Embedding e-Science Applications: Challenges from the eDiaMoND Case Study

Andrew Warr
Grace de la Flor
Marina Jirotka
Sharon Lloyd
Ralph Schroeder
Mustafizur Rahman

A report from the 'Embedding e-Science Applications: Designing and Managing for Usability' project
(EPSRC Grant No: EP/D049733/1)

oerc

UNIVERSITY OF OXFORD

DOMINVS ILLVMINATIO MEA

EPSRC
Engineering and Physical Sciences Research Council

Table of Contents

Executive Summary

This report is one of two, six month deliverables from the 'Embedding E-Science Applications: Design and Managing for Usability' project. In this report we identify and explore some of the challenges of embedding e-Science applications.

Section 1 introduces the report giving a brief overview of the e-Science vision and the use of grid technologies to enable this vision. This leads us to highlighting technical, social and legal challenges that now face the e-Science community. The purpose of the 'Embedding e-Science Applications' project is to address the challenge of embedding e-Science applications into scientist and researcher communities taking into account the effects of introducing new technologies on both individual work practices and organisational policies. We then describe the structure for the rest of the report.

Section 2 argues for the need to embed e-Science applications. We draw upon arguments made in the Computer Supported Cooperative Work (CSCW) community, that in order for technologies to be successfully adopted by their intended target community, developers need to understand researchers' work practice, culture and their domain of expertise. Because of this, e-Science project partners may need to expand their conception of the development process to include concerns beyond their normal sphere of activity. We further argue that this is particularly challenging in e-Science projects due to the large-scale and distributed nature of the projects, as well as the disparate group of stakeholders involved who have different working practices and motivations.

Section 3 provides an overview of our case study e-Science project, eDiaMoND. We describe the vision of eDiaMoND and discuss many emerging challenges that were new to the eDiaMoND project partners, the e-Science community and the research councils.

Section 4 describes the methodology we applied. We conducted a retrospective analysis of the case study, which included a number of semi-structured interviews with various stakeholders from the eDiaMoND project. The findings from these interviews were further triangulated with a documentation analysis.

Section 5 presents the findings from our retrospective analysis of the eDiaMoND project:

1. **Recalibrating the vision.** Project partners interpret the project vision differently and there is often a mismatch between the project vision and end-users' needs and expectations.
2. **Application developers as users of infrastructure.** Developers are also users of e-Science and grid technologies (e.g middleware). Therefore, we also need to account for the constraints developers encounter.
3. **Working with constraints.** e-Science projects have to work to tight deadlines. Therefore, activities often need to run in parallel. However, it is particular difficult to coordinate these activities when project participants are geographically dispered.
4. **Shared understandings between project partners.** Project partners on e-Science projects come from different institutions, disciplines and cultures. This can cause breakdown in collaboration. We therefore need to enable effective channels of communication for collaboration between project partners, which is particularly challenging in e-Science projects due to the number of stakeholders involved in e-Science projects and their geographic distribution.

5. **Sustainability.** If e-Science applications are not sustained they will not be adopted by their end-users and embedded in their working practices. Therefore, e-Science projects are faced with the challenge of how to sustain their e-Science applications. Currently, the sustainability of e-Science applications falls beyond the remit of the research councils who predominately fund e-Science projects.

Section 6 concludes the report summarising what has been discussed. Based on the exploration of the five challenges identified, we suggest a number of preliminary recommendations for e-Science projects in order to facilitate embedding e-Science applications.

1. The project vision needs to evolve to meet end-users' needs and their organisation's policies. This can be achieved through an iterative process between project stakeholders and end-users.
2. Infrastructure developers, application developers and end-users need to co-design e-Science applications, identifying how to meet the needs of end-users, whilst working within the constraints imposed upon the project.
3. Agile methods can be used to spread the delivery of milestones over time, thereby reducing time pressures. Furthermore, media policies can make sure e-Science project team give a clear and consistent message, thereby avoiding false expectations of e-Science applications to be delivered.
4. Encourage project participants to go beyond their normal sphere of activity collaborating closely with other project participants to facilitate the transfer of tacit knowledge between project partners.
5. Mechanisms need to be put in place to ensure the sustainability, use and maintenance of e-Science applications.

