JISC

Project Information			
Project Acronym	e-Uptake		
Project Title	Enabling Wider Uptake of e-Infrastructure Services		
Start Date	1 April 2007	**End Date**	30 June, 2009
Lead Institution	National Centre for e-Social Science, University of Manchester		
Project Director	Rob Procter		
Project Manager & contact details	Alexander Voss, National Centre for e-Social Science, University of Manchester, Arthur Lewis Building, Oxford Road, Manchester, M13 9PL		
Partner Institutions	Arts and Humanities e-Science Support Centre, King's College London, National e-Science Centre, University of Edinburgh		
Project Web URL	www.engage.ac.uk/e-uptake		
Programme Name (and number)	e-Infrastructure Programme, Community Engagement Strand		
Programme Manager	Matthew Mascord		

Document Name			
Document Title	*Deliverable D1.2, Community Engagement Report*		
Reporting Period	1 April 2007 – 30 June 2009		
Author(s) & project role	Alex Voss (project manager) Rob Procter (principal investigator) Marzieh Asgari-Targhi (researcher) Elpiniki Fragkouli (researcher) Stuart Dunn (researcher)		
Date	21.07.2009	**Filename**	D1.2 Community Engagement Report.doc
URL	*http://www.engage.ac.uk/e-uptake/e-uptake-deliverables*		
Access	☐ Project and JISC internal		X General dissemination

Document History		
Version	**Date**	**Comments**
0.1	30.09.2008	Initial version by Stuart Dunn and Elpiniki Fragkouli
1.0	08.10.2008	Revisions by Alex Voss, Rob Procter, Marzieh Asgari-Targhi, submission to steering committee
Final	21.07.2009	Final version by Alex Voss, Marzieh Asgari-Targhi and Rob Procter

ENABLING UPTAKE OF E-INFRASTRUCTURE SERVICES

COMMUNITY ENGAGEMENT REPORT

1 Introduction

This report presents the findings of a qualitative study of academic users of e-infrastructure services conducted as part of the e-Uptake project of JISC's Community Engagement programme. It is one of the deliverables of e-Uptake's work package one. We conducted 53 in-depth qualitative interviews with 55 researchers between January and April 2008. In a second round, we interviewed 57 respondents who we identified as 'intermediaries', e.g., people working in a research support role within information services departments or e-Science centres. The purpose was to provide an analysis of current barriers that are preventing academics from using e-Infrastructure, in particular, JISC-funded services such as: Access Grid Support Centre (AGSC); National Grid Service (NGS); Open Middleware Infrastructure Institute (OMII); UKERNA; MIMAS; the Digital Curation Centre (DCC); the National Centre for Text Mining (NACTEM); EDINA, and the Visualization Network (VIZNET).

Note to readers: this document presents an analysis of the data collected by the e-Uptake project focusing on uncovering some general themes and illustrating the kinds of issues uncovered under each topic. Such an analysis necessarily leaves out some findings and reduces the level of detail presented to aid readability. In order to provide access to the wealth of evidence collected by the e-Uptake project and to support the development of a knowledge base that further material can be added to and that can be further elaborated, we have established an online database of findings, which at the time of writing is being populated with content. A reference to this database can be found at:

http://engage.ac.uk/e-uptake/database-of-findings

2 Research Approach

In this section, we describe the conceptual background, the empirical approach and the implementation of the empirical work in workpackage 1 of the e-Uptake project.

2.1 Conceptual Background

The project took as a starting point for its study of the adoption and diffusion of e-Research approaches and supporting e-Infrastructure a conceptual perspective rooted in studies of science, technology and innovation. In Voss *et al*. (2007), we outlined this perspective, which is perhaps best summarised as a *mutual shaping* (Williams and Edge 1996) perspective. Mutual shaping takes issue both with technological determinist views that claim that technological development follows its own internal rules and 'impacts on' society on the grounds that technological innovation is not some exogenous process but is subject to the influence of social actors. It also takes issue with the contrasting perspective that technologies are entirely 'socially constructed' phenomena (Bijker *et al*., 1987) on the grounds that these fail to acknowledge how technical realities constrain the choices available to social actors. The mutual shaping perspective, then, assumes interplay between technological impact and social processes of adoption, adaptation and shaping of the direction of technological developments.

An important consequence of this view of mutual shaping is to see the decisions that different actors make as part of their work in a wider social and organisational context as being equally important as technological factors. Our investigation of factors inhibiting or enabling the uptake of e-Research approaches and e-Infrastructures therefore had to cover aspects such as organisational arrangements, career paths, training, funding, research policy, etc. While it is important to realise the role of the organisational and wider social context in the study of any technology, it takes on a specific significance when we consider the development of infrastructural technologies such as e-Infrastructures for research (Edwards *et al.*, 2007). Processes of infrastructural development are contingent on a large number of social and technical factors but while there is much uncertainty at any point in time, there are also “shared patterns, processes, and emergent lessons that hold widely true across the comparative history and social study of infrastructure” (*ibid*., p.1).

Edwards *et al*. suggest that “effective infrastructures are rarely ‘built’ in an entirely top-down, orderly, and blueprint-like way” (*ibid*., p.2) and that use of technologies, and in particular infrastructural ones, is often deeply embedded in a complex web of socio-material relations. Consequently, we also had to go beyond the consideration of single factors but had to try and uncover the complex interactions of different social and technical arrangements in configurations of technology supply, configuration, service provision, uptake, appropriation and usage.

2.2 Empirical Approach

To achieve the aims of the project we had to look beyond isolated, contingent or random problems that people encountered in deploying and using e-Infrastructure services. Rather, we sought to identify recurring, widespread barriers that could be overcome by a set of targeted interventions that the project was making or that suggested strategies, which might be followed up by e-Infrastructure stakeholders. Furthermore, the study had to reflect the diversity of the target population (research active members of the UK academic community), their different interests and possible uses of the services (from the Access Grid Support Centre to the National Grid Service) as well as the number of potential factors influencing uptake (from individual practices to organisational factors and wider research policy). It was important for us to sample not just the views of early adopters but also those of people who have not yet engaged with e-Infrastructure services in order to understand the factors underlying this. In addition, we needed to contrast the information gathered from academic end users with the views held by service providers and technology providers as well as intermediaries such as application developers, e-Science centres and academic hosting institutions.

The first step in our research was to review the existing literature on uptake of e-Infrastructure services in research and to analyse existing data collected as part of previous and ongoing activities of the project partners. The most significant of these were the Study of Users’ Priorities for e-Infrastructure for Research study (SUPER: Newhouse *et al*., 2007), the AHRC e-Science Scoping Survey (Anderson, 2007), the Log Analysis of Digital Resources in the Arts and Humanities study (LAIRAH: Warwick *et al*., 2006), the Accelerating Transition to Virtual Research Organization in Social Science study (AVROSS: Barjak *et al*., 2007), the Dealing With Data report (Lyon, 2007) and the Scoping e-Infrastructure Usage report (Mascord *et al*., 2007).

The SUPER study focused on identifying needs that could be addressed within existing funding streams. Using a series of unstructured interviews 2006 with representatives from UK e-Science projects, it uncovered five major areas of concern: distributed file management and policy relating to their curation; tools to support the creation and use of dynamic Virtual Organisations; the need to support projects through tools and training; the operational provision of authentication, software licensing, and reliable consistent environments across distributed resources; and user Interaction with e-infrastructure services through defined APIs, scripting environments, and graphical workflow environments.

The AHRC e-Science Scoping Survey focused mainly on assessing the opportunities for exploiting e-Infrastructure within the Arts and Humanities research communities. It did, however, also note that exploitation of e-Infrastructure was hindered by lack awareness.

The AVROSS study examined, via eight in-depth case studies, the barriers to the wider adoption of e-Infrastructure in the Social Sciences, Arts and Humanities within Europe. It reported a series of issues hampering wider adoption, including: technologies (e.g., lack of robustness, lack of interoperability), community engagement (e.g., poor outreach, difficulties in developing sustainable funding); academic cultures (e.g., disruptive effects of change, lack of incentives and career structures for innovation); and the absence of high profile success stories and exemplars.

Reviewing these reports resulted in a list of barriers identified in the literature as well as an initial list of enablers and candidate interventions to address them. The findings identified were organised through a typology covering and further detailing the various dimensions identified above. Based on the understanding gained through the literature review, we developed a two-staged fieldwork approach that focused on researchers in the first stage and on research computing services and other intermediaries in the second.

For the first group of respondents, we developed a questionnaire eliciting basic information about their background and research activities, their involvement in e-Research projects, their experience in using advanced ICTs, their use of JISC-funded and other services and their use of support and training mechanisms. The questionnaire is included in Appendix B. For phase I, the interviews were based on the information obtained through the questionnaire.

The interviews were conducted by telephone or, where feasible, face-to-face interviews and they would take the form of semi-structured open conversations with the aim to elicit more contextual information about researchers' use of ICTs. A proforma was developed to guide the interview process, based on a set of questions that pick up from the answers given in the questionnaire. The proforma is included in Appendix C.

The second round of interviews focused on members of staff within research computing services in research institutions as well as other intermediaries such as staff at e-Science centres. Here, the approach taken was to conduct interviews guided by a proforma oriented around questions about service provision at the research institutions studied. This proforma is included in Appendix D.

3 Implementation

This section describes the practical approach that was taken to collect data for the study, to process and to analyse it into usable outputs.

3.1 Stage 1: Researchers

The identification of suitable respondents was a major issue we faced. We knew that lack of knowledge of and experience with e-Research and e-Infrastructures would limit the value of responses from people chosen randomly from the research community, so we needed to define a sample of early adopters or people we could reasonably expect to be familiar at least with the principles of the use of advanced ICTs in research. Candidate respondents were identified using a combination of web searches, use of existing databases such as the UK research councils' databases of grants on the web and web mining. We found that it was relatively easy to compile long lists of candidates but that filtering them using our selection criteria (active in research and using at least one of the JISC-funded services) involved a large amount of manual work to compile the required information from publicly available data.

We therefore decided that it was practically impossible to define a suitably stratified and representative sample *a priori* and that an iterative approach was needed that would monitor the coverage achieved along a number of dimensions as the interview process progressed. In addition to the primary stratification by research disciplines, we were also aware of the fact that other dimensions would be relevant as they influence the kinds of barriers that researchers face and the way they react to them. For example, researchers at different stages of their careers may have different interests, attitudes towards technological innovation, skills as well as investments in standard methods and tools. Consequently, we were aiming to ensure that our sample includes respondents of different levels of seniority.

Before the main study began, a pilot study of four interviews took place, where the online survey and the interview schedule were tested in the field. There seemed to appear no problem with the tools themselves. However, the initial approach email had to become more personalised, in order to better engage the candidate interviewees. The design of the questionnaire and the interview schedule were substantially informed by earlier work conducted in collaboration with the eIUS project, involving 13 face-to-face pilot interviews and by two detailed meetings with Jan Rae, of the Open University's Institute for Educational Technology, who acted as an external advisor.

We conducted a total of 53 interviews with 55 individual researchers, yielding more than 25 hours of recorded audio. Before the interview, respondents were asked to fill in the short questionnaire so that the interviewers would have some baseline information to guide the interview process and in order to monitor our coverage.

Interview candidates were sent an initial approach by email presenting the aim of the project and asked if they would be willing to participate. When we received a positive response, we sent the participant an email containing the link to the online survey, a consent form that corresponded to the framework agreed with eIUS and ENGAGE and an information sheet regarding the community engagement projects. As soon as the participants completed the questionnaire and consent form, a telephone interview was arranged with them. The interviews usually lasted around 20-40 minutes each.

The questionnaire, which informed the approach to the individual interviews in phase one, aimed to provide an extensive overview of the characteristics of the services used by the participant, their use of support and training offered, and the possible enablers and inhibitors to their use of advanced ICTs for research. Although flexible and responsive to a degree, the subsequent interview was guided by the common interview schedule and the responses made to the questionnaire.

The interviews explored participants' use of advanced ICTs for research, their stance towards ICT as well as their community's perception and use of ICTs. Respondents were explicitly asked about enablers and barriers to using e-Infrastructure services. The interview schedule consisted of open-ended questions, informed by answers given in the questionnaire. The interviewees were given the opportunity to discuss the issues that concerned them, but the researchers aimed to ensure that all of the questions in the interview schedule were addressed.

Participants were asked which services they had used, and how they had found out about them. Where he or she had more than one role in a project; or was involved in different projects, the interviewer established which area of expertise the subject would be most comfortable discussing, and based the interview on that. The interviews focused on how use of e-infrastructure services facilitates or enables research. In particular, the approach was framed with regard to four specific stages of the research lifecycle: data collection, analysis, collaboration and publication. In order to reflect the study's terms of reference, the participants were asked to indicate if they felt any barriers encountered were directly concerned with a funded service, or whether it was related to external issues.

We asked participants if they had actively sought any kind of support in overcoming barriers in their research at these points. For example, it was asked if research support had been sought via local helpdesks. We asked what support had been received, and whether this was provided locally or institutionally or nationally, e.g. by JISC as a service or discipline specific support service; and if there was any factor or provision that could enable and/or facilitate *further* use of the services.

In our interviews we have used the different stages of the research lifecycle as a structuring device, from the birth of an idea through its recognition by being funded, getting worked on and validated by researchers, to the results being published and used in further research (Voss *et al*., 2007). It should be noted, however, that there is no single way of describing the research lifecycle and, in practice, the picture may be complex. For instance, the starting point for research can sometimes be influenced by political decisions within an institution or the development of new funding programmes. Project planning can also be determined by immediate needs of an institution and practice-led development of applications may be undertaken without detailed research. Practice-led research in the Arts is often heavily influenced by the personality of the artist/researcher and does not necessarily involve a detailed literature review or structured project planning. Nevertheless, the different stages outlined in the research lifecycle will take place at some point during most research activities, though maybe in a different sequence and priority. The lifecycle should be imagined as an ideal type, i.e., an abstraction that helps to understand the different stages of research and to organise this report.

However, we find it helpful to work with this slightly idealised model to structure our interviews. The starting point in any research area is an initial research idea or hypothesis that will lead to the creation of a consortium of researchers writing a project proposal. This involves defining the problem and developing research strategies. If the proposal gets the go ahead from funding bodies, there will be a phase of negotiation to work out the detailed division of labour and a project plan. Project execution often starts with a literature review followed by data collection, analysis of the data collected, and discussion of the findings. Collaboration between researchers is a feature of most of these activities although the extent to which activities are undertaken collaboratively may vary. Finally, dissemination of findings through publication and preservation of scientific findings is crucial for the sustainability and follow-on of the scientific work. In the conduct of the interviews, we have used the research lifecycle as a structuring device but we acknowledge that this is merely a practical choice. We refer the reader to the online catalogue of findings for a different, more dynamic and customisable form of presentation. For more details regarding the concept of the typology of barriers, please refer to Voss *et al*., (2008) and for a complete list of categories and their descriptions, please see the Typology of Findings report (Deliverable D1.1).

3.1.1 Questionnaires and Interview Coverage

The questionnaire data were used mainly ahead of the interview in order to acquire personal information about the participants. They also offered an initial picture of the participant's practices so that the interviewers were better prepared for the interviews and were aware of the issues that needed more discussion with each participant. Below we present diagrams produced from the questionnaire data that give some information about the participants such as their disciplinary affiliations and seniority.

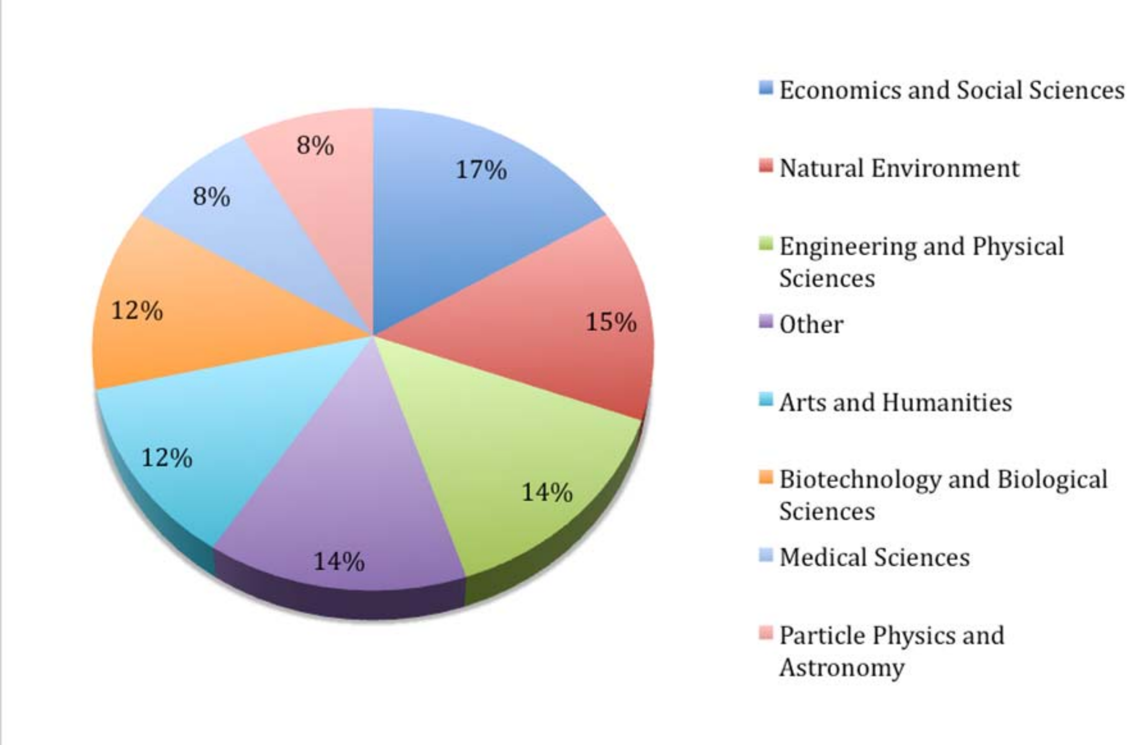

Figure1: Disciplinary affiliations of interview participants

The first diagram (Figure 1) shows the respondents by discipline. Please note that the categorisations here are the ones given by respondents in the questionnaire rather than the ones we used initially when identifying candidate respondents (these latter ones are used in the codes used in the presentation of findings below).

