktop. Because a day is a day.”*

*Interviewer: “So if the system is down for three weeks, two weeks, a week, it could be, in fact, a really annoying problem if they were relying on it for daily research purposes?”*

*Respondent: "Oh, yeah."*

*"The vastly more critical thing for us is availability of the resources to store the data on in the first place. If they go down for minutes or hours or days then we begin to lose raw data and that becomes a real problem because you can't gather that data again and it's hard to do."*

Availability of resources has a direct effect on usability in that systems which experience outages frustrate researchers and undermine their confidence in e-Science technology. There were few results from the research that indicated that availability was a pressing problem. To date, sufficient infrastructural resources have been allocated to prevent availability problem. Developers have largely authored applications that provide a quality level of service over this infrastructure. This perhaps can be counted as one of the major successes of the UK e-Science Programme.

### **3.10 Security in the Development Lifecycle**

In the course of the research, several unanticipated themes emerged that have a direct or indirect impact on the usability of e-Science systems. Many of these were the result of experiences that Developers encounter in their work. Others result from aggregations of recurrent concerns by many individuals.

#### **3.10.1 Security Not a Functional Requirement**

One of the emergent themes was that developers are so pressed to produce working prototypes that they are not able to adequately address security needs. This was generally perceived to be a problem among e-Science projects that have not made security a part of the development process from the outset, whereas projects which deal intimately with issues surrounding privacy and confidentiality have been very successful in avoiding this pitfall. Several comments illustrate how security can come to be viewed as a non-functional requirement, in the face of other pressures:

*"Security was something that was talked about all the way through the project but a direction for investigation was never agreed, was never defined. The right potential solution was never - I don't think there was sufficient clarity - There wasn't an obvious answer, and so, in the absence of an obvious answer, attention went elsewhere and this whole issues of "How are we going to do it?" was put back, and back, and back, and back."*

*"For them [Developers], there are more important objectives."*

*"We came to October last year and we didn't have any security at all. And we had a lot of people who said we won't use this software unless there is some level of basic security"*

*"So, for the reason that it was a problem that had no obvious solution - there were other areas of the project which had more obvious solutions and people wanted to engage with. And that the project was under time pressures. And under-resourced, in the sense that certain posts didn't manage to get filled. So there were not enough people to tend to the job at hand, so security just got pushed back, and never really addressed."*

*"I didn't seem an issue. I think a lot of these things - to be honest, people still see security as a bit of an afterthought and we were probably in a similar boat."*

*"These are scientific projects with scientific objectives and, although you can't avoid the issues of security, it's not the primary objective for these people to become experts in computer security."*

Developers may also implement temporary security scaffolding to address requirement that they have identified, but which, due to time pressures are never properly authored for production. At least one respondent described this scenario. The risk is that the issues will never be properly addressed as project deadlines loom, and the need to demonstrate results becomes more acute.

#### **3.10.2 Security vs. Usability**

Several respondents expressed the concern that when developers implement security, they fail to consider the task and situational context in which the security mechanism has to be used –e.g. time

constraints. While the security needs and the challenges they pose for usability vary from project to project, these observations serve to illustrate the tension that exists between providing a functional, secure e-Science solution and making it usable:

*“...there’s a danger that the over-paranoid will produce a set of hurdles that nobody even wanted you to jump over. And you’ve created a cost, a burden, a methodological challenge, a slowing down of the process...”*

*“Security should be something that a user doesn’t have to think about.”*

*“Bad news travels faster. And any discussion of not being usable or well-designed is news that will travel fast.”*

*Interviewer: “Are you saying that usability is more important than security?”*

*Respondent: “Absolutely. Absolutely. Security can be managed by careful practice and good administration. Ease of use is something that’s either there or it’s not.”*

### 3.10.3 Stewardship

Some of the respondents expressed the concern that in the rush to get functional, usable prototypes into the hands of researchers, little consideration has been given to how and whether these projects will continue after the funding expires. This leaves a cloud of uncertainty over the final status of a project and can have a negative effect on researcher perceptions of usability and quality. Should they invest effort to learn to use an e-Science tool and integrate it into their work process if its long term status remains in question? Several respondents identified concerns in this area:

*“No, I don’t think - that’s another question which hadn’t been answered: Who would ultimately be administering the system? It wasn’t going to administer itself, so beyond the funding lifetime of the project when it was built, assuming this information about the users had to be kept somewhere, it had to be managed, it had to be backed up it had to be secured - and the system was not going to do that on its own - it was not clear where the administrative support was going to come from.”*