Building upon the above challenges and recommendations we have identified a number of potential future studies. These studies include:

1. We aim to expand upon our retrospective analysis of the eDiaMoND project performing a number of concurrent analyses of ongoing e-Science projects. The aim of these case studies is to validate and expand upon our preliminary findings and recommendations.
2. We wish to perform a diary study of the use of communication/collaborative technologies, thereby increasing our understanding of how project participants communicate and what technologies they use and for what purpose.
3. We are conducting a number of in-depth studies into the use of tools to support distributed collaboration, such as Polycom PVX used to support multi-participant desktop collaboration; and AccessGrid used to support multi-participant room based collaboration.
4. We wish to study implicit collaborations through e-Science infrastructures, such as national grid services (NGS). Furthermore, we would also like to explore the challenges surrounding the sustainability of e-Science applications.

Overall, our research aims to understand how we can best support the embedding of e-Science applications.

1 Introduction

e-Science, or cyber-infrastructure as it is known in the US, promotes a vision of large-scale, collaborative and multi-disciplinary research (Hey & Trefethen, 2002). It is believed that this vision will give rise to new forms of multi-institutional and multi-disciplinary science that will allow new and more complex scientific questions to be answered (Hey & Trefethen, 2002). Not only will this vision create a significant difference to the practices and products of quantitative science, but also to qualitative scientific approaches and practices.

Through an infrastructure known as the *Grid* comprising of high-speed networks, supercomputers, cluster workstations, data repositories and expensive shared experimental facilities, this e-Science vision is made possible (Hey & Trefethen, 2003). While supercomputers are scarce resources only limited to a small number of sites, the Grid makes such resources available to the wider research community. Through federated databases and distributed data stores, the Grid makes a massive amount of data available: more than any single institution could produce or manage. Furthermore, in addition to e-Science projects developing tools to support large-scale, distributed collaboration, such collaboration is required in order to develop these technologies, with stakeholders consisting of requirements engineers, software developers and end-users, just to name a few examples. To support these large-scale, distributed collaborations, tools such as AccessGrid[1] provide a means to support the emerging large-scale, multi-disciplinary collaborations that are critical to e-Science work. Thus, the Grid is a middleware infrastructure between researchers and the resources they would otherwise not have access to from their institution alone. The middleware infrastructure of the Grid enables the construction of what have been termed 'virtual organizations' (Foster *et al*, 2001).

However, whilst the vision of e-Science promises to provide faster and better science (Hey & Trefethen, 2002), a number of challenges have emerged from attempts to turn the vision into reality (e.g. Hey & Trefethen, 2003; Jirotka *et al*, 2005; Lloyd & Simpson, 2005). One such challenge that we address in this report is that of *embedding e-Science applications* – facilitating the establishment and use of e-Science infrastructures and technologies in researchers everyday work practices. In order to start identifying and investigating the issues of embedding e-Science applications in researchers everyday working practices, we have adopted a case study approach looking at a major flagship e-Science project: eDiaMoND[2].

In this report we first set the scene by arguing the need to embed e-Science applications. We then present our case study e-Science project – eDiaMoND – to start exploring some of the challenges of embedding e-Science applications. Describing the methodology we applied to our case study project, we present our preliminary findings. We introduce each finding providing supportive data, which is then discussed. Finally, we draw our conclusions and highlight directions for future research.