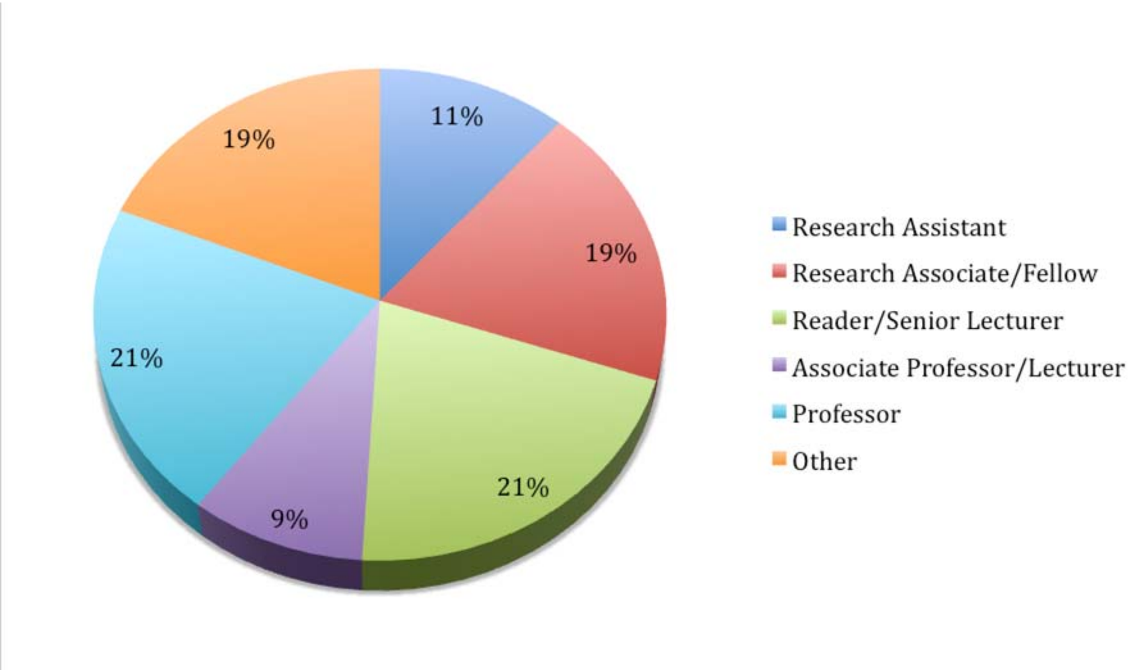

Figure 2: Respondents by position

Figure 2**Error! Reference source not found.** depicts the respondents' seniority. No attempt has been made to map free-form answers to the categories used to capture seniority although this might be possible in some cases. The diagram demonstrates that we achieved a good degree of coverage in terms of the seniority or respondents.

When asked about their experience with advanced information technologies in research and teaching, about half of the respondents said they had been using such technologies for 10 years or more. 30% had used them for less than 5 years and 18% for 6-9 years. At the same time, most respondents judged their related skills to be quite modest, with 75% judging themselves to be more novices than experts. This confirms that we have managed to go beyond the group of relatively expert 'early adopters', who have strong interests in e-Research activities and the necessary skills to pursue them. At the same time, we found that 89% of our respondents were currently involved in one or more projects involving advanced information technologies, so the group of respondents was also well placed to comment on the issues discussed.

Data obtained from the Higher Education Statistics Agency (HESA)[1] data shows that our interviewee sample broadly reflects the profile of the overall sector in terms of academic roles and disciplines.

Figure 3**Error! Reference source not found.** shows the awareness of services for each of the respondents. We asked respondents about 10 different services funded by JISC: Access Grid Support Centre, Edina, Mimas, UK Data Archive, Digital Curation Centre, National Centre for Text Mining, National Grid Service, OMII, UKERNA and VizNet. 16 researchers were frequent users of at least one service, 21 declared that they were using services only occasionally, 11 said they were aware of services but did not use any themselves and one respondent did not know about any of the 10 services we asked about.

While these figures seem to indicate that many researchers are aware only of a few services JISC funds, one needs to consider that different disciplines will be making use of different kinds of services. For the purposes of analysing our coverage, it is important to note that, on average, 68.3% of respondents were aware of any given service. Similarly, on average, 21.9% were at least occasional users of a given service. Given the expectable differences in awareness between different disciplines, these numbers again indicate that we have managed to cover the field. Even for some of the less frequently used and more specialist services, we managed to find respondents that were able to provide information.

3.2 Stage 2: Intermediaries

Respondents for the second phase of the fieldwork were recruited from research computing groups at research institutions as well as from groups of other intermediaries, e.g., from UK e-Science centres. The main emphasis was on representatives from universities with a large research volume and above average success rate. To ensure a degree of representativeness, we chose institutions from three groups. Firstly, we included institutions traditionally found at the top end of league tables, then institutions that are outperforming their peers and, finally, some small, specialist institutions that were included because they had strengths in particular research areas or were particularly active in e-Research. Figure 4 illustrates how we use the THES 2007/08 ranking and HESA funding data to select institutions. From each institution, we tried to recruit two respondents, one with a more strategic management role and one with a role that would bring them into direct contact with researchers in their everyday work.

[1] www.hesa.ac.uk

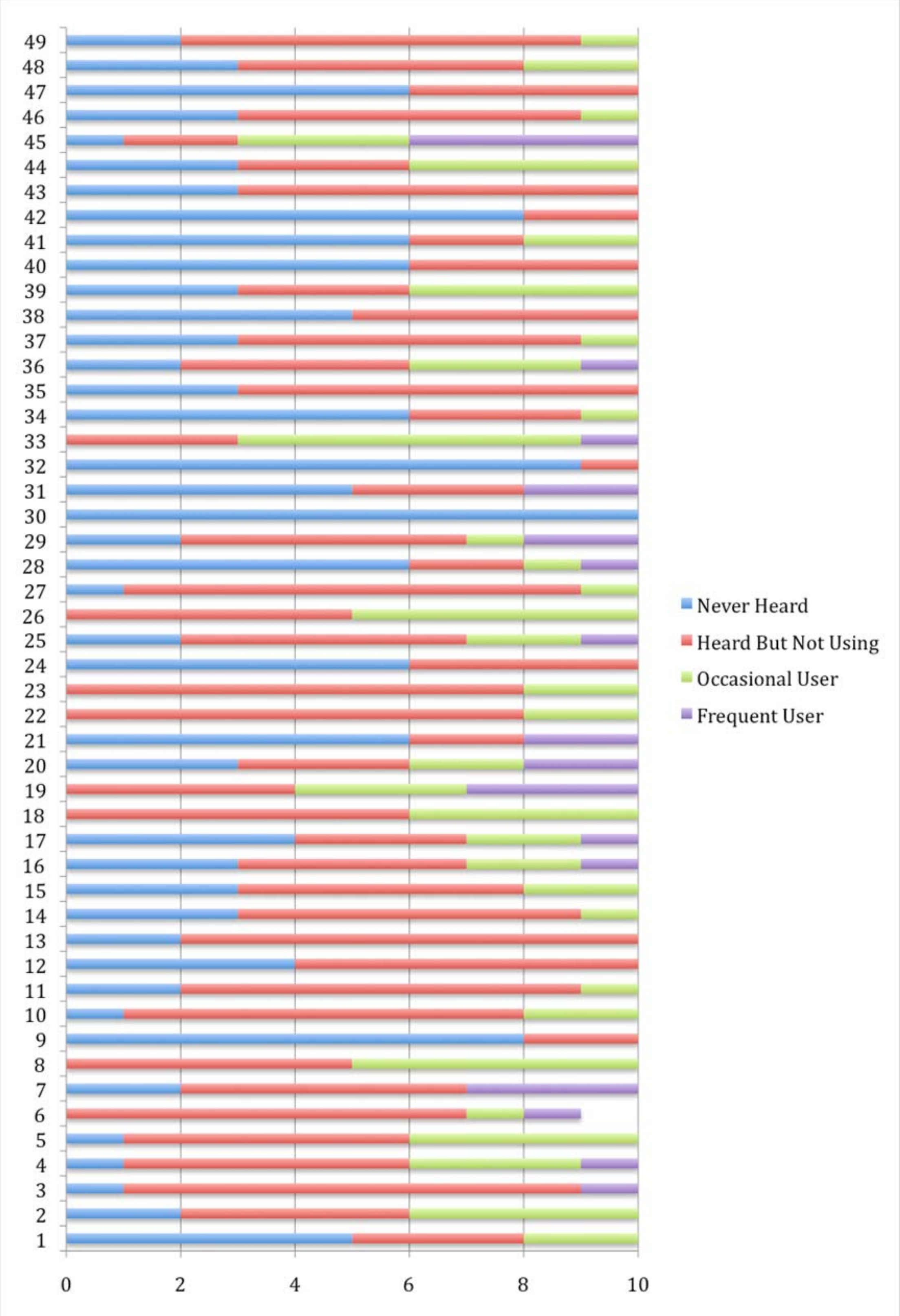

Figure 3: State of Adoption of Services – the x-axis represents the number of services

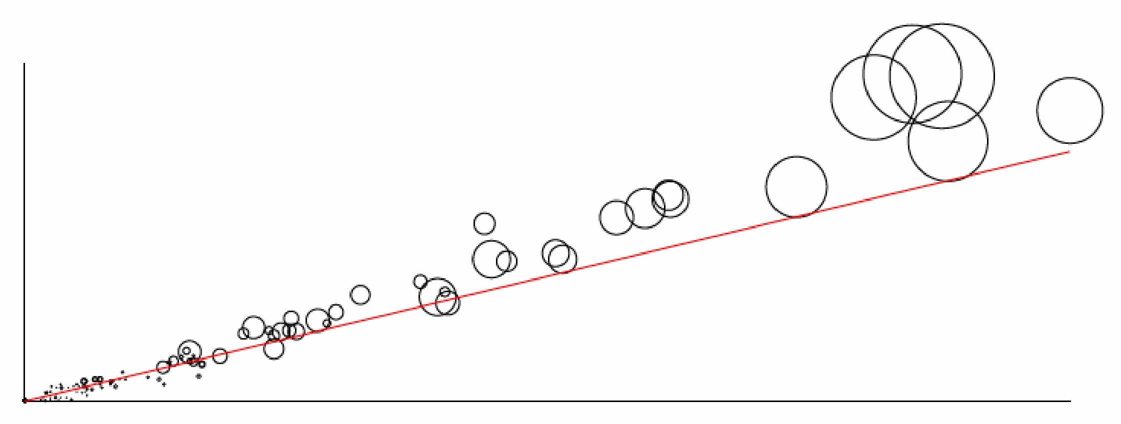

Figure 4: UK universities by THES 2007/08 ranking (number of awards on the x-axis and number of grants on the y-axis) and HESA funding (2006/07, bubble size). The red line represents the mean success rate of around 24%.

The interviews lasted between 20 minutes and over two hours in this second phase. Respondents were asked about the existence of a research computing group or other arrangements in their institutions as well as about the organisational role, contacts with other relevant groups and funding mechanisms. Respondents with a strategic role were asked about their institution's strategy for research computing, the drivers behind it and the embedding of this strategy in the wider context of the institution. The interviews then moved on to the local resources available, their usage as well as training and support provision. Respondents were asked to comment on their role in facilitating usage of national services, any local e-Science investments, the existence of a UK e-Science Registration Authority, issues they had experience with, unmet demands as well as what they thought JISC or the wider e-Science community should do to increase uptake. The interview proforma can be found in Appendix D.

3.3 Method of Analysis for Phase I and Phase II Fieldwork

All interviews were audio-recorded, resulting in more than 70 hours of recorded material. The interviews were fully transcribed using a commercial transcription service and quality assured by the researchers. The data from both phases of fieldwork was initially coded up using free coding (i.e., without a pre-existing coding schema) and the lists of codes generated were factored into the typology developed from the initial literature review. As the coding scheme contains in excess of 150 detailed codes, it was organised hierarchically to facilitate navigation.

From the fieldwork data, we extracted instances of inhibitors and enablers that were described by the respondents. Each has a short descriptive title, a longer description and is supported by at least one more specific example supported a quote from the fieldwork. In addition, inhibitors and enablers were linked to the typology and further classified as being either general or specific to a particular discipline, service or institution.

A review of traditional Qualitative Data Analysis packages (Voss *et al*., 2008, see outputs) showed that these tools did not support some of the functionality required by the project such as the integration of survey and questionnaire data or the collaborative use across organisational boundaries. We therefore decided to develop our own approach to representing the data. We initially used an XML schema developed by ESDS (Cummings, 2006; Milosavljevic, Grover and Corti, 2007) and developed a web-based interface for browsing and searching anonymised extracts of our fieldwork data on the basis of the questionnaire data and our typology. However, considerations such as the effort required to mature this approach and to achieve sustainability led us to adopt a more pragmatic approach in the end. We are currently using the Connexions system at Rice University as an alternate platform (www.cnx.org) for representing the data. We also intend to make this database available within the ENGAGE portal to allow future researchers as well as other stakeholders such as service providers and JISC programme managers to mine the information produced.

4 Presentation of the interview data

The presentation of the findings follows the typology of barriers and enablers outlined above. There are three main categories for barriers and enablers: social, technical and digital resources, each of which breaks down into further sub-categories. Where barriers or enablers are (particularly) relevant to specific services, this is noted and we use cross-references to reduce the amount of duplication where issues are mentioned under more than one heading.

All the interviews have been anonymised and a reference scheme is used in this document to make it possible for the reader to understand which discipline area (as defined by research councils) each of the respondents mentioned or quoted works in. Within the research council areas, which are indicated by the first two letters of the council's acronym, each respondent is assigned a unique number. For example, EP5 is respondent number 5 working in Engineering and Physical Sciences. Profiles of the candidates can be found in Appendix A. Similarly, respondent IT7 is an intermediary from the second phase of interviews and their profiles can also be found in Appendix A.

The report contains only a limited amount of information compared to the wealth of information available in the fieldwork material. We are working to make a more complete set of findings available in the online database of findings, which can be found at http://engage.ac.uk/e-uptake/database-of-findings. A definition of each heading can be found in D1.1, the typology report. We use *italics* to highlight quotes from the fieldwork material and **boldface** to flag key findings.

4.1 Social Issues

The following sub-categories were identified: training, education and outreach; user-designer relations and user requirements; understanding disciplines; collaboration; policy and funding; organisation of disciplines; individual issues; organisational issues; ethical and legal issues.

4.1.1 Training, Education and Outreach

Issues related to training, education and outreach seem to be major barriers, especially early engagement of researchers and the connection between different stages of engagement. Lack of **awareness** of services is a main barrier to their use. In general, there seemed to be a lack of a systematic introduction to **services and training,** which results in a lack of awareness as well as a lack of understanding of how services and methods can facilitate research and what different options exist. This was mentioned by a number of our respondents (e.g., MR2, EP3, AH4). One remarked that they are not aware of many of the JISC services we asked about in the questionnaire and that they might find that some of them are interesting but they have never heard of them before: "*one barrier is having not heard of these things*." (AH3).

Researchers appear to find out about these services through **events** such as the UK e-Science All Hands Meeting (AH6), through colleagues (AH1) or workshops (AH3, AH7) and it is often through their personal initiative that they acquire the skills to use these resources. Of importance are also the effects of **boundary spanning**, when people move from one discipline to another (cf. the account of an A&H researcher and their engagement over time with the e-Science programme recounted in section 4.1.5).

Embedding **education and training** related to e-Research methods into postgraduate courses was seen as essential for sustainability and wider uptake: "*one of the biggest challenges that faces us is translating all of this stuff into the next* *generation of suitably equipped PhD students and young researchers […] we need a much more structured approach to postgraduate training than we have at the moment*." (IT8). It is important, however, to make a distinction **between specialist and general e-Research skills** so that researchers are not burdened with information about low-level services they will not use in practice: "*it's never going to happen that you get clinicians to sit in front of a machine using very low level services. They just want to see something very simple with a few buttons*." (IT10).

In areas where the use of e-Research tools is already quite common, we find evidence of progress in the provision of **systematic training** of early career researchers: "*so we have some of the OMII people have been helping us run the little short courses, so two-three day courses on e-Science or Life Sciences and these are actual workflows for Life Science and Medicine and that sort of stuff, and they're incredibly useful just to some people like PhD students and post-docs. Just try and to get people up to speed*." (BB5). Clearly, such routine arrangements would be of immense value in other areas but they rely on having a **critical mass** both on the demand and the supply side. A medical researcher was in no doubt that there is an urgent need for capacity building within their field: "*Well, yes, biology is now a data rich science now in the same way as particle physics is because of its technological innovations and but the personnel hasn't caught up at all, there is still a massive shortage of computational biologists people*." (MR4).

One researcher (NE6) outlined the difficulties users face when engaging with grids for the first time. From their experience of being involved in cluster computing and grid applications, they commented on the importance of bridging the **gap between initial interest and training**. In relation to the kinds of training available they echoed the comments of the medical researcher (MR4): "*perhaps there's a general need for more training that's aimed at the domain experts rather than e-Science experts*." (NE6). Also, they expressed a need for more hand-on consultancy style assistance: "*unless someone […] actually goes and really works very closely with the scientist to set up things that they can use it really isn't going to happen […] there is a need for more people to sit down with scientists and work with them on their specific applications […]*" (NE6). When asked about provision of e-Learning material, an intermediary commented: "*I think hands on training is the best way forward, I mean you can have as much as you want of e-Learning training but you don't get anywhere unless you've got hands on training […] on distributed computing*." (IT30).

While training material on the technical aspects of using e-Infrastructure exists, several interviewees commented on the **lack of material** on the role of e-Infrastructure in the context of particular application areas and research methods: "*We provide training material but it's more based on the technical use of products rather then the researchers' requirements for the product, and as many people know if you throw a data set at some package like SPSS it will give you an awful lot of information whether it's what you wanted and whether you know what you do with it is another matter*." (IT20).

There is also a question of how outreach, training and support are provided and who should play a role in **developing new application areas**: "*the services themselves could probably cater for Arts and Humanities people or think about how they could be used by Arts and Humanities people and try to advance some things […] I know the Arts and Humanities e-Science Support Centre tried doing this, I know that the Arts and Humanities Data Service have tried to get people using these technologies but it would be nice if the central services themselves could actually have a think about how Arts and Humanities people could use them […] and run events based on their individual services*." (AH1). This call for more direct involvement of services is being met by the ENGAGE initiative and increased NGS user involvement activities.

An opportunity that has perhaps not been exploited to a large enough extent is to **link e-Research with ICT training programmes** that already exist within some disciplines. For example, one respondent mentioned that her department had been involved in a summer school on digitisation for a number of years: "*we used to run a very successful digitisation summer school for cultural heritage professionals […] and now we run a preservation summer school*." (AH4). Clearly, integrating e-Research training with events like this will serve to reach a wider audience of young researchers who would otherwise be difficult to reach.

Regarding raising awareness of services and of e-Infrastructure in general, it was suggested that intermediaries should play a **more active role in dissemination**. For example, one A&H researcher mentioned as a barrier the fact that there is no link on the library site for these services: "*if some of the JISC resources were a bit more prominent it might help people actually use them*." (AH3) They also advised that it would be very beneficial to the uptake of those resources if the librarians would be more involved in their dissemination.

Two engineering researchers suggested that the Post Graduate Certificate in Learning and Teaching in Higher Education could be a good route to promote the adoption of e-Infrastructure services. In that way early career academics would be introduced to these at the **early stages of their training** (EP2). Another suggestion was that service providers should take advantage of institutional staff development programmes to promote uptake: "*so I would suggest a sort of travelling roadshow […] give presentations, go round different universities, […] show them what's available and show them how it can be useful*." (EP5). Again, the importance of bridging the **gap between initial interest and training** was stressed (EP5), an issue that will need to be take into consideration in the development of the e-Research Roadshow initiative.