*“[GRID software]...needs to be kept up to date and maintained and whatever, and there aren’t really robust procedures for doing that yet. So maintaining the thing can be almost as installing it again, in some cases.”*

*“Having developed the services, who’s going to curate them? Who’s going to keep them up and running? Normally that’s the kind of – a lot of e-Science projects – that is out of their scope. I feel sorry for the people who expect that people who are doing these are developments are actually going to serve their needs. There’s this gap between research and production.”*

*“I’m struggling to understand what fruits of the UK e-Science programme translating into services which people within the UK can count on being there and fed and maintained five years from now.”*

## 4. Summary

### 4.1 General Recommendations

This report is the result of a survey of opinions gathered from a cross-section of the various users and user groups in the UK GRID community over a 90-day period of data collection and analysis. Whilst a single, brief survey cannot definitively identify all of the usability issues that are currently affecting the e-Science community, the responses of these participants have served to highlight some of the most salient and pressing ones. Deeper insight into the kinds of problems that e-Science stakeholders face in their work, whether they are researchers, developers, administrators, or managers, will require continued attention. Assessment of and feedback from the e-Science community can suggest solutions which will address their major usability issues.

The recommendations proposed are concrete actions that should yield positive outcomes for security on the GRID. Most of the recommendations were suggested by the survey participants, some have been inferred by authors. The authors would also like to point out that the many of the recommended actions provide the STF, or one of its agents, with an opportunity to provide leadership on security, and provide a one-stop shop for security issues, something that many of our participants crave.

The following general recommendations are summarised from the previous sections and ranked in order of importance.

- **Improve Documentation**

What is urgently needed is a unified documentation resource (e.g., a website) for all users. A single point of access for information, provided by an authoritative body, would reduce confusion about where to find information, whatever its quality. To do this, it is important to establish a first point of call, customised for each of the user types: Administrators, Developers, End-users. As more End-users come on board, this will become ever more important. Ideally, GRID projects should have a budget allocated to hire professional Technical Writers to create documentation when the project is over, part of which contains clearly written, jargon-free descriptions of how to use security. There is no sense in providing this resource without vigorously advertising it and allocating the resources to do so.

- **Promote judicious use of Digital Certificates**

Digital Certificates provide a reasonable level of authentication for very sensitive environments, but they are very difficult to obtain and use. As it is clear that Digital Certificates have become intertwined into the GRID community in a way that they will not be easily eliminated, clear guidance should be given to Developers as to when they should be used for a GRID project. Certain projects do not need Digital Certificates and should take advantage of alternative authentication options, such as those provided by Web Services. It would be useful to develop a specification that describes which specific requirements should result in the adoption of Digital Certificates and to ensure that all new GRID projects use it to evaluate what authentication scheme they will use. All End-users should be provided the same with consistent, clearly written instructions for acquiring and obtaining a Certificate. Improved support mechanisms should be provided to help them when they encounter problems.

- **Increase awareness, provide education and training**

*Security Awareness:*

Many researchers lack awareness of the value of assets (local and remote) and the underlying infrastructure, the threats they face, and why/how their own behaviour might put them at risk – e.g. that weak passwords or failure to virus-check emails can have far-reaching consequences.

*Security Education:*

Security education provides stakeholders with an understanding of their responsibilities for keeping their systems secure, of security policies, and what resources they have at their disposal. Education tells people what they should do, why it is important, and the consequences of failure. Different stakeholders have different security needs and responsibilities, and education must be targeted to address these.

*Security Training:*

Training provides the “how to” behave in a secure manner. For a researcher it might be how to create a secure password or obtain a Digital Certificate. A developer might need to know how to install the GLOBUS Toolkit. An administrator might have to grant or revoke authorisation to users or groups.

Awareness, education, and training each have a role to play in the provision of e-Science security. A matrix (see Table 1) representing these needs can serve to aid reasoning about the different information needs of each of these unique groups. The heterogeneity of e-Science stakeholders means that there are different user communities with different information needs. Their experiences are different, and so too are the specific actions that need to be taken to further the cause of security in e-Science.