[1] www.accessgrid.org

[2] http://www.ediamond.ox.ac.uk

2 Embedding e-Science Applications

By their very nature e-Science projects bring together a multitude of stakeholders, including requirements engineers, software engineers, project managers, researchers and end-users, to name just a few examples. These stakeholders come from different disciplines and institutions around the globe to work together on large-scale, distributed and collaborative projects (Hey & Trefethen, 2002; Hey & Trefethen, 2003). Such projects are extremely complex for a number of reasons: the size of the collaborations, the distributed nature of the collaborations and the unique knowledge, experiences and vocabularies of the collaborators. The size and distributed nature of e-Science collaborations can make it difficult for stakeholders to communicate with each other, as it is not always possible to meet face-to-face. This lack of communication can impact the coordination of stakeholders' activities. Furthermore, the unique knowledge, experiences and vocabularies of the stakeholders can further impact collaboration. It can be hard for a stakeholder from one discipline to understand a stakeholder from another discipline (e.g. Fischer, 1999) – for example, difficulties between end-users and developers has been well documented (e.g. O'Neill, 2000). Such differences are exacerbated due to the number of multi-disciplinary stakeholders involved in e-Science projects. These differences can also lead to differing misunderstandings and expectations (Lloyd & Simpson, 2005). These complexities can all impact e-Science projects and their subsequent outcomes.

The CSCW community who have long been trying to establish best practices and tools to support small group collaboration (e.g. Grudin, 1994; Gutwin & Greenberg, 2000), has faced similar challenges in the light of various projects developing "groupware" applications having been perceived as failures. Such failures were not of a technical origin. Grudin (1994) has identified the causes of these failures eliciting eight challenges for groupware developers:

1. **Disparity in work and benefit.** Groupware applications often require additional work from individuals who do not perceive a direct benefit from the use of the applications.
2. **Critical mass and the Prisoner's dilemma problems.** Groupware may not enlist the "critical mass" of users required to be useful, or can fail because it is never to any one individual's advantage to use it.
3. **Disruption of social processes.** Groupware can lead to activity that violates social taboos, threatens existing political structures, or otherwise demotivates users crucial to its success.
4. **Exception handling.** Groupware may not accommodate the wide range of exception handling and improvisation that characterises much group activity.
5. **Unobtrusive accessibility.** Features that support group processes are used relatively infrequently, requiring unobtrusive accessibility and integration with more heavily used features.
6. **Difficulty of evaluation.** The almost insurmountable obstacles to meaningful, generalisable analysis and evaluation of groupware prevent us from learning from experience.
7. **Failure of intuition.** Intuitions in product development environments are especially poor for multi-user applications, resulting in bad management decisions and an error-prone design process.
8. **The adoption process.** Groupware requires more careful implementation (introduction) in the workplace than product developers have confronted.

Overall these challenges call for a better understanding of work environments and for corresponding adjustments by developers (Grudin, 1994). The first five challenges require us to obtain a better knowledge of the intended users' workplace. The final three require change in the development process. The final challenge in particular requires developers to expand upon their conception of the development process and product to include concerns that have been outside their sphere of activity.

In e-Science the challenges faced by the CSCW community are amplified due to the large-scale, distributed, multi-disciplinary and multi-institutional nature of e-Science projects. Furthermore, the complexity of e-Science projects has caused new challenges to arise such as, how to manage the complexity of multi-disciplinary and multi-institution collaboration (Hey & Trefethen, 2003); trust between stakeholders developing e-Science technologies and applications, as well as trust between the researchers who use these technologies (Jirotka *et al*, 2005); and how to manage successful e-Science projects (Lloyd & Simpson, 2005). While technical challenges still exist, we are starting to see a number of projects emergong aimed at studying the social issues surrounding e-Science technologies. For example, this project – Embedding e-Science Applications: Designing and Managing for Usability (Grant No. EP/D049733/1) is aimed at addressing how to embed e-Science applications in researchers everyday working lives, and how to manage collaborations between e-Science stakeholders. A partner project - the Oxford e-Social Science (OeSS) project (Grant No. RES-149-25-1022) is aimed at addressing legal, institutional and social issues such as confidentiality, privacy and data protection, ownership of intellectual property rights, anonymity and accountability, and issues of trust, confidence, and risk in distributed collaboration. Therefore, if e-Science applications are to be successfully adopted by the research community, we need to consider more than just the technical challenges of e-Science. We have to also consider how technologies will be embedded into the everyday working lives of researchers. The shift from lab-based to digital science is a significant social development for scientific research communities, primarily because their working practices and collaboration activities will change as both are augmented with, or entirely transferred into, digital research environments.