Finally, respondents flagged the importance of **disseminating success stories**: "*It's been interesting […] when people start going out and about and just saying what it is possible to do just how much excitement you can generate, so I think having stuff where you can show that people have done really new science using those tools and using jealousy. […] it seems to be working quite well in terms of getting engagement and so the fact that you know, we're saying that other communities just like things like the systems biology communities are beginning to be very keen to play and join in*." (BB5). Clearly, it is important for the community to formulate clearly where e-Infrastructure usage has made a significant difference to researchers and to disseminate these success stories widely to inspire more researchers to start engaging: "*everything that we are about is not to provide thousands of CPUs to people – it's to make them do good research […] it's very important that we focus on the research outcome [… we have] demonstrated some but maybe not enough*." (IT49, also cf. IT20). The JISC-funded eIUS project, for example, has produced use cases and video presentations highlighting the crucial role of advanced ICTs in leading research projects.

Linking material that clearly describes pathways to adoption in specific application areas to concrete training material would also help to overcome the problem that researchers do not have **enough time for exploration of possible use cases and technical configurations**. The underlying problem is that potential users lack the time to develop the necessary skills and insights on their own: "*I can see that there are things there which we probably could be able to use in the future but first we'd have to work out how, if you know what I mean? There are projects for example like OGSA-DAI and OGSA-DAI has some features which I can see they would be useful if we had skills or if we had the time to actually be able to get far enough into the technology to be able to actually utilise it properly*." (EP5). Note also the related discussion in section 4.1.7 on the lack of funding for such exploratory work.

4.1.2 Technology Supply Relations

Respondents stressed the importance of **understanding user requirements** and of using these to inform the development of tools and services (e.g., ES5, ST1). Respondents had various suggestions as to how this might be achieved but the underlying message was the need for **closer collaboration** between discipline scientists, service developers and intermediaries. According to one environmental scientist what is required is **hybrids**, i.e., intermediaries with knowledge of research domains as well as computing: "*there is a need for more people to sit down with scientists and work with them on their specific applications [...] people that understand both, people that can understand the applications and also understand how to grid-enable them*." (NE6) (cf. also section 4.1.1).

Issues relating to **access to local expertise** were flagged repeatedly by researchers as well as by intermediaries. Often, developing working systems involves more than merely putting in place technical components but involves developing complex socio-technical configurations: "*we have [a developer], who's been working for a couple of years on building a portal for [a research institute] for doing qualitative trait analysis and I think they were the first one to issue a portal certificate, which was alien to the NGS because it means you don't know who the user is […] that means the uptake is much higher because suddenly hundreds of people are using the portal and thereby using the NGS*." (IT10). What is important to note here is that the development of the portal involved a change in policies on the side of the service provider, not just the installation of a technical system.

Active **outreach and engagement** through direct contact can be an important enabler, especially to overcome the initial hurdles in taking up e-Infrastructure services: "*we have what I could probably term as almost a campus champion for Grid computing who is the person that goes out and works with new users. This person also then works with those users and assesses their requirements and then, if necessary, links them into the national infrastructure so hand holding them through getting membership of the UK NGS and getting their application to run through the campus Grid infrastructure onto the NGS system.*" (IT48, also cf. section **Error! Reference source not found.**).

Communicating effectively what e-Research might be and how intermediaries such as research computing services or e-Science centres might help is a challenging task. As a consequence, researchers may **not know what to expect**: "*if you buy a car, you know, most people know what you get and what you should look for. But if you come to us what can we do for you is always vague because we can do quite a lot we've got a lot of expertise, but you know if it was write it down in a brochure it would be kind of be a long-winded one, and often they don't understand all the jargon that we have and all that sort of thing*." (IT10). Users may not **understand the technical challenges** of implementing applications, so intermediaries need to spend time on educating users managing their expectations. This, of course, is related to the observation that researchers often lack an understanding of what distributed systems are and how they can support research activities (cf. section 4.1.1).

An important concern in the collaboration between researchers and computer scientists is the **lack of a common language** and the technical language used by technology developers and service providers (e.g., ES8, ES5, EP5, IT10). A respondent observed that the use of computer science terminology without prior introduction in e-Research projects alienates researchers: "*it's one of the things that's sort of continually frustrating in the field is the assumed terminology if you know what I mean? That there's a lot of terminology that's come over from computing science which is never designed for the rest of us who actually do the science […]*"

(EP5). For their part, intermediaries acknowledged the importance of them making greater efforts to help create a common language: "*now of course that's a bit bad for me to say I would like them to understand me but I should also try to understand their needs*." (IT10). What a lack of a common language exposes, however, is not simply the need for intermediary-hybrids but also the need for the development of a common ground that supports a discussion about requirements and other issues on the basis of a better mutual understanding and appreciation of concerns.

Focus groups and other methods from the social sciences can potentially be adopted to establish a dialogue between technology providers, support staff and researchers. "*The good thing about that is that it meant that we had to do the focus groups with the consultants and it meant that we could be closely involved with all these people coming in and saying what were the barriers to them using a national data service, but it was the contact with the users that was great*." (IT50). While focus groups are suitable to explore issues in depth, other methods such as surveys can be used to elicit information about the wider importance of issues and other quantitative information: "*one of my objectives each year is to devise, in conjunction with colleagues, a survey of researchers. I know that those surveys will take different forms so, for instance, we're just completing one which is looking at research centres.*" (IT56, also cf. IT45, IT48).

While these methods are useful for systematic data gathering, the development of **personal contacts** should not be neglected. Individuals can act as 'champions' who have a key role in promoting the uptake of e-Infrastructures in their departments or in the wider disciplines and can channel the communication of requirements into service provision and support: "*I think they've started to develop a kind of model of a school champion, so a champion from every single school in the University […] so the idea is to bring someone in who can represent their school*." (IT22).

4.1.3 Understanding Disciplines

Concerns related to who the stakeholders are and what they are doing were expressed especially by Arts and Humanities researchers. Respondents mentioned that there is an impression of **services being designed for scientists** and services' primary users being scientists with necessary technical skills (rather than considering Arts and Humanities researchers who need further introduction to the services). Such an impression can inhibit A&H researchers using digital resources as they may **not perceive them as relevant** for their particular discipline. An A&H researcher offered the view that there was a lack of fit between the services offered and the needs of their discipline: "*I think the main barrier is that we need to develop more services for the type of research I do*." (AH6).

A respondent from an information services department commented that it is difficult for them to go beyond the provision of infrastructure and get involved in discussions about the applicability of particular methods for specific research purposes or about research questions themselves: "*We can certainly help you get the results, what we can't help you do is understand what to do with them or what you were looking for in the first place*." (IT20).

Some researchers reported finding that e-Research technologies did **not provide sufficient immediate benefits** to justify the effort: "*We tried to look at using some of the technologies to do some work with historical census records. You know, theoretically it would be great to be able to pull in all the different datasets from all the different providers using grid-based stuff but, actually, in reality it was going to be much easier to collect them and have them locally [...] and we weren't going to be able to [add] any data to that because we would have had to digitise it ourselves*." (AH1). Clearly, even if the vision of how e-Research technologies could be applied is convincing, the potential benefits might not be tangible enough to warrant a concrete investment of effort in the short term. Researchers might prefer to adopt a wait-and-see approach, continuing with existing practices that, while seen as less than perfect, can be made to work. Similar observations were made in the social sciences where downloading datasets and using them locally is still the predominant mode of use (e.g., ES1, ES3, ES7).

However, respondents in disciplines with a history of using advanced ICTs also reported that their needs are not being met. An astronomer commented that e-Infrastructure services are too **general for their specific needs**, noting that astronomy is quite self-reliant as a discipline: "*astronomy is fairly technologically sophisticated, a lot of services that these provide we provide for ourselves anyway and have done so for a long time and do so internationally*." (EP3). Another astronomer referred to the lack of specific support for the needs of their discipline and suggested that all too often what was being built was general solutions rather than the specific solutions they needed (ST1). They went on: "*They [NGS] clearly didn't think of the size of the data we are used to […] clearly [NGS] would not be expecting […] queries to databases that returned gigabytes of data, but we routinely do that in physics and astronomy. That was one critical flaw in the design*." (ST1). This example demonstrates how specific research needs that go beyond the usual requirements can **challenge assumptions** made in the development of e-Infrastructure and potentially give rise to new developments.

While in some disciplines or more specific research areas, use of e-Infrastructure services is an **option**, in others it is clearly **essential**: "*the research work I do does involve use of electronic resources and it involves processing of them using software applications and it's linguistic research that wouldn't be possible without computational tools*." (AH6). Another researcher commented: "*text mining for us is absolutely central and crucial. Most of the knowledge is out there in the literature and there is a huge need to actually pull that in effectively*." (BB5). Research areas in which the use of e-Research methods has become essential often provide a template for the development of similar practices elsewhere. However, the translation from one area to another is not necessarily straightforward or a linear one.

Other differences are related to the relationship between tools and research methods. While in the physical sciences, ICTs have long played an important role and their use does not normally raise any concerns about their relationship to the **generation of scientific knowledge**, this is different in other disciplines. For example, a social scientist described their problem getting e-social science work published because the **traditions of publishing** in the social science differ significantly from the sciences: "*when you look at research papers in other disciplines perhaps in medical research, there's just a cross reference to some previous description of the dataset or previous description of the work and [in] my experience in the social sciences publications we are not so happy with that [cumulative] approach we usually want every article to more or less be able to stand on its own and be accessible in its own right which is a real challenge to describe the complex research*." (ES7). It can be more difficult for social scientists to describe complex e-Research projects drawing on multiple datasets and using relatively complex and less established methods because the scholarly traditions favour papers that are self-contained.

4.1.4 Collaboration

Collaboration has often been flagged as an important aspect of e-Research. Clearly, researchers' willingness to collaborate is an important precondition for many aspects of e-Research, especially when it comes to the sharing of resources that are not nationally provided. This **collaboration readiness** (Olson *et al.,* 2008) is dependent on a number of factors and can play out in different ways. While it may be of advantage for researchers to form consortia to bring together sufficient skills, human effort and other resources, there are also disincentives for collaboration: "*To a certain degree research centres are in competition. So there is a certain reluctance I think in individual research centres within the same discipline sharing what they're doing. They may well be bidding for a research project in competition with each other*." (IT20). So, while there are incentives for collaboration and while often research can progress only through collaborations, **research is competitive** and the alignment of interests achieved between different researchers or research groups remains a temporary one.

The scaling up of synchrotron resources means a scaling up of data volumes and potentially also an increased interest in **remote instrumentation**: "*that's the particular point at which we have large amounts of data and they're collected remotely so we wanted to control the synchrotron remotely, we want to be able to control that from the lab which is, again, quite difficult, quite new, that's where some of the secure communications come in*." (BB2).

However, remote instrumentation is described as being still **under development** and the respondent also points to the logistical issues involved in using such a facility: "*In the past until 6 months ago you pack everything up and you take a plane or a train and you go there and you work overnight and then you come back again, so you usually have a shift of 24 hours that you're expected to work 24 hours solid, and you go with a small team of maybe three people to do that work […] well the technologies for shifting crystals at cryo temperatures, liquid nitrogen temperatures, has improved so it's now feasible to ship crystals to Grenoble and now the software is beginning to develop where you can start to control the beam line remotely so we need to send fewer people and you can have people at home doing some of the work*." (BB2). This example clearly shows how the development of e-Research practices is part and parcel of **developing disciplinary practices** and cannot be separated out from the development of tools and methods. The example shows how other developments such as improvements in transport techniques combine with the new affordances to create quite a significant change in working practices of researchers.

Collaborations among different disciplines were considered potentially problematic, due to general **cultural differences** (MR6) as well as specific difference in attitudes, skills and professional languages used. As one respondent explained: "*when you are working with sociologists they might tune out when you are talking about quantitative analysis, or likewise some quantitative scientists may not fully appreciate the qualitative analysis*." (MR2). There is clearly a need to develop a "*multidisciplinary culture and knowledge to be able to manage projects that goes beyond technology*." (MR2).

The importance of choosing the right **technical solution** for a particular collaboration situation was a common theme. Researchers reported uses of the Access Grid and reasons why they preferred it over options such as email or phone conferences in some circumstances: "*it wasn't possible to arrange a face-to-face meeting and we were able to meet […] in order to set up a project consortium. It was very useful because when you have to negotiate you need a face-to-face meeting really or at least with the Access Grid it worked, it would have been very difficult to organise by email of by telephone*." (AH6). Another respondent made a similar point: "*if you want the sort of meeting where you can have a proper argument and a bit of a row, then Access Grid is good, if you're just touching base and keeping in touch, then Skype […] it's more effort to get an Access Grid meeting going than just a Skype but you can get a better meeting*." (BB5). It seems that some researchers were quite aware of the **trade-offs** involved in choosing on technology over another in terms of the effort involved and the likely impact the choice would have for the meeting (cf. IT31).

4.1.5 Individual Behaviour and Attitudes

Our interviews show how researchers' **personal careers** can shape the uptake of e-Infrastructure. As another A&H researcher put it: "*before I was at [my current institution], I was at an engineering department at [other institution] and so I was kind of aware of a lot of these things that we are talking about – Access Grid, e-Science*." (AH1). They recounted how they kept in touch with the e-Science programme from its early days even when switching institution and discipline area, for example through attending conferences. They used their experiences to come up with ways of applying e-Infrastructures in new contexts: *"[even] before the funding programme for e-Science had gone out I was aware of that, that it was happening and I thought 'oh, that is something we should look at for Arts and Humanities', so when the opportunity came for us to do something using e-Science technologies, I kind of hassled the research computing people [at my institution] to tell me about it*." (AH1).

The amount of time researchers are able devote to finding out about new possibilities can vary by enormously. One researcher commented: "*there are some of the scientists that are interested in computing grids and enjoy finding out about some of these things*." (NE6). A number of respondents (EP2, EP3, ES4) pointed out, however, that researchers do not have time to search for and explore new services or to attend events: "*as an academic it's difficult to […] keep track of what new things are happening because we're so busy doing so many different things that I do know there are a lot of really good bits happening but it's just getting time, although, you know, I don't have time to check the website*." (EP2, cf. IT3). Another respondent commented: "*I have loads of other things related to my research plus I do teaching, undergraduates and postgraduate, I also have huge admin responsibility so I don't think I would get round and do training really, there won't be much time left*." (ES4). Clearly, **time constraints** are an important factor, especially for outreach activities. These need to be carefully designed to address the right kind of audience at the right time and with opportunities for following up interest. This respondent also suggested that leveraging institutional incentive structures was likely to be more effective than a direct approach to individual researchers in raising interest (EP2).

However, **maintaining confidence** in the benefits can be difficult as this statement shows: "*I haven't really used [some services] much, I just think that's the nature of my research, […] if I was able to actually do things with the data that I had then I probably would have but it came back to the state where I have got questions about what I'm going to do and how useful will it be? […] in the end we didn't really implement anything. There was really little need to use the various services on offer*." (AH1). Clearly, where the value proposition of a service is marginal to the research undertaken, other, simpler, better known solutions will be preferred, even by those who can see the potential benefits.

Finally, at least one respondent referred to the effect of the **Research Assessment Exercise** (RAE) on incentives to work collaboratively: "*[It] encourages individuals to publish independently, to keep things secret while they can be many advantages to their career no matter if they have been funded publicly or not, because by doing that, they appear to be better by the criteria used for measurement of the research assessment exercise, that's the major cultural problem because it makes it too difficult to perused scientist to be open with their data, they fear losing it, and therefore their current position*.'' (MR5). The RAE also influences career opportunities for those involved in the development of e-Research tools and techniques: "*what's alarming is […] we take people on as research assistants but then we end up asking them to do development […] sometimes quite industry focused development which isn't really research and doesn't produce papers. And we also don't reward them for collaborating, we never reward them for that*." (MR9). This kind of approach, which is common in academic software development, creates an inherent tension between individuals' career goals and the need for production quality software engineering.

4.1.6 Disciplinary Factors

Collaboration within and across disciplines was another issue that was discussed by many researchers from different disciplines. Collaboration between researchers in A&H was presented as problematic as A&H research is very "individualised" and accordingly the things people ask questions about are very individualised as well; so the problems A&H researchers are facing are not "generic" enough to be "*a socially recognised category*." (AH8).

Another researcher points out that "*there are fundamental barriers to using these technologies for the Arts and Humanities but that's to do with the discipline […] rather than [with] not understanding the technology, not being able to get the technology. It's a matter of 'what can we do with these technologies which are useful, which are methodological questions and also a kind […] about what is Arts and Humanities, what kind of research are we doing?*" (AH1). These statements call into question the assumption that e-Research technologies are **disciplinary neutral**. Rather, their adoption may have significant implications for research practice for some disciplines and, as a consequence, be strongly resisted. "*Being able to fit an e-Science paradigm into Arts and Humanities is the problem rather than whether we can use the technology*." (AH1). This latter statement indicates that there is a potential clash between some established disciplines and the ideas behind e-Research.

One problem mentioned was an impression on the part of researchers that the e-Science programme was not aimed at doing research but rather had a technology focus: "*money came out of the e-Science programme and obviously when you look at the remit of the e-Science programme, it was to do infrastructure and to build middleware, that's what they wanted to do, and in fact the e-Science programme was very clear that you shouldn't, you weren't being funded to do science, you were being funded to do e-Science and I think that was a critical problem from the start, sending people down the road of doing software development, because it had to be hardcore software development and it discouraged people from actually doing science*." (ST1). Given the inevitable need to focus on developing basic technologies at the start of the programme, it is not surprising that the perception of some researchers was that the e-Science community was **dominated by technologists**. While this issue is being addressed more and more, it is difficult for these trends to be reversed and to build up a focus on research problems and enabling new practices instead.

This issue is closely tied to the problems of **multidisciplinary work** which were mentioned a number of times by our respondents (cf. section 4.1.4). These were related, for example, to different terminologies used. Respondents pointed to a number of practical ways of addressing this problem such as increased emphasis on regular interactions: "*I suppose you just talk more, you talk to each other until you have developed something resembling understanding*." (MR4).

Another issue that caused problems was the **lack of researchers with computing skills** (MR4, MR7). It was suggested, for example, that although biology has advanced considerably in terms of its use of data and computational resources, the staff skills have not advanced equally: "*well, yes, biology is now a data rich science in the same way as particle physics is because of its technological innovations – but the personnel hasn't caught up at all, there is still a massive shortage of computational biologists people*." (MR4).

There is also a concern regarding **publishing multi-disciplinary work** in academic journals. As one respondent put it: "*publishing in non-physics journals tend to get less recognition in physics community, and probably vice versa publishing in physics* *journals would not get much recognition in medical community so that's an issue to do with interdisciplinary working, it's probably fairly universal*." (MR6). A common solution to this problem is publishing the same material twice with different disciplinary emphasis but this does not seem to be the best way to foster genuine inter-disciplinary work. Although there are some interdisciplinary journals, there is still a lack in many emerging fields, like cancer informatics which, "*is a growing field and not ideal for publication yet*." (MR5).