- **Give Clear Guidance to Developers**

Developers are seeking clear guidance about which frameworks and toolkits will solve their security related problems. A planning tool that clearly describes which technologies should be considered and under what circumstances would be very helpful. It is also important to support developers' learning of Web Services and GLOBUS by providing a single point of contact for information, such as a UK Developers' Network for e-Science projects that is sponsored and funded. Networking and learning events should receive enhanced priority, as the personal contacts that are developed provide valuable resources when developers encounter problems. Communication must be enhanced through a website for information and community building so that people will become aware of these resources.

- **Continue Support for Outreach and Community Building**

Information and events that help e-Science stakeholder groups to communicate with one another are vital for the development and acceptance of complex new technology. Useful information is one of the key supports to overcoming security problems for all strata of e-Science users. Respondents indicated that the existing network events have been extremely helpful. But there is still room for improvement. Thus, it is vital to support and enhance information exchange through continued networking events, web resources, and advertising. Technologies such as web seminars and online communities may also enhance community building, and should be considered. These efforts will not be useful to the community unless they are thoroughly promoted.

- **Consider Security from Day 1**

For everyone, time is in short supply. This means that e-Science project teams should account for and allocate sufficient resources to develop applications and services that are both usable and secure. Security cannot be considered a "bolt-on" solution that is added after the core functionality is achieved. In most cases, this means that one or more competent individuals on a development team should be assigned the primary and sole responsibility for developing security that is effective and usable. While some e-Science projects have done this, it is very often the case that developers must wear many hats, leaving them insufficient time and resources to address security adequately. At all phases of the development lifecycle, the need to produce results should not take precedence over the development of usable security.

- **Plan for Stewardship**

While the ambitious goals of many e-Science projects have been achieved and concepts proved, little consideration has been given to how such successes will be nurtured in the long term. If users are not confident in the long term viability of their e-Science applications, confidence in the technology and its security, is undermined. In the course of developing secure, useful e-Science applications and services, managers should cultivate relationships that will support the adoption and care of their the projects for the long term, which will support user acceptance.

These general recommendations address the key problem themes identified in the study. The remainder of this report details the Method and Findings for particularly salient areas. When

pertinent, additional *Specific Recommendations* which would help to mitigate particular problems are suggested.

**Table 1: A Framework for Reasoning about Security Information to Provide to UK National GRID User Groups**

	<b>End-Users</b>	<b>Developers</b>	<b>Administrators</b>	<b>Others</b>
<b>Awareness</b>	<p>An End-user centred website which offers information of interest to new and seasoned GRID users. Contains details about emergent security risks (e.g., viruses, etc.) and directs users to information about how to secure their computers against threats.</p> <p>Introduction Packs/Pamphlets with plain-English descriptions of the GRID, Digital Certificates, and creating secure passwords.</p>	<p>A Developer-centred website which offers current information about authoring tools. There is a facility for announcements of important information such as upcoming networking events, notices of new version releases, bug reports, patches, etc. Capability for online community building and individual participation. Contains links to external sites and communities.</p>	<p>An Administrator-centred website which offers information about new security threats and patches. Contains links to external sites and communities.</p> <p>Follows the model of CERT (Computer Emergency Readiness Team <a href="http://www.cert.org">www.cert.org</a>), but is GRID-oriented. Perhaps a "GERT"?</p>	<p>To some degree, this may be already provided by the National e-Science Centre.</p> <p>GRID Security policy and strategy information is provided in this central repository. Capability for online community building and individual participation.</p>
<b>Education</b>	<p>Information on User-centred website describing security policy information and procedures. The scope of information is provided by the home institution and the NGS. End-users' rights and responsibilities are described.</p>	<p>Information on User-centred website describing security policy information and procedures. The scope of information is provided by developer consortia and the NGS. Developers' rights and responsibilities are described.</p>	<p>Information on Admin-centred website describing security policy information and procedures. The scope of information is provided by the home institution and the NGS. Admin-users' rights and responsibilities are described.</p>	
<b>Training</b>	<p>Information provided on a website, in pamphlets, emailing lists, and/or other means about networking events, training sessions, tutorials, and workshops that are open to end users.</p>	<p>Documentation on using tools such as the GLOBUS Toolkits, Web Services, and authoring tools with links to external communities.</p> <p>Information provided on a website, in pamphlets, and/or other means about networking events, training sessions, tutorials, and workshops that are open to end users.</p>	<p>Information provided on a website mailing list, and/or other means about networking events, training sessions, tutorials, and workshops that are open to Admins.</p>	

**DRAFT**