The focus of this project is to understand how best to embed e-Science applications in our users' working lives, thereby contributing towards addressing some of the social challenges to technology uptake facing e-Science. To identify and investigate issues with embedding e-Science applications, we have adopted a case study approach, looking at a major flagship e-Science project: eDiaMoND. In the next section we describe the eDiaMoND project and our reasons for choosing it as our case study e-Science project.

3 The eDiaMoND Project

The eDiaMoND project was a 2-year, flagship UK e-Science project, jointly funded by the EPSRC (Grant No: GR/S20956/01), DTI and a multi-national ICT company in 2002. Oxford University and the major industrial partner, in collaboration with three other universities, four hospitals and one other ICT company, fundamentally drove the project.

The aim of the eDiaMoND project was to apply the e-Science vision of distributed data sharing and collaboration to the medical field of breast screening. The project vision was realised through a UK archive of mammography images and associated meta-data, which was to be utilised by four core applications:

1. A screening application
2. A training application
3. An epidemiology application
4. A research application

A critical component of the project lay in the Grid's ability to provide access to a federated database of mammography images. In the UK at present, there is a growing shortage of radiologists. Through means of a screening application and access to a federated database of mammography images, the increasing workload of radiologists could be spread across the various UK breast screening units (BSUs), rather than individual BSUs having to cope with disproportionate volumes of work. In addition, access to a federated database could aid the training of radiologists. Rather than being limited to the mammography images of a single BSU, interesting cases from around the UK could be utilized in training. Access to digital records would also prevent the need for shipping mammography images, a process that is costly, time consuming and prone to loss. Epidemiologists could use the federated database with appropriate meta-data to perform more refined analyses of the potential causes of breast cancer. The research application could be used for "find one like it" searching (i.e. a search mechanism that would return *n* mammography images similar to the one being viewed) and temporal analysis of changes in the breast (i.e. a comparison of a history of mammography images for a patient). Finally, with such a large collection of mammography images there was a further technical vision of developing learning algorithms to assist radiologists through the computer-aided detection of breast cancer.

The eDiaMoND project, like many e-Science projects, was ambitious, even though it was fundamentally a proof-of-concept: an innovative flagship first phase e-Science project. The project was responsible for 'looking at generic requirements for e-Science projects' (Hey et al, 2003: 1822) in the development of a grid infrastructure and federated databases to demonstrate the capacity of the Grid. It also hoped to support and improve the work practices of radiologists, epidemiologists and researchers. Furthermore, whilst the vision of the eDiaMoND project, like many other e-Science projects, aimed to bring about large-scale, multidisciplinary, collaboration in the scientific domain, such collaboration was also critical in the ongoing development of the eDiamond project. As the eDiaMoND project was a flagship project, many challenges emerged throughout its lifetime for the first time. Hey & Trefethen (2003) describe some of the technical, social and legal challenges that emerged from first-round funded e-Science projects, such as eDiaMoND:

- *Standardising e-Science and grid technologies:* Multi-disciplinary research can inevitably lead to tensions – application scientists want to do applications research and computer scientists want to do computer science research, for example. Successful collaboration between disciplines requires a delicate balancing act. A core part of the UK's Core Programme is to look at generic requirements for e-Science projects to develop a robust and effective middleware for use by all.
- *Establishing intellectual property rights (IPR) within virtual organisations:* Virtual organisations are made up of many institutions that work together to create new IP. A challenge for e-Science projects is to establish who owns the IPR.
- *Secure access to e-Science and grid technologies:* Security protocols are usually designed to protect unauthorised users from breaching an institution. However, in virtual organisations security mechanisms (e.g. digital certificates) need to be able to allow users from multi-institutes with different access rights to access their resources.