4.1.7 Policy/Funding

Policy and funding issues raised covered a range of concerns. A medical researcher pointed to a lack of understanding on the part of research councils, where e-Science gets labelled as infrastructure: "*The fundamental barrier is support of research funders for this sort of discipline, because they see it as an infrastructure and they are really organized to fund hypothesis driven research, what we are trying to do is establishing the infrastructure that will make the hypothesis driven research much more efficient, so there is [a] funding issue*." (MR5)

Funding arrangements for services and (perceptions about) their **sustainability** were mentioned as determinants for researchers' decisions in the adoption of services: "*most academic software seems to have the lifespan of about a year before it disappears or doesn't get updated or something, and we just can't work like that so the reason why we would always want to work with people like OMII and the Text Mining rather than some of the academic groups within the space is that we know that there's longer term support available, so the stability matters to us*." (BB5). Clearly, availability of software as a service or the provision of longer-term support mechanisms and ongoing maintenance of software through OMII is important for researchers whose work critically depends on it (cf. also IT3 and IT30).

Withdrawal of funding from services can have negative impacts on the adoption of e-Infrastructure services that can potentially be crippling and long-lasting: "*most of the services in the area where I work are being withdrawn at the moment, funding is stopping for them, for the AHDS, [indst.] and the Methods Network. That would be a problem for the people in the subject area in which I work*." (AH6).

Research computing groups do not tend to be **well-funded** compared to the potential need for training and support: "*But that's a lot of effort and, well we'd never have enough. At the moment we've got essentially one and a half people dedicated to that. We've got another two posts coming of which one post will be dedicated to that*." (IT49). Other respondents have commented that their research computing groups are mainly funded through research grants with only a small number of key posts being funded by institutions. This means that there is little room for providing support outside the boundaries of funded projects and a potential **conflict between research and service provision agendas**: "*on the same point as we are trying to be liaison between IT Services and the faculties, we are, at the same time, doing research projects*." (IT24).

This begs the question whether institutions alone can be relied on to provide the **funding** to create the environment for e-Research practices to flourish or whether funders need to finance programmes aimed at increasing uptake. As an intermediary put it: "*with repository development, for example, there may be a need to actually have some sort of way of actually providing additional resource into higher education institutions […] I think there might actually need to be some priming funding to accelerate the level of uptake and deposit […] at the moment there's quite a strong focus on research and development in relation to e-Research and data curation which is good but I think some of that needs to be redirected now to actually providing training materials and support [and] arranging meetings. I think to expect individual institutions to bear that burden in its entirety is probably unrealistic*." (IT32).

Funders increasingly demand that researchers **curate** their primary data but it is as yet unclear how the necessary storage mechanisms and curation procedures will be put in place and funded: "*one of the concerns that we have at the moment is the requirement coming out of the research councils for storing primary data, and how on earth we're going to meet that need and if it's going to be met you know centrally by JISC, or whether the University is going to have to make their own provision*." (IT17). The same respondent expressed a concern about the current state of practice: "*data will be held in an inappropriate way within the departments because they won't have the resource to store that kind of data properly, back it up, keep it safe. So they'll either be working on ad-hoc solutions in the departments or we're going to be asked to provide a huge amount of storage*" (IT17).

Another issue related to the issue of continuity of funding is the **state of software** produced by projects with time-limited funding (cf. section 4.2.1.2). More often than not, this software is not ready for general use and requires significant investment in time and skills on the side of users to actually deploy. OMII-UK has been funded to address some of these issues but it may be the case that they need to be addressed in a more fundamental way by changing the way that funding is provided for projects that develop software in order to improve their ability to develop more products rather than prototypes.

The **time and funding** to explore services was mentioned: "*It's a resourcing problem largely, we can only engage in that sort of collaboration when we've got the resources to do it and finding those is often difficult [...] it would be good to have, you know, some pilot funding to really spend some time with the Digital Curation Centre to run a pilot experiment to see when there are problems could be addressed by some of the method data tools*." (BB3). There does seem to be a lack of resources to facilitate the initial exploration of possible uses of e-Infrastructure services on the side of potential users. This also relates to issues described in section 4.1.1 on the lack of connection between initial outreach and specific training. A related problem is the **time** it can take to get access to high performance computing resources: "*the grant peer review process […] I think is the largest single inhibitor […] I think the selection is good and right and necessary, it is just the time associated with it that is the problem*." (IT8). While access to the NGS is not mediated by peer review, the process of gaining access can also take significant time: "*one of the problems our users find is it takes so long to get started on the NGS, and you know from when you go and chat to someone and say okay well your best move is to use say the National Grid Service and you have to get a certificate and do that and then you have to do the application process which takes at least a week*." (IT42).

Funding in general was, perhaps unsurprisingly, a recurring theme in the interviews. One aspect that was stressed particularly was the issue of getting funding for building up an infrastructure and critical capacity: "*in order to actually get time assigned to do the necessary work to set up the infrastructure here we would need to be part of a project or there would need to be some kind [of] funding for it and as such we don't have any such project*." (NE4). The use of national e-Infrastructure services often needs to be complemented by the development of local infrastructure (e.g., for data management and curation) as there is as yet no complete set of services that would cover all areas. As a consequence, usage of services can be dependent on funding for local infrastructure even if the national e-Infrastructure service is free at the point of use.

At the same time, existing **local service investments** can act as a barrier to innovation. For example, because NERC institutes have invested in HPC clusters, there is "*almost [a] disincentive to want to bother with grids*." (NE6). However, researchers recognise the benefits of linking local resources together, e.g., to provide overspill capacity (NE6), cf. also section **Error! Reference source not found.**.

A medical physicist (EP1) commented on the problem of **costing** compute services (also cf. IT12). Many such services are free at the point of use at the moment but change is underway to include them in full economic costing: "*at the moment effectively the cost of facilities are built in but I guess at some point this is going to change, e.g., currently dedicated large computer services are for free but there was a note on the [local] compute facilities recently indicating that using them is going to be costed, the note read any funding for proposals from now on should include the cost of using compute facilities. That's a chicken and egg thing, until we have explored the technology we don't know what technology can do for us, we are not going to buy technology, we are talking about research process here, before you get the big funding you need to do pilot studies which we have done here […]. Money and funding is potentially a barrier at some point what can you pay for, or if you put that cost in to your proposal it's likely not to get the funding. The trouble with the computing is that it is seen as hidden costs, if you add the cost of computing into your proposal it would lower your chance of getting the funding. Some of their funding bodies, like a national charity […] it is very hard to get from them the perspective that computing costs are real, if we said that we need a couple of PCs they would understand it but if we asked for computer cluster facilities they had difficulties to understand that computers are research equipment rather than just consumable pieces of kit*." (EP1).

4.1.8 Institutional

Institutional structures and arrangements for research support are of crucial important to the uptake of e-Research practices. The **availability of local IT support** was flagged a number of times as vital to the use of national services. Local support was considered by one researcher (ES3) as more vital from the support provided from a "disembodied third party", e.g., a helpdesk. Local provision can **vary**, however, with some researchers describing their environment as being very supportive while others being dismissive: "*it's not just my department, unfortunately, it's the entire faculty […], there's a support person […] but to be perfectly honest I never do [contact him] because I know as much about computing if not more than he does*." (AH3). Here, we can observe that the support arrangements available to researchers from different disciplines vary most widely but that this may also be influenced by the way IT support is structured within institutions.

The concept of **research computing support** as a distinct function in institutions is a relatively new and ill-defined one (IT44, IT50, IT51). As one respondent commented: "*it's difficult exactly to know what's going on in research computing. I mean how do you define research computing even?"* (IT51). The way that support for computing in research is embedded into the wider information systems function is crucial as it influences the awareness both of researchers and of other IT support staff: "*we should in theory be contacted via the University IT support. Unfortunately, at this time we have […] the problem that faculty IT support doesn't necessarily know about [us]*" (IT24). The relationship between the overall information services function and research support has also been discussed by another respondent: "*we're also very user focussed so in thinking about information management, we think about the information that's required for researchers and that isn't just the systems that they use […] I think the main disadvantage of doing it this way is that it becomes isolated or can become isolated from the information systems strategy […] if they're developing something in support of a research system, we wouldn't necessarily know about it, even though we should know about it and we have to be proactive in finding out what's going on in order to make sure that we're providing the best environment that allows people to get at all of these things*." (IT50).

While some institutions provide explicit and active support for researchers, others lack these facilities. A respondent from a traditional research-intensive institution described the very **active user engagement** arrangements in place: "*information services has […] academic liaison directors whose task it is to speak to the users and their colleges. [There are] monthly stakeholder meetings, and [liaison staff] go out and meet with the research groups*." (IT8 also cf. IT48). Clearly, active user engagement by research computing services has the potential to increase the awareness of e-Research practices and provide the support necessary to overcome many of the issues outlined in this report. The inclusion of research computing support in **institutional strategic planning** can help to ensure consistency of support and service provision as well as to facilitate the exchange of experience, knowledge and skills. The organisational context can influence the success of such efforts. As one respondent put it: "*You might find that there are cultural differences with different institutions. Institutions like the [institution] have people from all over the world who have experience of many different approaches, different ways of doing things. So you might find that there's a richer approach to using different approaches here than maybe a University that has more of a local market or a local research establishment. However the [institution] may also be fragmented in the way it approaches things, in which case you might find one group doing something very good and another not aware*." (IT20).

A respondent working in a big biomedical research institute said they did all their computing in-house and were not using the NGS because of the **overheads** involved in engaging with it but they added: "*I think we may look to share facilities between departments here as a way of increasing power and keeping things locally managed. That would be my guess for the next five years at least*." (BB2). This may point to an opportunity for the organic growth of **national compute resources** with some larger users eventually becoming resource providers (it is interesting to note how concisely the respondent captures the rationale for the development of grids). In the short term, there may be a need to provide advice on setting up inter-departmental grids and in setting them up in a way that can lead to their eventual integration with campus grids and wider grid infrastructures.

Later in the interview, the same respondent reflected on the provision of compute resources: "*clusters is something and again that maybe something where a grid type system which means cluster was managed elsewhere and you just access them remotely all the time, you know, that might be a good way forward but not something I have great experience in*." (BB2). Clearly, again, they understand the principles of and motivation for using grid-based compute resources but are **not ready yet** to investigate this further.

Sharing of resources is seen as a good approach by many researchers but there can be potential **barriers at the departmental level**: "*I looked […] to see where if you could offer [a local compute resource] as part of an NGS resource and it was too difficult for me to get that put on and go via that way to submitting jobs on to it, because the way I understood it is that we are all put our machines on it and then the NGS would be a much bigger computer, so I looked to try to see whether the resources that we had we could stick on it and actually there were various political barriers to doing that [… Parts of the University] harbour resources basically and stick up barriers towards their open use, and I guess that's probably true about any University, I don't know*." (ES3). Clearly, unless there is support for resource sharing by departments and institutional research support services, initiatives by individual researchers will be frustrated.

Other respondents commented on the policy that their institution had for **charging external users** for use of the Access Grid. They also commented that they had problems using their local node because of **inadequate booking procedures**, adding that the were not sure all bookings justified the use of an Access Grid Node: "*There are one or two people using the access grid for interfacing in the research* *context rather than a meeting context, one of the problem is relating to the priority ordering, what kind of meeting and what kind interchanges being done on the grid […] it's a local issue of scheduling but probably it's an issue about spreading the technology more*." (ES1)

One NERC researcher commented that national compute resources used to provide them with the largest compute capacity they could access but that this was no longer the case: "*until recently the national services provided the most computer power that we could get our hands on but that's no longer the case, so that's providing a shift from national resources to local resources […] they're not quite so large but we can get the larger share […] we are in the fortunate position of a good local resource*." (NE5). Whether or not the problem of getting a **large enough allocation** of national resources is relevant for researchers will depend on the nature of their computational problems. As the researcher interviewed mentioned HPCx and HECTOR, we may assume that they were mostly concerned about larger-scale, closely coupled models as opposed to throughput computing.

Availability of the Access Grid service is a problem for some researchers: "*We do have a nearby Access Grid node but we now use a commercial videoconferencing solution […] most of the places that we want to videoconference with are not physics laboratories but hospitals and the like and they don't have equivalent facilities*." (BB2). Other respondents also reported that they use alternatives because some of their partners did not have Access Grid facilities (BB5, MR7). Clearly, communication relies on all parties having access to the same technologies and on a sufficient networking infrastructure. Since the successful use of Access Grid is closely tied to the availability of local support, the mere availability of an AG node is not sufficient (cf. sections **Error! Reference source not found.** and 4.2.1.2); the quality of the installation and the support arrangements plays a crucial role.

A medical researcher raised the problem of a **lack of access to resources for NHS research staff** working on collaborative projects with universities (MR7). This issue is usually addressed through the use of honorary contracts but these can be difficult and time consuming to arrange. Also, even where access is guaranteed through an honorary contract, researchers sometimes still have the problem of accessing resources from the NHS network: *"all sorts of services don't work across the firewalls between the NHS and the university and so we have repeatedly had to do things in different ways and build things in different ways*." (MR7). The NHS HE Forum is working to try and address these issues by working towards adequate organisational and technical arrangements that allow more seamless access across the NHS/HE boundary (MR7) but from our own experience we can say that the focus of their work is on teaching, not research. We should also note that this problem exists also for those working in other research institutes outside the UK higher education sector. While the solution to the problem of access may be a technical one, the problems of implementing solutions are mainly to be found in the inter-organisational politics between NHS Trusts and HEIs.

4.2 Technical Issues

Our interviews uncovered less technical issues but this may be because our interview process was not designed to tease out detail on technical aspects. Comments were categorised under 5 different headings: infrastructure, security, rate of change, cost of adoption and scale.

4.2.1 Infrastructure

4.2.1.1 Operating Environment/Reliability

There were mixed views about the **reliability** observed in Access Grid sessions. Some respondents found Access Grid to be beset by problems: "*it failed more often than would be ideal it means that you can't really rely on it if you have to have an important meeting*." (NE2, NE5). Others reported that the Access Grid sessions they had been part of worked mostly without problems (EP3, EP5) although they may be using other technologies where these work better – for any of a number of reasons (AG not being available to partners, need to conference from offices, ad-hoc meetings, etc.).

Clearly, the availability of **easy to use** desktop conferencing tools such as Skype, Adobe Connect or Conference XP raises the expectations of Access Grid users who feel that they would like to see less overhead associated with running Access Grid meetings: "*thing is, they don't always want to or need to go to a big room, you know, and hire it out and find a technician who can, you know, support them but it would have to be very easy, you know, sort of Skype download it and it's ready and off you go.*" (AH7). Another researcher referred to the lack of support related to technical aspects when there are issues of access, interoperability of tools and format (AH6).

Network problems were frequently referred to in relation to the use of the Access Grid. As an experienced user of Access Grid explained: "*The problems I get are often that the bridges are either overworked or don't work and setting up multicast on a network is not trivial so we have, the only thing we use actually from provided by Janet in this respect is the multicast beacon to check the multicast connectivity, I think it's a Janet service, […] if there's any sort of network average for any length of time it gets fixed because network admin see the network's down and they fix it otherwise they get a lot of complaints, whereas if multicast connectivity goes down for any reason they probably don't notice it and probably nobody notices until you try and connect to your Access Grid meeting and realise that oh it's not working. Inevitably it's because somewhere there's a router that has fallen over, it's been brought back up but multicast hasn't been turned on […] it's actually an institutional problem so we end up not having a multicast somewhere at the university*." (NE2).

The potential for such network trouble to affect meetings means that a certain percentage of meetings cannot happen: "*I guess this means that something of the order 10%, 5% of meetings don't happen because, you know, it works every week for 6 weeks and then the seventh week you turn it on and you can't see anybody and that sort of stuff […] I mean the problem that you were into is sort of scaling to more people. For smaller meetings we've moved entirely over to, say, 2, 3, 4 people over to using instant messaging and video-enabled instant messaging so just, you know, on a Mac iChat or equivalent*." (NE2). Clearly, technologies that do not use **multicast functionality** do not provide the scalability of Access Grid but work more reliable for smaller meetings. They are therefore preferred for some types of meetings involving fewer participants.

Respondents also referred to the set-up time required to check the workings of a set of Access Grid nodes before a meeting: "*There were various issues when the bridge would go down or there was a certain amount of instability in the system but generally it works fairly well [...] Access Grid meetings would start off with about ten, fifteen minutes of people struggling with the technology*." (NE4).

In relation to these problems, researchers expressed the desire for easier personal access to an Access Grid Node, preferably from one's PC with a webcam (ES6). In general, they would like JISC support to make more use of the web as a medium for video-based inter-personal communication (ES1).

4.2.1.2 Support

Documentation was one element that was presented as a concern as "*it is quite hard for new users to start up using the software tools*." (BB6). In addition to this, **support** following training days and documentation was regarded as necessary because researchers might feel helpless once they return to their workplace: "*they do offer the training days which are really quite good for hands on help and advice and showcase of the software and how to use it you know, and what you know example tutorials of how to build the work flow they're quite good but it's just once you get back to your laboratory and you're stuck on your own*." (BB6). This latter point also relates to the discussion in section 4.1.1 on the need to link up engagement at different stages. Given the importance of documentation, the experience of researchers using it did not always match expectations: "*essentially it was through searching for handouts and searching for presentations that were scattered about over various websites*." (ES8). Having a clearly identifiable set of **easily accessible documents** would clearly be a key enabler that is currently missing.

Advanced IT support for specific services was again a prominent issue. As mentioned in section 4.1.1, successful uptake requires the filling of gaps between initial outreach, training and specific support through arrangements that help researchers to work out the details of how they might be using services. **Co-location** with support and development teams was presented as an important factor in respondents' successful use of e-Infrastructure services. Often, researchers referred to the fact that they were collaborating with the providers of services or tools and that this closeness and the exchange of expertise it enabled was very important for them (e.g., BB5, BB6). As a respondent put it: "*Barriers aren't a problem if you can go and knock on the door of the appropriate people. So, I don't know quite how accidental it is that most of the JISC stuff we use is the stuff where there are people at [institution]*." (BB5). When asked about the reasons for collaborating with a service provider, one respondent commented: "I mean it was because of their expertise I think as much as their services. I mean this is a research project and we see them as a, you know, as an important friend to be involved with and they did seem to have particularly interesting technology so I suppose you could say it was because of their services as well." (BB3). Another respondent remarked: "*they can learn from working with us and whether it be a use case or a bit of code that goes into the core code or what-have-you, there's benefits to […] we're all very close to the collaborators [DCC, NaCTeM, NGS and OMII] and I get something out of it as well which is a product either for my service or my community, but I think that's quite important thing to come out to be honest […] it's the fact that those centres exist is the reason, you know, I can go to those centres of expertise and help me with my problems and hopefully feed something back*." (EP6).

Support request **turn-around time** was mentioned by one respondent as an issue: "*there's just little things like, for example, it's very outdated versions of Python that are installed and we put a support request in ages ago and nothing ever happened so but that was also round about the time that the NGS2 migration was starting to happen as well so I suspect that things kind of got lost in the mix in there somewhere*." (EP5). Clearly, service provision can be adversely affected by transitions in the service arrangements, especially where 'services' are funded on a project grant basis. Another respondent reported good experiences with support through mailing list but said that being able to attend regular meetings provided a more rapid response (BB6).