- *Protect sensitive data:* Sensitive data such as medical data needs to be protected from unauthorised persons. The Grid and its related services cannot be successful deployed until issues over data privacy and anonymisation are resolved.
- *Establishing trust relationship within virtual organisations:* Researchers have trust relationships between their colleagues and the technologies they use. However, within virtual organisation, how can you establish trust relationship with someone you have never met? Such relationship are further complicated when we consider agent-based systems (i.e. automated processes).
- *Facilitating and promoting the sharing of research data:* The e-Science vision promotes the sharing of research data. However, this is a controversial issue as researchers are often reluctant to share their data, as they have spent the time to obtain the data and they do not wish others to socially loaf over their hard work. Therefore, issues around IPR exist. Furthermore, if this data is to be collated, appropriate meta-data needs to be generated in order to facilitate the sharing of this data. Who will create the appropriate meta-data?

The challenges encountered in eDiaMoND prompted a number of current and future research projects, such as:

- *The 'Generic Infrastructure for Medical Informatics' (GIMI) project* is a large-scale collaborative research project, which is concerned with the development of a secure IT infrastructure to support medical research. The project is developing software to allow secure and ethical aggregation of data from distributed remote sources to support a number of applications, including support for patients suffering from long-term conditions such as diabetes and asthma, image analysis for cancer research and mammography training and auditing.
- *The 'Distributed Intelligent Learning Environment for Mammographic Screening' project* aims to develop and test an Intelligent Tutoring and e-Learning Environment (ITeLE) to provide instruction, support, practice and feedback for trainee radiologists intending to specialise in mammography.
- *The 'Copyright Ownership of Medical Data in Collaborative Computing Environments' (IMaGE) project* aimed to develop a UK-based ownership model premised on IPR for the sharing of digitised data of breast imaging and associated data in a collaborative computing environment.
- *The 'Oxford e-Social Science' (OeSS) project* aims to focus on understanding the legal, institutional and social issues such as confidentiality, privacy and data protection, ownership of intellectual property rights, anonymity and accountability, and issues of trust, confidence, and risk in distributed collaboration, in the domain of e-Science.
- *The 'Embedding e-Science Applications: Designing and Managing for Usability' project* aims to understand how best to managing requirements for usability and developing collaborative approaches to system development that allow for communication of these requirements in the domain of e-Science.

These challenges that arose from the eDiaMonD project were not only new to the researchers on the project, but also to the e-Science community and the research councils as a whole. Due to the novelty of these challenges, the eDiaMoND project was seen to provide an extremely rich set of issues for investigation and thus, a suitable a case study for our project.

In the following section we describe the methodology we utilised to specifically identify the challenges of embedding e-Science applications that emerged from the eDiaMoND case study data.

4 Methodology

We conducted a retrospective analysis of the eDiaMoND project as a case study to identify and analyse challenges of embedding e-Science applications. A series of semi-structured interviews were conducted with project participants that focused on descriptions of collaborative activities within the project. Nine interviews were conducted and the data provided descriptions of events, as project members themselves understood them. Patterns where identified from their descriptions (Ackroyd *et al*, 1992: 155), which highlight the issues and challenges that emerged during the design process.

Semi-structured interviews were used to obtain qualitative data addressing a core set of questions, whilst allowing the interviewer and interviewee to explore unexpected issues. The interview framework allowed the interviewee to describe:

1. The eDiaMoND project
2. How the project aimed to embed e-Science applications
3. The issues that arose trying to embed e-Science applications

The complete interview framework can be found in the appendix.