Local support can be essential in mediating the engagement with and use of e-Infrastructure services as one intermediary pointed out: "*you know it can be hard to persuade the NGS to give you a 700 gigabyte […] to actually get the data there. So I think you have to know the right people and sort of ask them very nicely and sort of* *you know send an email as someone who knows about these things to the NGS people and say yes actually this is […] a serious request and they do need this amount of data and actually it's really only […] temporarily for them so they can actually get the data on there*." (IT42). The potential role described here contains an element of managing expectations on the part of the researchers by pointing out to them that the NGS does not provide long-term storage for data and at the same time a degree or assurance for the service provider that the resources requested match a genuine requirement that has been assessed by someone with sufficient expertise.

There is a general impression that people are resourceful in getting the support they need and that they are becoming more self-reliant over time. Some may even **actively contribute** to the development of e-Infrastructure services and tools **through reporting bugs and stating requirements**. As one social scientist put it: "*in the case of individual surveys, there is sometimes the case that I have downloaded a file and then found out perhaps an error within the data file or lack of priority on a particular variable or a variable missing within the data file that's refer to in the documentation and occasionally I have reported that sort of example to the help desk, and they've usually been, well in fact pretty much every time they've been able to get back and come up with the solution quite rapidly so I find them a very helpful service in that respect*." (ES7).

As one of our respondents from the group of intermediaries pointed out, there is a balance to be struck in terms of the level of **guidance** given to researchers: "*we know that some people use [Condor] badly but when we find out we talk to them and say, look, you need to understand a little bit more how this works […] if you restructure your work like this you'll get a much bigger benefit […] it's a tension between if you wanted to guarantee that your users used it perfectly you'd have to hand-hold them all the time and they would find that unbearably restrictive and wouldn't use it, so that would be stupid*" (IT49).

The Access Grid Support Centre (AGSC) appeared to be very helpful in facilitating academics' use of the Access Grid. Because of the nature of the service, respondents' comments were mostly about their experiences as users of Access Grid rather than as users of the support service. Researchers are not necessarily **aware** of the role AGSC plays in the running of an AG meeting: "*I use Access Grid quite a lot, I got training through my local node […] and I haven't had anything to do, I don't think, with the Access Grid Support Centre – that might be a lie, maybe when you are actually setting up a meeting you have to go through that. I don't know because I just use the system that I know how to use locally*." (AH1). This was confirmed by a number of respondents (e.g., BB5) and may not be an issue in general but rather a sign of success for AGSC, given its brief of being mainly involved in setting up nodes and providing initial information about running and facilitating meetings.

It does, however, mean that use of the Access Grid is crucially dependent on **local support**, which consequently was a recurring topic (AH1, AH6, NE2, NE3, NE4, NE5, NE8). However, this support was not always ready to hand: "*I kind of hassled them until they taught me how to use it, there wasn't anything organized […] I was the first person from the Arts and Humanities at [institution] to actually use Access Grid, so as a result they showed me how to use it and trained one of my PhD students up as well*." (AH1). Other researchers reported that their institutions provided regular practice and training sessions (AH6), so the picture is at least somewhat mixed in this respect. The support provided by AGSC was viewed as very satisfactory by all participants who had been in contact with it. One of them specifically said: "*the support provided by the AGSC has been good enough to be invisible – I've wanted a meeting, it's happened*." (BB5).

However, **lack of support for multicasts** was mentioned as a barrier (NE4, NE2). One researcher explained that the problems arise because not too many people use multi-casts and the administrators don't notice that it has failed so they haven't fixed it: "*I'm sure that the network administrators would find a way of making it work and fixing it when it breaks but my impression is that because nobody notices when it breaks and there's still a people use multicast, nobody notices when it breaks and it takes you to find out. You are the person who finds out it failed more often because you're one of the few people who use it*." (NE2). Clearly, active monitoring of multicast functionality can help but IT services do not always provide this.

It is perhaps ironic, given that Access Grid was developed for large-scale meetings, that it suffers from the problem that the likelihood of problems occurring and affecting the whole meeting increases with the number of participant sites: "*a large number of sites, more than 10, probably about 15 sites joining working at the same time and that's the real problem. Yeah, so I think our main problem is it's unreliable when it comes to big meetings […] if we have a choice we don't use Access Grid for any big meetings*" (NE1).

Software support was raised as a concern a number of times**.** An engineering researcher was referring to issues they were facing with GridSAM and commented the state of support for this software: "*unfortunately we found that GridSAM has some problems and unfortunately GridSAM is no longer actively developed […] Depending on what happens over the next 3 or 4 months, we'll probably either start coming up with […] solutions that work better for our purposes or try and fix GridSAM ourselves I would guess. I mean [OMII has] been trying very hard but obviously it's very difficult because the developers are employed on something completely different so it's quite hard to get the support for a lot of the OMII software although they are trying hard to improve that. Unfortunately most of the software that's in the OMII stack isn't actually finished, that's the main sort of problem. I mean I think OMII is a very good idea but they're in a very difficult position because obviously most of the time a research project is funded for 18 months, 2 years 3 years maybe, it produces results, it demonstrates a technology but the, you know, the EPSRC in the UK were not very good at funding software development projects*." (EP5).

4.2.1.3 Usability

A number of interviewees, including both researchers and intermediaries commented on the usability of middleware and other infrastructural components. Top of the list were issues to do with the acquisition and handling of certificates, which we will discuss in section 4.2.4. In general, there are questions about the effort spent on either training on the one hand or the development of simplified user interfaces that can help to increase uptake: "*if you want a wider benefit then you would have to make it easier for them to use not just keep on training them because you have to not just consider the cost of training but also the cost of them not being engaged for a long time with their own stuff because they're being trained to use this, and of course e-Infrastructure will change again and then they have to be re-trained. So what you want you want to do is make it so easy for them to us that you don't need training*." (IT10).

Another important issue that affects researchers' perception of the usability of a service is what happens in the event of a **service failure**: "*When something goes wrong it's difficult to find out what's going wrong if it's an application that's fired off as a pure grid application in a sort of fire and forget type of thing which was in a sense what the project was supposed to be, if it folded somewhere along the line it was impossible to really find out what was going on*." (ES8). Clearly, while it may be desirable to hide details of a job execution in any normal situation, when a job fails it becomes necessary for the user to be able to investigate what has gone wrong. It may well be that the best thing to do is not to add complexity to the user's interface with a service but to improve their ability to get help in such situations by better enabling 'post-mortem' analyses of specific jobs by support staff.

A number of issues were raised by another respondent that are related to the configuration and quality assurance of NGS nodes. They mentioned the problem of not having the **same environment** on different NGS nodes and therefore having to compile code repeatedly for different nodes: "*So, the problems we faced from the user's point of view, I guess, the biggest problem is that the machine in terms of libraries and compiler versions aren't kept in sync. So it's not – MPI libraries and so on – so it's not always possible to compile on one machine, especially MPI executable, and then take it to another of the NGS machines and run it and this is a significant problem because it means that […] where you use multiple machines you have to long into each one in turn and compile your code which takes time and the more grid resources you have the more time that takes*." (NE2).

Specific issues to do with the way that the NGS middleware stack handles job submission and management were raised by an engineering researcher: "*We don't actually use the NGS compute resource ourselves […] predominantly because the best method for us to actually utilise the NGS resources is to log into the head node and do a queue sub which kind of defeats the purpose really […] we need to monitor a hundred thousand jobs, that's practically impossible using Globus because Globus just doesn't want to work like that. Although [another resource] is actually NGS compliant, we have a secondary submission method there which is via Ganga [which] gives a bit more flexibility in terms of submission and monitoring although there is still a Globus back-end on it at some point. Ganga takes a lot of pain out of managing the jobs*." (EP5). Problems with **submitting many jobs** to the NGS were mentioned by another respondent (NE2), who commented that quality assurance did not seem to cover such scenarios: "*testing of these clusters pretty much comes out to can you run a single Globus job, oh yes you can, that means it works*." (NE2). We already mentioned the related problem of tracking failed jobs above.

There were difficulties also with using different types of Access Grid nodes (like IOCOM versus Access Grid Toolkit) where respondents had problems connecting with other types (NE1). One researcher suggested aiming to develop **interoperability** with other communications technologies: "*if somebody's in their office and could actually connect using a different client – Skype […] iChat anything that's sort of a video capable client, that would be very useful*." (EP5).

4.2.1.4 Matching Research Needs to e-Infrastructure Provision

The match of infrastructure provision with the needs of researchers was commented on a number of times. For example, the match between **computational requirements and resource provision** was mentioned a number of times: "my impression was it was quite slow in getting the job finished so if you're running lots of small jobs the overhead in submitting those small jobs is larger than often running them" (NE8). Clearly, some problems do not map well onto the grid paradigm.

On the other hand, the provision of relatively expensive shared memory multiprocessors can be dwarfed in terms of sheer computational power provided by clusters or desktop grids and smaller institutions can face the problem that the kind of HPC resource they can afford can become obsolete quite soon: "*when we've looked at the power of desktops, actually utilising some of this on unused desktops is almost as fast now […] it is not a sort of large clustered service, that […] dwarfs everything else we have*." (IT20). Clearly, for resource provision to be viable it has to provide a service that provides more than an incremental improvement over what can be provided locally.

4.2.1.5 Application Programming Interfaces

Another issue concerns the **availability of application programming interfaces** to resources. One respondent stated that they could not use the NGS because it did not allow them to build applications on top of it: *"[…] the other reason is [...] what I would like to do is build applications that use the National Grid Service underneath and the NGS has the wrong kind of interface at the moment so it's difficult for me to build those applications […] it'd be useful if the interface was a service, at the moment it's a log-on I have to long on to the system and type my own commands*." (NE1). This comment points to the need to enable usage of e-Infrastructure services from researchers' own systems or services. While the idea of such usage is very consonant with the spirit of e-Research and distributed computing, there is currently a lack of adequate interfaces. The NGS is currently developing a GridSAM-based endpoint for job submission but this is provided only as a pre-production service.

4.2.2 Security and Access

There were several issues mentioned concerning the need for **registration** in order to access services, the **lack of standardised** rules and systems to access services**,** the **lack of group access to services** for teaching purposes and the **complexity of acquiring and managing certificates.**

More specifically, one of the interviewees underlined that registration is a barrier not only for them but for many other users: "*I know there has been a barrier for me and also for a lot of users which is the requirement that you register before you can download anything from the Data Archive and I understand why that is but it can be very difficult and very off putting and a lot of people just give up before they download stuff*." (AH3). They point out that sometimes there are alternative resources providing a similar service that do not require registration and are therefore preferred by researchers (also *cf*. NE4).

Other researchers commented negatively on the procedure for acquiring a **UK e-Science certificate**. They suggested that "*there are many places where the security gets in your way and what it does it puts people off getting involved. It is easier to use a computer at your university that is free and easy to access*." (EP9). On the other hand, researchers complained about the need of using **different passwords** for different services (EP2).

One respondent reported difficulties while trying to teach her students about Edina and UK Data Archive as they had to have their Athens password and not all of them had it: "*students have to go to get their own password from the library. It's one more step they have to do to get access to resources. It depends where they are – a lot of students want to use […] from home, so they have to have the Athens password from there and it would be quite nice sometimes if they [they could go] straight into the objects*." (AH4). Obviously, the move towards **Shibboleth-based authentication** ought to address this issue and others related to the more general problem of VO management.

In addition, each service has different **access rules and procedures**, so that there are several different routes which it is necessary to take to access services and this may lead to researchers giving up trying to access the information through this route, reverting to web searches or use of other systems instead: "*It sometimes takes more time to try to figure out how to access the information as it's presented there than it does to actually just, you know, the worst comes to worst I just do a search on Google and I find where the papers are*." (AH8).

4.2.3 Rate of change

The rate of technological change that researchers experience can be a problem in that it can leave them and those who support them trying to catch up with developments. The technical possibilities often evolve faster than the ability by institutions to adapt to them and support them. One respondent identified an inter-generational gap as a potential cause of these difficulties: "what gets in the way of our strategy, there's also the cultural thing that the committees of the University tend to be run by the more senior academics who all date back to the days before computers took over the world, and so talking to their research students and their junior post-docs is practically speaking a different language compared with speaking to the professors" (IT3). Another respondent commented on the problem of technical development outpacing their ability to provide training: "things were moving so quickly it was very hard for people to keep up, so even in the training context it's very hard for people to keep up, you know, with what they're training and what people are learning" (IT14). A related issue is that training is often provided for what is perceived to be the dominant technology or version of a technology, whereas researchers often have reason to adopt specific solutions: "*our goal all along has been to do generic middleware that could be used by many different people so we've always fought very hard against that but the users want very specific versions*." (IT14).

On the other hand, one NERC researcher commented on their perception that the development of services was too slow. Specifically, they mentioned that AG has not been developed as fast as it could and is being overcome by other emerging technologies such as EVO (N3).

4.2.4 Costs of Adoption

There are **direct** as well as **indirect costs** involved in adopting e-Infrastructures that need to be weighed against the possible benefits. Both sides of the equation are often determined by technical factors. For example, one respondent explained: "*It's the technical aspect of running the models I'm mainly thinking about, so there are barriers to grid use that are related to how much effort's required to actually adapt your scripts for example, to running on the grid and if it's too much effort then there's a tendency just to stick with what you've got*." (NE6). The costs of adoption are not only relevant for researchers but also for those supporting them: "*the particular technology that you need to deploy to use the Grid […] has a stiff learning curve so we need to learn about and train people so it's not a straightforward thing yet for us at least*." (IT54).

4.2.5 Scale and Complexity

The increasing **scale of computational models** is one factor that has the potential to drive forward the use of e-Infrastructures even in disciplines that can afford relatively large local resources. For example, in environmental modelling, there is a "*need [for] more than just the NERC clusters […] some of the regions that were going to be modelled in parallel can be run on smaller clusters like ours for example at […] but there is a need to actually run on Hector and HPCx for some of the groups of regions that require much closer coupling between them during the simulation [...] I've also worked with some scientists on Condor, trying to actually help me to use Condor on the [institution] Campus grid*." (NE6). While some models can be effectively sub-divided and run on multiple clusters or even collections of workstations, others require large-scale resources that offer better support for closely coupled computational processes. Computational grids may help to provide a way to use different kinds of resources transparently.

e-Infrastructures need to be **scalable** along a number of different dimensions, both in terms of the technical infrastructure and in terms of the human and organisation infrastructure into which they are embedded. It needs to cope with increasing "*complexity and volume, number of users [and] number of interactions they have with each other*." (IT10). Clearly, running a single supercomputer for a limited number of users is different from running a comparatively modest grid infrastructure that comprises resources from different organisations and is used by hundreds or thousands of users from different organisations. In addition to the increased scale of computational models, which makes more sophisticated forms of computation and data management necessary, their increasing complexity adds to the need for improved **collaborative practices** between different stakeholders to ensure adequate resource usage and the effectiveness of e-Research methods to produce valid results.

4.3 Applications

4.3.1 License Management and Costs

Commercial software and other digital resources are often distributed under license agreements entered into by individuals, departments or institutions. The costs associated with software licenses can be significant if they cannot be split amongst a larger number of parties. "*Software is the other one; we do use CHEST where possible, although I do know that not all of the software suppliers work through CHEST. Some of the MatLab software is potentially of great interest to some of our researchers but the way it's currently licensed we can't really afford it. Perhaps at a national level better deals could be struck with the suppliers of some of that sort of software*." (IT26). "*MatLab is available in [the] college but they have to pay for their own licenses… each department is responsible for its own licenses at present*." (IT18). Scaling up the licensing arrangements to the national level would potentially allow the UK's HEIs to make savings and to widen access to specialist software packages.

In addition to the benefits of cheaper licenses, the central management of license arrangements is also cheaper to administer: "*The second strand is that our user community have a range of packages and commercial software and it's more cost effective for us to manage licenses to those packages centrally. So we provide the applications as a service for our user community*." (IT33). Similar issues arise with respect to the licensing of e-Journals and other digital resources (e.g., IT3).

4.3.2 Availability

Researchers often need access to **specialised software packages** that may not be generally available through routine software distribution mechanisms in institutions and even where licenses exist there may be problems as researchers may be unaware of their existence or they may not exist in the operating environment used. As licenses are often acquired by departments, central services can struggle to provide effective support, as one respondent commented: "*I think there have been a few issues with the software for [...] GIS type stuff, that sort of mapping services and things. I think there have been some requests for some extra software but I believe that there is some software available and it's on the [local] computers but they also have it in the Geography Department, where they know about it but that's one of the things we do want to sort of catch up on and make sure that we're supporting it as much as we can*." (IT51).

It is sometimes not immediately obvious what **potential uses** software products may have, so support may be limited to departments that are their obvious users. One intermediary commented on the use of Mathematica: "*I assumed it was just used in the maths department and then I discovered it was used very much in the biomedical and health department […] in modelling, protein modelling and genetic stuff and* *things like that so there's a lot of sophisticated stuff going on among our academic staff using IT in various ways and sometimes in places that I don't immediately associate with it*." (IT51). Similarly, the respondent commented on the use of advanced ICTs in the Arts and Humanities, which are often taken to be late adopters or even disinterested. Clearly, gaining a good understanding of what particular applications are potentially of interest to different disciplines is a prerequisite for effectively making them available.

4.3.3 Cost of Implementation and Maintenance

The **initial cost** of implementing applications is often covered by time-limited research grants. The level of completeness, usability and quality of what can be achieved given these arrangements is often limited. While initial development can often produce applications that are mature enough to be adopted by researchers, the initial use often gives rise to further requirements and to a need for ongoing maintenance that is often not sustainable: "*It's difficult because a research group […] has a number of post docs and PhD students and it's not a software engineering company. So when you finally deliver something it's going to be a prototype it's never going to be a finished product. But they then start depending on it and then the problem arises because you don't have the funding or even the capacity to maintain [it …] whenever they want to change something, and they don't have the money to do the same thing because they can't just say oh here's an extra amount of money just to add this feature that doesn't exist so it's going to be very limited*." (IT10).

A compounding factor is that the applications are often based on other **digital artefacts** that have been produced through projects that themselves may have an uncertain future: "*So you're working with products that come out of research rather than out of a software factory, and often these will have problems or they'll be half finished or they don't really fit together yet […] it's difficult to decipher what the risk is before you start*." (IT10). This lack of a dependable basis increases the initial costs of developing applications as well as subsequent maintenance. It also increases the risks that maintaining and sustaining the application will not be feasible.

The risks associated with application development can be mitigated by carefully managing the dependencies between codes used and adopting simple, lightweight solutions where possible. The fact that many of the codes used are open source can further help to mitigate the issues, especially when a solid model for managing contributions exists for the components used. Organisations like OMII-UK provide a certain level of assurance as they provide a framework within which software products can continue to mature.

4.4 State of the Art

The advent of multi-core CPUs and the increased availability of multi-CPU systems and clusters is bringing issues concerning the **exploitation of parallelism** to the fore. As one respondent commented: "*what do I see as the biggest problem in research computing at the moment […] it's the end of Moore's law […] the approach of going down the multi-core route means that if you're going to solve the computational problems of a decade's time […] you're going to be looking at 10,000 or 100,000 processors and frankly if you look at the code base they don't scale so the numerical tools for example that we have today need a thorough overhaul and that's an enormous task*." (IT8). The lack of software making good use of compute resources results in poor utilisation: "*when we analyse what's running on their HPC cluster, approximately half the jobs don't appear to be making good use of it*." (IT53).

At the same time, the move to interdisciplinarity in the study of complex systems means that **computational modelling practices are becoming more complex**: "*if* *you're designing an aircraft wing, in the past we've split it up in a way that you'd have one set of people or one set of codes looking at the structural integrity of the wing and another set of people and another set of codes looking at the airflow over the wing whereas actually you know that you want to combine them because different airflows could produce different strains on the wing and therefore different issues for structural integrity and so on and so forth so there is a lot more thrust towards inter-disciplinary working that we have to drag together and that just makes the computational task much harder*." (IT8).

Verifyability of computational models and simulations is an important issue that needs to be addressed. While in some areas results can be verified experimentally, in others this is not possible, so the question of trust arises: "*It is very discipline specific but I don't believe that most people actually understand how big a problem this is*." (IT8). The respondent points out the experiences of the particle physics community, who have put much effort into validating their computational models while other disciplines have not spent anywhere near the same effort. **Trust in computational methods** can either be established through verification by observation or experiment, by long-term experience and comparison between different computational models or through more formal verification and validation: "*what we would like to do is take verification to a more fundamental level and to do it not simply by comparison with observation but also take a sort of mathematical theoretical approach to verify the algorithms*." (IT8).

4.5 Digital Resources

Respondents raised a range of predictable but nonetheless important concerns about digital resources, most commonly in relation to research data but also including resources such as learning materials. Broadly speaking, these collectively point to a number of perceived gaps in digital resource infrastructure provision. The following six categories were the most prominent ones: discovery, sharing/re-use, curation and legal/ethical.

4.5.1 Data Management

The issue of management of data within groups of researchers and with a view to its **long-term usage** has been flagged up by a number of respondents from information services. Information services and research computing groups are often aware of needs but can find it difficult to respond to them as it is difficult to define generic functionality beyond the level of raw storage capacity, which many institutions provide in the form of Storage Area Networks: "*We have demand from groups particularly in Medicine saying that they have data that they want to store. We can provide them with raw volume, but we would like to provide them with more than that. As I say advice on curation and all the rest of it, and that is still at the planning stage […] I suppose the problem is that it's a complicated area, different people have data which is structured differently, and I suppose we're grappling with whether you can give generic advice or if it's got to be disciplined specific or if it's got to be […] project specific.*" (IT11). Another respondent commented that they are "*preparing a briefing document on our research support […] one of the research groups […] in Neuro-Science did come to [us] and they were asking us about data storage and whether it was possible to save their files to a central place and have a centralised backup and things like that and that was quite some years ago now and I really think that is one of the prime needs that people are aware of*." (IT51).

Issues of **data management and curation** were mostly raised by intermediaries and often seemed to be related to strategic developments within institutions. For example, one respondent commented: "*issues surrounding research data management such as backup and long-term storage have come to the fore recently* *following a report from the University's internal auditors*." (IT44). Intermediaries also reported on requests by researchers for general data management functionality (e.g., IT45) while researchers tended to mention more specific issues they were facing such as access restrictions, storage capacity or performance.

Collaboration between different partners involved in different stages of a research process is an important aspect that, when combined with the exchange of **non-trivial amounts of data** needs to be supported using e-Research technologies: "*But it's quite a distributed process so one lab might be interested in a protein and do all the biological assays on it. One lab might make it and […] then the synchrotrons such as Diamond may be where the data are collected, which are a few tens of gigabytes, and then we would go back to maybe yet another lab to be processed and the structure produced. So it's very distributed work and so the messaging and secure messaging and looking after the image data that's all important to us*." (BB2).

4.5.2 Discovery

In the idealised research process, the data collection phase marks the beginning of the data lifecycle. For many researchers, however, data collection is complemented or even substituted by the discovery phase where a search is conducted to see if relevant datasets already exist. For example, repositories such as the UK Data Archive, Mimas and EDINA provide a vast range of datasets to the social science research community. Respondents confirmed that these services are perceived as being very valuable but that they also feel the **discovery process is not always reliable**. In a number of cases, it was the quality of metadata that was a key concern: "*it is quite difficult to find all the data that exist […] There is metadata there to be sure but you cannot query it in a way we want, that would facilitate the research, because it's a laborious part of the research which is not that exciting. So a better metadata and better ways of searching the metadata is what's needed there*." (ES6).

Part of the problem is seen by respondents as stemming from a **lack of agreement over metadata standards** within research communities: "*people do not use controlled vocabularies, and ontology, that also causes difficulties sharing data, because meanings of the terms used is often are different from individual to individual or even the same individual on different trials, they may use the same word to mean different things. That was the biggest problem we identified*." (MR5).

4.5.3 Integration

Integrating different digital resources is often not a simple exercise as it raises technical as well as methodological problems. As one of our respondents observed: "*you have to get the right methodologies for the right data but also sometimes the data is distributed out from multiple organisations and there might be different types of data that you want to bring together and you want to integrate first and then when you've integrated it you can finally do the actual extraction of knowledge […] making sure that that process becomes easier is an important challenge*." (IT8). The respondent pointed out that data integration often requires expert knowledge and is often done on a one-off basis. Providing **integration solutions** that work over time is important in order to achieve benefits of scale and re-use. The problem of **orchestrating** data integration processes is not solved in general although practical solutions exist that may work sufficiently well in particular areas: "*orchestrating the whole thing when you conceptually have it written down that's the next challenge then and that's what's still very hard to do so you need tools to make it easy […] Taverna of course is a typical tool for a particular discipline*." (IT10).

Integration solutions are often **complex socio-technical arrangements** involving collaborations between different stakeholders such as data providers, infrastructure providers, data integration specialists and researchers: "*all of these somehow need* *to interact or help you interact with particular bits of these processes*." (IT8). Visualisation mechanisms and collaboration support can help to make it easier to make sense of complex relationships, of the provenance of the data and the data itself.

4.5.4 Sharing and Re-use

Even when suitable datasets have been discovered or are known to exist, where re-use entails sharing with other parties, researchers may encounter a number of barriers which inhibit sharing and re-use. For example, researchers often lack **robust mechanisms for sharing data** across organisational boundaries.

4.5.4.1 Legal and Ethical Issues

A number of researchers commented on the **legal and ethical issues** in relation to conditions applied for the access and use data and on various ways in which these might inhibit the research process.

In many fields, the sharing of data is subject to policies which are designed to protect IPR (e.g., where commercial collaborators are involved). In some cases, these policies were seen as being **too restrictive**: "*we had lots of licensing issues and that so one of the ideas was that you could use other people's data but not necessarily download it, its licence agreement kind of allowed it. And then you could run your model regardless […] you could just get the results from the model and that sort of thing so there was a lot of discussions on how to deal with that as well. We got a system in place eventually but the idea was that outside users could come into project and use it in that way but if it was allowed that the partners would share the data to make it easier to run on their machine if it was there, you know, part of [the] deal*." (NE7).

In other cases raised by respondents, **licensing policies** are still in their formative stages. This posed a problem for medical researchers in particular: "*Where there are intellectual property concerns there are more restrictions, we have started working with the industry […] to try and establish how much data can be in pre-competitive area so they would be happy to share it, so we trying to find out what is commercially valuable so they can keep that secret and let that be adequately be protected and release the rest […] there is some understanding in the industry that we are in a reasonable position, that's at the very early stage*." (MR5). "*We have spent endless hours, essentially one person full time on a big collaborative project negotiating these issues, and it's just about got away through, but the fact that there is no national policy has cost us [...]*" (MR7).

A critical issue for data sharing agreements in the social and medical sciences is the protection of data subjects' privacy. Some researchers see imposing **confidentiality preserving measures** as an impediment. One medical researcher commented: "*barriers we sort of uncovered, because we did run a survey in trial centres, there is a barrier with data sharing which is based on concerns about confidentiality*." (MR5). Often, sharing policies for confidential data stipulate that data subjects must be anonymised. Some respondents felt that such measures were sometimes overly restrictive and can inhibit research. A social researcher remarked: "*in my view some of the survey data is unnecessarily reduced in its detail, sometimes I can fully understand why do this, sometimes I don't think it's necessary*.'' (ES7). A medical researcher commented: ''*I would argue that perhaps there is too much bureaucracy, it does mean that sometimes you cannot do things with data that, if it wasn't for that it would be perfectly reasonable to do, in my view*.'' (MR6). Another respondent took a more pragmatic view: "*there is a barrier with data sharing which is based on concerns about confidentiality, there are means of dealing with that*.'' (MR5). In a similar vein, a social scientist revealed: "*It's really important for me that I use real* *data and current data and that means that I work closely with people who are actually monitoring what's going on in society today, like telecommunication and financial services […] their data […] represents real consumption processes and it's not subject to any of the problems of upside or downside, if you like, of archive resources which are often dealt with in a particular ways to make them more consistent perhaps by to take things out by making them subject to all sort of privacy concerns*." (ES2).

Where research would be impossible if datasets are anonymised, researchers need access to **secure environments** ('data enclaves') for analysis and storage: ''*if we want to link individual level data and link things like census data and council registries then we need to have names and addresses […] have you heard of pseudo-anonymisation where you create an ID from a name and address, for example, and then those pseudo-anonymised IDs can be […] kept in a very secretive file*.'' (ES3).

More generally, a **lack of trust in the security** of distributed computing systems was noted as a barrier: *"other projects (not necessarily medical ones, but also engineering ones) where the organizations involved saw the Grid as great looking solution but didn't want their data to leave their network […] it has happened before where companies involved didn't want to go beyond some toy examples, despite the project being able to solve a lot grid related problems*.'' (EP1).

4.5.4.2 Unfamiliar or Heterogeneous Formats

Once barriers to sharing have been resolved, the existence of **heterogeneous data formats** may still cause problems. A social scientist respondent remarked: "*I mean it's my problem really, but, you know, it's getting the stuff in a format that you actually know what to do with it basically, so, you know, they present some boundary files in, you know, ILP format or maybe mapping format or what have you, but you still need to know, obviously, what to do with that, that's different and it isn't always in the form that you wanted basically, because they put it in the form that they wanted receive it, rather in the form that I as user want […]*" (ES3). Another respondent lamented the lack of "*common data formats so that you don't have to know a hundred data formats*." (NE1). This respondent continued: "there is a need for "core data services which serve fairly raw data and also value added services top of that that package up the data in a way which could be more valuable for certain clients [...] you just get the data that you need in a format that you expect it." (NE1).

4.5.4.3 Quality

Most kinds of automated data manipulation and analysis require data to be of good quality, regular, well-defined and well-described. Very often, though, data in the Arts and Humanities, the Social Sciences and in Medicine (e.g., hospital records) is highly irregular, lacks adequate metadata and is of varying quality. Consequently, automated processing cannot be applied without further effort, workarounds or methodological compromises. For example, an A&H researcher said: "*the e-Science access to census holdings [project] kind of died a death because the data which was available wasn't good enough to use any of the tools that social scientists use to look at the data, to manipulate it because the nature of the data is that it is fuzzy, it is not scientific data. […] that is on hold until we can get better data*." (AH1).

4.5.5 Curation

Curation, i.e., the preservation, archiving and maintenance of digital resources, is becoming an issue, which researchers increasingly expected to grapple with. For example, if new (or derived) datasets are to be discoverable, researchers must prepare them for deposit according to accepted standards. One respondent pointed to the problems that depositors face owing to **current repository practices** not being 'user-centric': ''*what we've been trying to do is to look very carefully at the interaction and the user models offered by the standard, kind of repository practices up to now and seeing how they need to be adapted. So, for example, the repositories tend to provide a view of the world which is very much the librarians' view […] it emphasises the description of an item rather than the item itself and what we've been trying to do is work out the simplest way for people to upload their digital materials […] we have tried to make the actual files or documents themselves […] much more at the heart of dissemination process, rather than the metadata, the title or the abstract, the authors, all those kind of things which are given much more priority in the librarian standard view of how a repository should operate*.'' (EP8).

In some fields where depositing is not established practice, **archiving for re-use** is seen currently as being expensive: "*We do do some archiving which we can afford to maintain in-house or on the campus so we've spent some money there. It's simpler for us to do that because we need a fairly low level archive, we don't particularly need to go back to it; we just want to archive it. You know, the Atlas store gives you fairly rapid access back again but that comes at a cost premium so we didn't go for it*.'' (BB2). Another respondent commented on the lack of adequate **financial support** makes proper curation infeasible: ''*Another barrier is the long term support of databases […] research funding bodies are proved to be quite reluctant to do that, the US government and NIH has been the best by quite a long way, and the European Union does help through European Bioinformatics Institute, the other major funders want scientists to share data but they are not showing enough evidence to me that they are actually putting money into a resource in which it can be shared. They recognised the problem, they say they want data to be shared but they are not actually funding them*." (MR5).

A number of respondents raised the issue of how the problem of **inadequate metadata** might be tackled: "*so it's richer, and it has more information, more tags if you like, more labels, so that whoever deposited that data would also indicate other data and related data, expansion of more data, and then, of course, rapidly more build up of data and people as well*." (ES5).

Another respondent, noting that curation must increasingly embrace outputs over the whole of the research lifecycle commented on whether it is feasible to capture sufficient **provenance** to make research replicatable: "*in terms of dealing with relatively complex data and relatively complex analytical techniques, at least complex to the perspective of social scientists, there is this basic tension between describing things clearly and from an introductory level and having enough space to go into the more details, detailed output*." (ES7).

One of them spotted the additional problems arising from the relationship between research and research infrastructure between the NHS and the University, which was described as terrible: "*They are physical walls between them on the equipments, they are mostly about privacy and confidentiality and the lack of a common framework for well-negotiated between the NHS and the universities for handling research other than bio-clinical members, other than my people who have joint appointments in a trust and are clinical members of a trust*."' (MR7).

5 Summary

The completion and analysis of two rounds of fieldwork has helped us gain insights into current research practices, the state of adoption of e-Infrastructure and e-Research methods, and brought to light a great variety of issues, barriers and enablers across the research fields which influence their uptake. The majority of these issues were broadly in alignment with the typology of barriers, which results from our initial desk research, while others were added to the typology during the analysis of the data.

The typology of barriers and enablers has been useful in that it has provided a systematic way into the data collected. However, it does not allow us to capture completely and understand the way that factors implicated in the uptake of e-Research are interrelated. In reality, these do not lend themselves so readily to the simplifications imposed by hierarchical representation. In the case of e-Research, we might expect and, indeed, are discovering, that there is a network of interrelated factors, some with complex or subtle interdependencies, embedded in a wide variety of situations or circumstances, which impact on the uptake of e-Research services and resources.

We take up the implications of our findings in Deliverable 1.4 (Community Engagement Recommendations).

6 References

Anderson, S. (2007). AHRC e-Science Scoping Study. Available at http://www.ahessc.ac.uk/scoping-survey

Atkinson, M. *et al.* (2009). Century of Information Research (CIR): A Strategy for Research and Innovation in the Century of Information. *Prometheus* 24(1), pp. 27-45.

Barjak, F., Wiegand, G., Lane, J., Poschen, M., Procter, R. and Robinson, S. (2007). Accelerating Transition to Virtual Research Organisation in Social Science (AVROSS), Final Report.

Bijker, W.E., Hughes, T.P. and Pinch, T.J. eds. (1987). The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology. Cambridge, MA: MIT Press.

Charmaz, K. (2008). Reconstructing Grounded Theory. Chapter 27 in Akasyytari, P., Bickamn, L. and Brannen, J. (eds.) *The SAGE Handbook of Social Research Methods*. Sage Publications Ltd.

Lee, C., Dourish, P. and Mark, G. (2006). The human infrastructure of cyberinfrastructure. *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. Banff, Alberta, Canada, November. pp. 483-492.

Lyon, L. (2007). *Dealing with Data: Roles, Rights, Responsibilities and Relationships*. UKOLN, University of Bath, June. http://www.jisc.ac.uk/media/documents/events/2007/06/liz_lyon.pdf

Mascord, M. et al. (2007). Scoping e-Infrastructure Usage: Interim Report, December. Available at: http://www.eius.ac.uk/scoping/index.xml

Newhouse, S., Schopf, J.M., Richards, A. and Atkinson, M.P. (2007). Tech Report UKeS-2007-01, April. http://www.nesc.ac.uk/technical_papers/UKeS-2007-01.pdf

Strauss, A. L. (1987). *Qualitative Analysis for Social Scientists*. Cambridge University Press.

Voss, A., Mascord, M., Arguello Castelleiro, M., Asgari-Targhi, M., Procter, R., Fraser, M., Jirotka, M., Halfpenny, P., Fergusson, D., Atkinson, M., Dunn, S., Blanke, T., Hughes, L. and Anderson, S. (2007). e-Infrastructure Development and Community Engagement. *Proceedings of the 3rd International Conference on e-Social Science*, Ann Arbor, MI, US, October.

Voss, A., Asgari-Targhi, M., Procter, R., Halfpenny, P., Dunn, S., Fragkouli, E., Anderson, S., Hughes, L., Mineter, M., Fergusson, D. and Atkinson, M. (2008). Widening Uptake of e-Infrastructure Services. *Proceedings of the 4th International Conference on e-Social Science*, Manchester, UK, June.

Voss, A., Asgari-Targhi, M., Procter, R., Halfpenny, P., Fragkouli, E., Anderson, S., Hughes, L., Fergusson, D., Vander Meer, E. and Atkinson, M. (2009). Adoption of e-Infrastructure Services: inhibitors, enablers and opportunities. *Proceedings of the 5th International Conference on e-Social Science*, Cologne, Germany, June.

Appendix A: Respondents' Backgrounds

PHASE ONE

AH1 is a senior lecturer in electronic communication in the department of information studies at one of the leading UK universities that is a member of the Russell Group, where she teaches internet technologies, digital resources in the humanities, and web publishing. Her research interests include humanities computing, digitization and digital imaging, artificial intelligence, palaeography, knowledge elicitation, and internet technologies.

AH2 is a senior lecturer and the head of the department of French in one of the UK universities belonging to the Russell Group. His research focuses principally on medieval French historiographical literature. He is also interested in the application of e-Science methodologies (digitisation and various forms of Grid and internet dissemination) to the study of medieval manuscripts and edited texts.

AH3 is a reader and programme director in electronic communication and publishing in the department of information studies at one of the Russell Group universities. Her PhD was in English literature and her research is in the area of digital humanities, particularly in the development and use of electronic texts and digital libraries. She is also interested in the social impact of publishing technologies, and especially the Internet.

AH4 is a director of an institute concerned with computing in humanities and resource development officer within the school of history and archaeology at one of the UK Russell Group universities. She teaches on a range of undergraduate and postgraduate options including arts and media informatics and the MPhil in history and computing. Her research interests include the recovery of digital data, interface and multimedia design, evaluation techniques and digitisation projects.

AH5 is a member of one of the UK Russell Group universities visualisation lab at the Centre for computing in the humanities, where she represents 3DVisA, the 3D visualisation in the arts network. Active in the UK visualization community, she has recently convened various workshops. She has also initiated a joint archaeological computing laboratory project, combining the time map/heurist environments to map visualisation research networks, and subsequently refined this approach to map other projects in cultural heritage across different calls and funding programmes. She holds a PhD in mathematical sciences with a focus on abstract visual notation in mathematics. She has also held a faculty position in the school of interactive arts and technology before their current post.

AH6 is the head of the text archive in the research technologies service in computing services at one of the UK Russell Group universities. He is also on the executive board of an important project building a pan-European research infrastructure for research with language resources in the humanities. He has recently taken up the role of liaison between the research technologies service and the relevant e-Science research centre, in which role he has a particular interest in developing e-Research projects in the arts and humanities.

AH7 is working as a project manager in one of UK Russell Group universities. The project aims to build virtual research environment for the humanities in the UK. She is responsible for the day to day running of the project, including the conduct of the detailed user requirements survey.

AH8 is a reader in the intelligence, agents and multimedia group in the electronics and computer science department at one of the UK Russell Group universities. Her main area is Human Computer Interaction (HCI) / Human Factors. Her research focus is on interaction and information systems design to support knowledge building in desktop, mobile and increasingly pervasive environments. In other words, she is interested in looking at how we can design tools to help connect information on massive information repositories like the Web so that we can make better use of the information that's out there. That is, design ways of accessing, representing and manipulating this information to find out what we want to know, learn what we want to learn, more effectively so as to increase knowledge, delight, discovery, quality of life.

BB1 is a researcher in the school of geography and the environment at one of the UK Russell Group universities. He has a BSc in botany, an MSc in plant ecology, an MSc in forestry and its relation to land use and a D.Phil. in tropical forest diversity and conservation. His research interests focus on three areas of investigation, each addressing a number of questions: 1) Conservation beyond protected-area boundaries. 2) Cultural and spiritual values in conservation. 3) Climate change and conservation. His current research involves a project investigating legacies of human management of a cultural landscape in different parts of the world.

BB2 is a researcher at the division of structural biology at one of the UK Russell Group universities. His research interests have moved from a background in chemistry and computation to structural biology and structural bioinformatics: the application computational methods to the study of biology. His work has focused on methods development from bioinformatics through protein production and data tracking and structure determination to structure analysis and display.

BB3 is a researcher in the department of biomathematics and bioinformatics at one of the seven institutes sponsored by the Biotechnology and Biological Sciences Research Council (BBSRC). His research is involved with the National Centre for Text Mining, they're collaborators on the grant and they have expertise in text mining technology which is explored and exploited within the biology projects. Also he is part of bioinformatics groups that works on systems biology projects and are collaborators in systems biology projects within BBSRC funded institutes.

BB4 is a research fellow in the school of computer science at one of the UK Russell Group universities. She is a biologist by training, and works in a bioinformatics project which has developed a comprehensive loosely-coupled suite of components specifically to support data intensive *in silico* experiments in biology. Workflows and query specifications link together third party and local resources using web service protocols. She works with users, helps with workflows and gets involved with the user community. She also does work with semantic discovery services and workflow construction. Her role is entirely user-centric, she is not involved in programming. She has been using advanced methods in ICT for over 6 years.

BB5 is a professor of biology and bioinformatics, he is a faculty member of School of Life Sciences and affiliated with the School of Computer Science at one of the UK Russell Group universities. He describes himself as mostly a computer scientist. His role is to 'catch problems', meaning that he is the person on the team who understands the life sciences and medicine sufficiently to be able to pose problems for the informatics core. He isn't working on any specific projects that use ICT tools and methods, but the field of bioinformatics is very computer-based, and all of his work requires the use of a range of JISC and other services.

BB6 is a research student currently working towards a PhD in Bioinformatics at one of the UK Russell Group universities. In his research he is looking into the use of web services and workflows in aiding Genotype and Phenotype correlations. This is being conducted through the MyGrid project utilising the Taverna workbench as the means of building such workflows. The benefit of using this web services based technology is that it provides a systematic, un-biased, and explicit means of analysing vast amounts of scientific data. This project so far has been able to link the genotype to the biological pathways influencing the phenotypes expression. Work is currently underway to determine the exact role (if any) each pathway may play. This involves the use of text mining methods, conducted through the use of web services and workflows.

BB7 is a researcher at the European Bioinformatics Lab at one of the UK Russell Group universities. The Bioinformatics lab is primarily responsible for archival storage and curation of biological sequence information (DNA and Protein sequences). He is part of a research team whose main interests are in methods for the analysis of DNA and amino acid sequences to study evolution. Past work has been on the theoretical basis of phylogenetic analyses of this data, aiming to understand and improve methods, and on the development of statistical methods that test the accuracy of current mathematical models. More recent work has concentrated on devising new, better, mathematical models and the application of these models and methods to comparative genomic data.

BB8 is a researcher at one of the UK's leading crop research centre. Her research focuses (along other areas) on root trait responses to soil physical conditions, including changes in border cell production, exudate production, changes in root meristem and overall root structure. The main goal of her current project is to use image analysis of confocal images to understand the growth of plant roots, so she employs confocal microscopy images and using computer software that has been designed enables her and her research team to analyse roots, root growth and start to understand the differences.

NE1 is a researcher in an e-science centre; environmental system science centre at one of the UK's 1994 Group universities. His research interests are on e-Science methods in environmental sciences and numerical modelling of volcanic processes. His current work is concerned mainly with computer modelling of the flow of gas through permeable volcanic rocks and how this process affects volcanic eruption dynamics. As well as working within the environmental sciences community, he is interested in how Grid computing can be made easier to use by all scientists with little technical background. He retains a strong interest in physical volcanology, the subject of his Ph.D. He is involved in many projects, to just mention one of the project's goal here: it is to build up a sustainable pan-European capacity in providing timely, quality assured marine service (including data, information products, knowledge and scientific advices) in European coastal-shelf seas.

NE2 is a researcher at one of the UK Russell Group universities. His research focuses on the computer modelling of imperfect and disordered minerals in general and on mantle minerals in particular. He is particularly interested in understanding the microscopic processes such as diffusion and the movement of dislocations that ultimately lead to global scale convection in the Earth's mantle. A second theme of research involves improving our use of modern computer methods in science. As part of this work he develops methods and tools to manage large numbers of calculations and handle the resulting data deluge.

NE3 is a Professor of mineral sciences at the department of earth sciences in one of the UK Russell Group universities. His main research interests concern understanding the behaviour of environmental materials at a molecular level, using a mixture of atomistic computer simulation (theory) and neutron scattering (experiment) techniques. Recently, he has led major projects to apply grid computing technologies to materials modelling under the national e-Science initiative.

NE4 is a remote sensing scientist researching at a specialist research institutions associated with a UK university. He is a member of the research team representing one of the largest clusters of expertise in marine science and technology in Europe. This cluster of teaching, advanced training, research and knowledge transfer sits within a region of excellence in bio-medical sciences and environmental technologies, with organizations such as the UK MET Office and the Peninsula Medical School.

NE5 is a lead programmer for a project working at the department of Geographical Sciences in one of the UK Russell Group universities. The project's aim is to build simplified and faster-running models of the Earth's climate system, and make them easier to use and more widely available to other people who want & need to use them. State-of-the-art climate models (such as the excellent Hadley Centre model) work on quite detailed scales in both space and time, and are consequently big and slow, and cannot be used to simulate more than a few centuries of climate on anything other than a super-computer, and even so may take months to give results. His research team aims to model the climate for many thousands of years, and they aim to do this using more widely available computers (including top-end PCs) so that more scientists who are not computer experts can explore their ideas.

NE6 is a grid software developer and programme manager at one of the e-Science centres that is associated with an UK Russell Group University. He develops e-Science tools for environmental science projects involving computational grids. His background is in gas industry research and development. He is mainly concerned with Grid middleware development for a community of scientists for running climate modelling programmes. He is actively involved in researching solutions to wider issues including sharing, comparing, and analyzing large datasets. The Grid middleware tool he is involved in developing is concerned with usability issues surrounding how best to use them on remote resources. Therefore his research involves both climate modellers and e-Science experts.

NE7 is a research associate at the department of civil engineering and geosciences at one of the UK Russell Group universities. She is interested in the following interconnected research areas: Catchment hydrology and climate change impact assessment, Shallow landslide modelling and Mineral physics & the earth's deep interior. One of the main e-science projects that she is involved in is modelling of all aspects of the water cycle. The project brings together a number of teams from around the world, inviting them to integrate their models in order to look at the way in which different aspects of the ecosystem impact upon each other. The different teams from around the world have contributed models in areas such as fire, erosion, etc. e.g. the fire modelling team would create a map of the burned area and then the team at Newcastle could use this information to look at the impact of the fire on the hydrology. This project made use of Grid technology, but she was not involved with the technical side of the project.

NE8 is a senior research fellow at one of the e-Science centres; environmental system science centre, at one of the UK Russell Group University. He is involved with diagnosing atmospheric models for storms, particularly in relation to climate change. One of the projects he is working on was co-ordinated by his research team and involved all the major climate modelling groups in Europe. The project was supported by the European Union. The aim of the project was to produce a long time-series of three-hourly global window channel thermal infra-red images of the Earth and to test the feasibility of using this in evaluating atmospheric General Circulation Models. Its archive currently spans the period 1st July 1983 - 30th June 2006. Responsibility for maintaining and updating the archive resides with NE8 research team.

EP1 were a group of four: researchers, senior researchers and technical professionals, grouped as medical physics, working at the Cancer Research Centre at one of the UK Russell Group universities. Their current project aims to develop a prototype of a national facility which could be used by medial professions requiring high throughput Monte Carlo based radiotherapy planning in diagnosing various type of cancer. This is essentially collaboration between radiotherapy physicists based at the cancer centre and Computational Scientists. At this stage of their project they have active engagement with prospective users from outside their institutions but are not yet running operationally on that basis. Their research is on development and evaluation of techniques for the dosimetric verification of intensity modulated radiation therapy. Radiotherapy involves the use of high energy X-rays for the treatment of cancer. Intensity modulated radiotherapy (IMRT) is a sophisticated form of radiotherapy in which each beam is split into many smaller 'beam-lets', resulting in 'beam modulation', and is the desirable treatment in situations where the tumour volume is in very close proximity to or impacts on 'organs at risk', such as the spinal cord. Verifying the accuracy of treatment delivery is of utmost importance in radiotherapy physics, and the more sophisticated the dose distribution, the more complex the verification solution. This project aims to develop methods for accurate and efficient verification of IMRT.

EP2 is a senior lecturer at the school of engineering sciences at one of UK Russell Group universities. He is a member of aerodynamics and flight mechanics group. His research interests are, to name just few areas: unsteady aerodynamics and aeroacoustics, racing car aerodynamics and flight simulation. He has won the ExxonMobil Gold Medal and the School recognized as a Centre of Excellence in Engineering Teaching by the Royal Academy of Engineering. His current activity aims to seek funding to develop, build and test a near-real time processing system for quantitative aeroacoustic measurements using microphone arrays.

EP3 is a researcher at the department of physics and astronomy at one of the UK 1994 Group universities. He is currently working on an EU funded project, a design study which aims to complete the technical preparation for the construction of the European Virtual Observatory. The European virtual observatory is an integrated and coordinated program designed to provide the European astronomical community with tools, systems, research support, and data interoperability standards necessary to enable astronomers simplified access to the information they need to complete their research. The idea of the Euro-VO is to make it feel as if all the astronomical data and tools are available on the astronomer's desktop, even though they are actually located on systems spread out over the whole of Europe and even the rest of the world. The project is responsible for completing the design work and feasibility studies on the backbone software components that will make the Euro-VO possible.

EP4 is an honorary fellow at the department of chemistry at one of the UK Russell Group universities. His research interests are, to name just a few areas: theoretical chemistry, equilibrium structure, molecular electronic properties and quadrupole coupling. He is currently involved mainly in two projects as he is retired. One is the electronically excited states of aromatics and alkenes where they have an 'experimentally complicated spectrum' and his role is to interpret this spectrum by means of theoretical calculations. The other is Interpretation of NQR and microwave spectral Quadrupole Coupling Constants.

EP5 is a research assistant working in the device modelling group at the Electronics and Electrical Engineering department at one of the UK Russell Group universities. He received a Bachelor of Engineering Degree in Electronics with Music and then completed a PhD on 3D Simulation Techniques for Biological Ion Channels. He is currently developing 3D simulation software in C/C++ and Python to simulate large biological molecules.

EP6 is a lecturer at the chemistry department at one of the UK Russell Group universities. His PhD was in structural systematics and molecular modelling, which he undertook whilst working part-time for the national crystallography service. On completion of his thesis he worked with the Royal Institution, but based elsewhere he worked with a collaborative team to build the highly successful Small Molecule Single Crystal beamline, 9.8. In 1998 he joined forces with colleagues in the same field to establish a new laboratory and manage the national crystallography service. He has successfully been awarded a number of grants in the area of Information Management, eResearch and eLearning, has served on a number British Crystallography Association committee positions and is on the Editorial Board of Supramolecular Chemistry, Journal of Coordination Chemistry and Chemistry Central journals.

EP7 is a researcher at the informatics departments at one of the UK 1994 group universities. Active member of HCI group at her institution, her research investigates how novel and familiar technologies can be used in teaching and learning contexts to challenge and engage science students. The project she is currently working on investigates how these technologies can be used to support teachers to create motivating and engaging science in and beyond the classroom.

In building upon the previous project (1) what resources and support do teachers need to scale-up to full classrooms for creating science sessions which incorporate mobile sensing technologies, scientist collaborations and ICT support infrastructures? (1) Learning through e-Science) was an exploratory study investigating the potential for using e-Science and GRID technologies in schools for scientific learning. We worked with a local and remote school to investigate air quality in the surrounding area, whilst looking at potential for use of wireless and mobile devices in data collection, scientific enquiry and intra-school connectivity tasks. These projects build on work exploring external representations and interactions in mixed reality environments with colleagues in the Interact and Ideas Labs. Mixed reality environments involve a virtual world and elements of the physical world in a space that can represent both simultaneously.

EP8 is a senior lecturer in intelligence, agents, multimedia group at one of the UK Russell Group. He is also a fellow of the Web Science Research Institute. One of the project he is currently involved with aims to develop a prototype demonstrator that synthesises research information from heterogeneous sources (institutional repositories and research council information systems), resolves name co-reference issues between the sources, and presents it to research-focused end users through an interface that will allow them to explore the state of the research landscape in UK higher education.

ES1 is a professor of planning at the centre for advanced spatial analysis at one of the UK Russell Group universities. Among many other positions held, he was professor and head of the department of City and Regional Planning in another UK Russell Group university. He is a Fellow of the British Academy as well as a Fellow of the RTPI, CIT and RSA. His research is in the development of computer based technologies, specifically graphics-based and mathematical models for cities, and he has worked recently on applications of fractal geometry and cellular automata to urban structure. He is the Editor of the journal Environment and Planning B and was awarded the CBE for 'services to geography' recently.

ES2 is a senior lecturer at the management school of one of the UK Russell Group universities. He also holds a chair of electronic business at the business School at the same institute. He has researched widely in the general technology arena, from basic research into microscopic phenomena to novel applications of these phenomena in innovative devices for computing, telecommunications and display applications, with publications in journals ranging from the IEEE Journal of Quantum Electronics to the International Journal of Innovation Management. He holds higher degrees in Physics and Business and has participated in a large number of interdisciplinary research programmes, with funding from the EU, UK and USA. Current projects include a UK Economic and Social Research Council-funded project that links grid resources in Australia to those at an UK parallel computing centre, enabling Grid data fusion and mining of commercial banking and telecommunications data drawn from each country

ES3 is a professor of spatial analysis and policy at the school of geography at one of the UK Russell Group universities. His major research interest is in simulating social and demographic change within cities and regions, and in understanding the impact of these changes on the need for services like housing, roads and hospitals. He is a member of the editorial board of Transactions in GIS and on the programme committee for both the European Social Simulation Association and the International Conference on e-Social Science.

ES3 is a researcher at the school of geography at one of the UK Russell Group universities. He is generally interested in methods for describing, modelling and analysing earth based systems and the processes that interact in the ever expanding Anthroposphere or humanosphere (what he means by the space-time region that humans influence). There is a growing belief that something has to be done to control socio-economics else environmental consequences could be catastrophic to a great deal of life on Earth: Although nature is likely to redress balance, what are the likely costs? Arguably, there is much that can and should be done to mitigate climate change impacts and preserve, conserve and sustain environments and habitats for future generations to enjoy. He is also involved in the development and application of tools for analysing, visualising and disseminating geographical information. In particular, he is interested in developing geographical analysis tools for generalising and exploring data at higher levels of spatial, temporal and attribute detail for larger geographical regions.

ES4 is a senior lecturer in human geography at the department of geography in one of the UK 1994 group universities. She is a feminist geographer with interests in Africa and cyberspace. She has three interconnected research interests: Geographies of development in sub-Saharan Africa, cybergeographies and feminist geographies. These research interests are pursued within the geography department's research group postcolonial worlds and through undergraduate and postgraduate teaching within the department.

ES5 is a professor and director of institute for rural research and a member of staff at the department of geography and environment at one of the UK Big 36 Group universities. He was promoted to professor of transport and environment in 2001. He is interested in the field of transport and environment, and rural inter-disciplinary research with special focus on: Rural digital economy, Rural accessibility and social exclusion, Rural social, economic and environmental sustainability and Inter-disciplinary research in the evidence-based policy context.

ES6 is a senior research fellow and PGR director at the centre for census ad survey research in the school of social sciences at one of the UK Russell Group universities. He has worked at the centre since 1996, mainly in the field of statistical confidentiality, founding the international recognised confidentiality and privacy research group in 2002, and has managed numerous research projects within centre’s remit. He is one of the key international researchers in the field of statistical disclosure and has an extensive portfolio of research grants and publications in the field. He has extensive experience in collaboration with non-academic partners, particularly with national statistical agencies (e.g. Office for National Statistics, US Bureau of the Census, Australian Bureau of Statistics) where he has been a key influence on disclosure control methodology used in censuses and surveys and where the SUDA software developed in collaboration with colleagues in computer science is currently employed. Aside from confidentiality, privacy and disclosure his research interests include psychology and sociology of personal relationships and social network analysis.

ES7 is a lecturer in sociology at the department of the applied social science at one of the UK universities. His work has a particular remit concerned with using quantitative research methods in the social sciences. His research involves conducting sociological research which makes use of quantitative secondary survey datasets, and methodologically oriented teaching and research. He has worked as a researcher at two UK Russell Group universities before joining his current position. His research interests are, to name a few areas: structures of social stratification, the uses of occupational information and occupation-based social classifications, social stratification, ethnicity and immigration, and e-social science.

ES8 is a lecturer in econometrics at the school of social sciences at one of the UK Russell Group universities, he has worked as a full-time researcher at various (Russell Group universities) and abroad. He has also been a statistician for the NHS. His main research interests lie in the estimation and misspecification testing of micro-econometric and limited dependent variable models. Recent areas of application include experimental economics (binary panels), and observational labour market behaviour (grouped count responses, determinants of pricing for non-standard occupations). Present research activity is concentrated on using an e-Research infrastructure for combined data modelling.

ES9 is a professor of health psychology in the school of psychology at one of the UK Russell Group universities. She is a joint editor-in-chief of psychology and health, journal of the European Health Psychology Society from January 2007. She is the head of the centre for clinical applications of health psychology, and director of the MPhil/PhD in health psychology research and professional practice. Her current interests are in using the internet to support self-management of health, chronic illness (especially dizziness and falling) and attitudes and adherence to treatment and rehabilitation, including physiotherapy and CAM. An important focus is on empowering people in the community to take control over their illness and treatment, and she is involved in several clinical trials and other interventions. She has expertise in using and combining a wide range of both quantitative and qualitative methods, and has edited several books on research methods. Much of her previous research has focused on prevention and treatment of balance problems and the experience of dizziness and vertigo, especially the links between balance and anxiety disorders (e.g., psychophysiological and cognitive processes in disorientation and panic, perceptual-motor and cultural factors in balance and agoraphobia).

ES10 is a senior research fellow in urban planning and geographic visualisation at the centre for advanced spatial analysis in the faculty of the built environment at one of the UK Russell Group universities. The main topics of his research involve developing visual techniques for urban planning, three dimensional city visualisations and public participation by means of web-based systems utilising GIS and CAD. He has been instrumental in developing his research team into a world leading research lab in the area of visualisation and transferring research into the commercial and private sector, most notably through the recent projects which he manages.

MR1 is a research associate, and a biologist by training at the institute of human genetics at one of the UK Russell Group universities. She is part of a research team working on human reproductive cell biology and developmental genetics. The programme on reproductive cell biology focuses on mechanisms that regulate chromosome segregation during oogenesis. The human developmental genetics research is focused on developing a pioneering of high resolution gene expression in embryonic development. The group is also using animal and cellular models to investigate the molecular basis of several human malformations as described under their mammalian and vertebrate developmental genetics research theme.

MR2 is a senior lecturer in primary care at the e-health research network at the centre for public health & primary care research at one of the UK Russell Group universities. She is convenor of e-health interdisciplinary research group and academic director, MSc in health informatics and MSc in health information governance. Her research interests are in exploration and evaluation of e-health interventions, including organisational and social influences on implementation and clinical and psychosocial impacts. She is also interested in the following: systematic review and policy analysis of e-health risks and benefits, evaluating complex interventions, linking policy and research, consumer and patient engagement, and therapy adherence in chronic disease.

MR3 is a reader in biological resource management and the archive director at the centre for the integrated genomic medical research at one of the UK Russell Group universities. His primary current research interest is in biobanking. This is an emerging discipline within biological sciences that takes a rigorous approach to the handling of biological resources. With the completion of the human genome project, very large, high quality, accessible, sustainable resources have become essential to biomedical research translating the human genome sequence into health benefits. This then has implications for sample handling methodologies, for bioinformatics, for operational management and their setting in an appropriate collaborative, governance, ethical and societal framework. Intergovernmental bodies have advised on the need to develop global and European biological resource networks. He is actively involved in this work as associate coordinator for the EC-funded European biobanking and bioMolecular resources infrastructure. He is also involved for the UK DNA banking network which is funded by the MRC. Through these activities he has many hundreds of collaborators in the UK, elsewhere in the EU and internationally.

MR4 is the head of bioinformatics and the MRC human genetics group in an NHS hospital at one of the UK Russell Group universities. His research interests broadly speaking straddles two areas: computational genomics and molecular evolution, using methods from these fields to shed light on human evolution and disease. Such activities have been classed as 'evolutionary systems biology'. His current research projects include: 1) Systems-level insights into human disease genes using statistical meta-analyses of functional annotation. 2) Human chromatin structure, gene expression and evolution. 3) The evolution of mammalian transcriptional regulation networks. 4) Discovery and evolution of novel components of the mammalian innate immune system. 5) The comparative genomics and evolution of the ubiquitination pathway.

MR5 *is* a professor of experimental therapeutics; antibody-based target therapy of the cancer research UK targeting and imaging group at the cancer institute at one of the UK Russell Group universities. He is a group leader and his research interests are: 1) Improving the understanding of cancer and the care of people with the disease. 2) Targeted cancer therapy and imaging. 3) Treatment of gastrointestinal cancer, germ cell tumours and lymphoma. 4) Bioinformatics and mathematical modelling in cancer therapy

MR6 is a senior researcher at the bioPhysics group in the physics department at one of the UK Russell Group universities. One of his current research project concerns constructing a prototype truly simultaneous combined PET and MR scanner and is exploring design concepts for clinical PET/MR. This is a collaboration between the brain imaging centre and the department of physics. PET/CT has shown the potential of bringing two scanners together, albeit sequentially. MR6 research group unique concept of the PET/MR scanner will allow for the first time high resolution PET and MR to acquired simultaneously. The system is designed for molelcular imaging studies, although other applications are possible, such as process engineering. The prototype currently under construction is the first stage in the development of a human system, using a novel technology to deliver state-of-the-art molecular imaging.

MR7 is a professor of medical informatics in the school of computer science at one of the UK Russell Group universities. Over the past twenty-five years he has led a series of projects on clinical decision support, medical records, and medical terminology including the ground breaking on intelligent medical records sponsored jointly by the UK medical research council and department of health. During the 1990s his work focused on medical terminology and ontologies. His work on clinical terminology and ontologies provided a key stimulus for the technologies underpinning the use of ontologies for the Semantic Web. Increasingly his work has focused on the development and application of the Web Ontology Language, OWL, and its development environment. He currently leads two industrial collaborations using OWL to enhance clinical systems. One is US based and the other is with an UK university on adaptable forms for pre-anaesthesia assessment and other specialised information gathering tasks. This work leads back to a concern with methodologies for using OWL and other Semantic Web technologies effectively in practical applications and the relationship between OWL and object-oriented and frame based modelling systems. He has participated in the EU-funded semantic mining network of excellence and semantic health roadmap projects. He has been active in various national and international committees including, the JISC Support of Research Committee, the National Cancer Research Institute's Informatics Initiative, the Joint NHS/Higher Education Forum on Informatics, and the Board of the Academic Forum of the UK Institute for Health Informatics.

MR8 is a senior clinical research fellow and professor of epidemiological psychiatry at one of the UK Russell Group universities. He is also an honorary consultant psychiatrist, at the mental healthcare NHS Trust associated with his institute. He is the chief investigator on the international trials project and director of the medical research institute research center in treatment trials in bipolar disorder. He is principal investigator on the UK medical research council funded project that will use e-Science technology to facilitate large scale neuroimaging studies in clinical neuroscience.

MR9 has a role of research facilitator and project manager at computing laboratory at one of the UK Russell Group universities. She has joined academia from industry having spent many years working in the IT industry. She was project manager for an eScience project, and is currently project manager for a range of multidisciplinary/multi-institution research projects. She is research investigator on another eScience related project which is looking at management techniques for ensuring usability and pursues her own interest in determining new project methodologies for these diverse and complex projects spanning industry and academia. She also works as research facilitator for the university computing laboratory and supports the department in increasing its research income.

ST1 is a professor of astrophysics at the institute of cosmology and gravitation at one of the UK universities. He has been a researcher in the US for many years prior joining his current institue. He has experience in using NGS, and he was involved in the AstroGrid project. His main research interests are: 1) Glaxy evolution studies, primary using data from the Sloan Digital Sky Survey (SDSS). He has been a member of this survey since 1993 and has full data rights and he is a "SDSS Builder". 2) Constraints on cosmological parameters, again mostly using the SDSS and WMAP datasets. Recently, he has also become involved in the SDSS-II

Supernova Project. 3) Mapping the large-scale structure in the Universe using novel techniques and large datasets. 4) Construction of new surveys of galaxies and clusters of galaxies. 5) Multi-disciplinary research into new statistics and algorithms for the analysis of massive astronomical datasets.

PHASE TWO

Fifty seven candidates (IT personnel) from twenty five institutes, two from each institution were interviewed. The aim was to interview one person with a strategic role and another with an operation/application role. The institutions were as follows:

Seventeen were members of the Russell group universities.

Three were members of the 1994 Group

One was chosen from a specialist research institution.

Four were chosen from the middle to lower ranking universities on the Times Higher Education ranking list (2006-07)

References:

Russell Group. See: **http://www.russellgroup.ac.uk/about.html**

Described as: "The Russell Group is an Association of leading UK research-intensive Universities committed to maintaining the highest standards of research, education and knowledge transfer."

Million + group. See: **http://www.millionplus.ac.uk/index.htm**

Described as: "We are a university think-tank. This means we work to help solve complex problems in higher education and to ensure that policy reflects the potential of the UK's world-class university system."

1994 Group. See: **http://www.1994group.ac.uk/memberinstitutions.php**

Described as: "Established to promote excellence in research and teaching. To enhance student and staff experience within our universities and to set the agenda for higher education."

Appendix B: Questionnaire

Introduction

The e-Uptake project is a major JISC-funded study of use of its services.

This questionnaire aims to gain background information on academic users of advanced ICT methods for research and teaching (e.g. video based interaction; synchronous and asynchronous e-research; 3d scanning; modelling; user interface design; graphical collaborative publishing; data mining; image segmentation, etc), and their experience of JISC funded services.

The answers you give in this questionnaire will form the basis of a structured telephone interview, to be conducted at a time convenient to you.

This questionnaire should take you around 5 minutes to complete. Please remember that you need to fill in the questionnaire before the interview.

A copy of your answers will be emailed to you shortly after you complete this questionnaire.

Thank you in advance for your help in our research.

Data Protection Statement

All data collected in this survey will be held anonymously and securely. No personal data will be retained.

Your name is needed only for administrative reasons.

Cookies, personal data stored by your Web browser, are not used in this survey.

Personal Information

1. Name: […………………….]

2. Institution / Department (please write in) […………………….]

3. Please give your current primary disciplinary area (Please tick one)

   [ ] Engineering and Physical Sciences

   [ ]Arts and Humanities Sciences

   [ ] Economics and Social Sciences

   [ ] Biotechnology and Biological Sciences

   [ ] Medical Sciences

   [ ] Natural Environment Sciences

   [ ] Particle Physics / Astronomy

   [ ] Other (please specify): […………………….]

4. Please indicate your current position (Please tick one)

   [ ] Professor

   [ ] Reader / Senior Lecturer

   [ ] Associate Professor/Lecturer

   [ ] Research Associate

   [ ] Research Assistant

   [ ] Other (please specify):

5. How long have you been using advanced ICT tools and methods for research and teaching? (Please tick one)

   [ ] 1-5 years

   [ ] 6-9 years

   [ ] 10 or more years

6. Are you currently working on a project that involves advanced ICT methods? (Please tick one)

   [ ] Yes

   [ ] No

   If yes, could you please specify what project(s) you are working on? (Please write in)
   […………………….]

7. Out of 10, how would you rate your own ICT expertise for research and teaching? (1= Novice and 10= Expert)(Please tick one)

1	2	3	4	5	6	7	8	9	10
[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]

8. Does your institution provide support for advanced ICT methods use specifically in research and teaching? (Please tick one)

   [ ] Yes

   [ ] No

[ ] Don't know

If your institution provides support for advanced ICT methods use specifically in research and teaching, please specify what form this support takes (Please write in)

[……………………..]

9. Have you attended any training/course for researchers in the use of advanced ICT methods? (Please tick one)

[ ] Yes

[ ] No

Use of JISC Services

10. Awareness and use of JISC services (Please tick all that apply to you)

	I am a frequent user of it	I am an occasional user of it	I am aware of it but never used it	I have not heard of it
Access Grid Support Centre				
Edina				
Mimas				
UK Data Archive				
Digital Curation Centre				
National Centre for Text Mining				
National Grid Service				
OMII				
UKERNA				
VizNet				

11. Have you ever sought support in using any of these services? (Please tick one)

[ ] Yes

[ ] No

If yes, for which services have you sought support? (select all that apply)

[ ] Access Grid Support Centre

[ ] Edina

[ ] Mimas

[ ] UK Data Archive

[ ] Digital Curation Centre

[ ] National Centre for Text Mining

[ ] National Grid Service

[ ] OMII

[ ] UKERNA

[ ] VizNet

If you sought support, what kind of support did you use? (select all that apply)

[ ] Emailed helpdesk

[ ] Contacted local representative

[ ] Asked more experienced colleague

[ ] Read the guide provided

[ ] Attended a training course provided locally

[ ] Attended a training course provided from the service provider

[ ] Other (please specify): [……………………..]

12. Do you know people that use these services? (Please tick one)

[ ] Yes

[ ] No

If yes, which ones do they use? (select all that apply)

[ ] Access Grid Support Centre

[ ] Edina

[ ] Mimas

[ ] UK Data Archive

[ ] Digital Curation Centre

[ ] National Centre for Text Mining

[ ] National Grid Service

[ ] OMII

[ ] UKERNA

[ ] VizNet

13. Are there any other JISC services that you use? If yes, please specify (Please write in)

[……………………..]

14. Are there any services that you use and that are not provided by JISC? If yes, please specify (Please write in)

[……………………..]

15. Have you ever contributed data or uploaded a project to any of these services? If yes, please specify (Please write in)

[……………………..]

16. Do you have any other comments on JISC's services? (Please write in)

[……………………..]

Final Page

Thank you very much for your time and help

Appendix C: Interview Proforma for Phase I (Researchers)

JISC e-Uptake project

Interview Pro Forma

Notes for interviewers are given for each question to allow them to further elaborate the questions where necessary.

The interviewers should ask permission to record the interview making clear how the information will be used: interviews will be anonymised and if their name needs to be mentioned, they will be contacted beforehand (see consent form).

At the beginning the interviewers will need to inform the interviewees on the areas of questions we will ask.

Can you please describe which JISC services have you used?

> In case the interviewee has more than one role in a project or is involved in different projects, the interviewers must acknowledge it and let the interviewee set the framework for the discussion.
>
> If the interviewee cannot tell if the services she uses are JISC, then she can speak of any e-science services she uses. In the questionnaire we ask about their use of other JISC services that they might use (excluding the ones we mention) and other non-JISC services. In case they mention any, we need to follow up and ask how frequent users they are and description of their use.
>
> N.B. From the questionnaire we have already which services they use. It is useful to take them one by one and ask all the questions of our proforma referring to each one of them if possible (as Marzieh did in her pilot interview with [anonymised]).

How did you find out about them?

How does using these services facilitate/enable your research?

> Mention the notion of a research lifecycle and ask interviewees to cover the whole breadth of their activities. Specifically, ensure that the data collection, analysis, collaboration and publication stages are covered.

Did you seek any kind of support (e.g. talked to a local representative, emailed the helpdesk)?

What support did you receive? Where did it come from (e.g. talked to a local representative, emailed the helpdesk)? Did it come from the service or from a local provider? Was it helpful?

> If people have indicated that they have taken training, this should be followed up in the interview to ask what training that was and what their experiences were.

Have you faced any barrier while using these services? Was it a barrier directly related to a JISC service (e.g. difficulties in navigation)? Was it a non-JISC barrier (e.g. lack of resources)?

> When researchers tell you about a barrier or difficulty they have faced using services, always ask for an example. That will allow you to dig deeper into the barrier, and it does make you understand the problem better.

How have you overcome them?

What would further enable your use of these services?

> For example: more training, better or different kinds of support, additional resources, changes in research policy, etc.? How could these be provided or achieved?

Is there anything else you would like to mention/add regarding your use of the services?

> If they have completed the project they were working on, were they happy with the results in the end? What would have been the ideal situation in terms of the problems they encountered?
>
> And if they have any plans for their future research, given what they have been through, would they do it again? Specifically ask about any problems they foresee for their future research? And what would be the ideal solution?

In all questions, consider whether there are expressions of particular attitudes towards tools or services, i.e., how people see them in relation to what they take to be important in their work and its organisation. For example, people may say that a tool breaks established routines that they have invested a lot of effort in.

After the participants add any comments, the interviewers turn off the recorders and they notify the interviewees. They try to sum up and inform them about the event we will organise to promote the use of advanced ICT methods and if they could provide any colleagues that could benefit from it. We need to mention that we face difficulties in approaching people who do not use ICT and that we need their advice on how to do that.

N.B. Interviewers could make use of the following list of services and their characteristics and the questions included during the interviews.

[this was followed by a crib sheet describing he services and their important functions]

Appendix D: Interview Proforma for Phase II (Intermediaries)

Questions to be asked IT services respondents

We need to ask about the role that IT Services currently play as well as the potential for them to play a role as providers or intermediaries. We should focus on research-specific services.

Specific questions (please tick off those that are covered):

- Is there a dedicated research computing support group in you institution? If not, how are the special IT needs of research met? If yes, where in the institution is this function located, e.g., in a central information services department?
- Are there another department in the institution that you work with very closely?
- Are there local (research) IT support groups in the different faculties / schools?
- Ask someone with a strategic role: what is the strategy for supporting research computing needs, what do they see as driving this? How does it fit into the overall university strategy? Do you have regular interactions with members of the university developing its research strategy? What is the funding model for research computing services?
- What local resources exist?
  - HPC compute resources
  - HTC compute resources
  - Grid computing
  - Special services for commercial codes
  - Resouces for long-running serial or memory-intensive jobs
  - Visualisation resources
  - Scientific collaboration (networking, data depository, outputs repository)
  - Distributed scientific collaboration (Access Grid, EVO, etc.)
  - Storage resources (SANs, storage resource broker, etc.)
  - Database resources
  - Specialist digitisation services
  - New technologies and experimental services
- How many users are using these resources? Do you have users from a range of disciplines or only from specific ones?
- What support and training do you provide for the use of research computing resources?
  - Training and support for use of specific services
  - Assessment of research computing needs at proposal stage
  - Application development

  - Help with procurement and hosting of departmental resources (clusters, HPC, etc.)
- Do you provide support and liaison for national services? If so, how do you do this and what issues do you encounter as part of this?
  - Compute services such as National Grid Service, HPCX or HECToR
  - Data services such as MIMAS, EDINA, AHDS
  - Collaboration support, e.g., Access Grid
  - Advanced networking, e.g., lightpaths (UKERNA)
- Are there local e-Science investments such as a local e-Science Centre? If so, what is its function and how does it operate? If not, do you have relations with a regional e-Science centre?
- Is there a UK e-Science Registration Authority (for certificates)? Is there a local NGS node? Are there any other e Science related resources?
- What do you see as the main strategic strengths and weaknesses of research computing in your institution and how do these affect researchers?
- What role do researchers play in formulating the service provision agendas? What do you do to elicit their requirements?
  - Are you proactive in developing research computing or are you waiting for people to come along with problems?
- Do you have experience with specific issues that researchers have encountered in using advanced IT resources or do you have examples from your own experience? Have these issues been overcome and if so, how?
- Have you encountered problems with specific services (including regional or national) that you could tell us about?
- What could JISC or e-Research community do to bring about your ideal world?
- Do you know of an unmet demand for training or support in the area of research computing? If so, what subjects would need to be covered? Would you have expertise and infrastructure to provide this training?
- What services would be helpful in supporting your institution in providing distributed computing training (e.g., repositories of material, specific eLearning resources)?
- Do you consider that the provision of a national infrastructure supporting teaching in distributed computing where worked examples, simplified access and student sandboxing are available, would be useful in helping your institution?