Of course, such questions were answered from the perspective of the interviewee. Therefore, to get different views, a sample of key stakeholders were interviewed:

- A requirements engineer & ethnographer
- Two requirements engineers/software engineers
- A project co-investigator
- A software engineer/ethnographer
- Two software engineers
- The project manager
- Radiologist

Project documentation was also analysed including; the project proposal, requirements and project management documents.

In the following section we discuss the preliminary findings drawn from our retrospective analysis of the eDiaMoND project.

5 Preliminary Findings

A number of key challenges relating to embedding e-Science applications were identified from our analysis of the semi-structure interviews and eDiaMoND documentation. A list of the challenges identified is as follows:

1. Recalibrating the vision
2. Application developers as users of infrastructure
3. Working with constraints
4. Shared understandings between project partners
5. Sustainability

In the following sub-sections we shall introduce each of these challenges, present supportive evidence from our data, identifying the occurrence of these challenges and discuss each in turn reflecting upon its implications on the eDiaMoND project and the development of e-Science technologies in general.

5.1 Recalibrating the Vision

The Principle Investigator (PI) developed the project vision for the eDiaMoND project which, reflecting the e-Science vision, was to create a shared repository of digitised mammography images using grid infrastructure technologies and four core e-Science applications (see section 3). Through the use of the Grid and the four e-Science applications the eDiaMoND project was envisioned to facilitate the working practices of the BSUs.

However, there were a number of challenges in turning the initial vision into a reality. First, there were different interpretations of the vision between the project partners. Second, there were conflicts between the PI's perceived use of applications and the radiologist concerns of how those applications would affect their day-to-day work practices. We shall now discuss each of these sub-challenges in turn.

5.1.1 Interpreting the Project Vision

Many research projects start with a high-level overview of what they hope to achieve – a vision. The same is true of e-Science projects. However, due to the ambiguity of such a vision, it is often open to interpretation. This was certainly true of the eDiaMoND project. The academic partners were looking to produce innovative research, whilst the commercial partners were more inclined towards producing tangible deliverables. The ambiguity of the project vision fuelled conflict between project partners at the beginning of the project.

> "There was a lot of arguing at the beginning – What are we doing: a demonstrator; a proof-of-concept; or a production level system?"

A number of factors lead to misinterpretations and varying expectations between project partners. First, it can be hard to specify an e-Science project proposal in any detail. e-Science projects are multi-institutional and multi-disciplinary, involving many project partners. By the time the proposal has stated the partners involved, the project background and aims of the project, there is often little room to specify how the project aims are to be achieved, other than providing a high-level overview. This overview was presented within the *Project Proposal* and the *Critical Success Factors* documents. Therefore, in the case of eDiaMoND, it was uncertain whether the aim of the project was to develop a demonstrator; a proof-of-concept; or a production level system? The academic project participants considered the project from a purely technical aspect – developing a grid infrastructure and a number of exemplar applications that could demonstrate the use of the Grid. The commercial project participants considered the project as a proof-of-concept to the NHS BSU, which could lead to the development of a production level system. The perceived outcome of the eDiaMoND project – whether a demonstrator; a proof-of-concept; or a production level system – would impact the activities and enthusiasms of the project participants and their expectations for the project.

A second cause of misinterpretation is due to the "hype" that is put on project proposals. The Usability Task Force (UTF) (Grant No. GR/S79084/01) have argued that the hype around projects and associated expectations has put tremendous pressure on e-Science projects. In order to get funding from a research council the research must be highly innovative and contain a research component. In contrast to this requirement, to bring onboard commercial partners the project had to be able to attract commercial interests, specifically, the potential of offering a production-level deliverable or some other commercial gain such as media publicity. Due to the differing goals of project partners, expectations also differed. When different project partners have different goals and expectations, negotiations need to take place to develop shared understandings between project partners and establish a common vision/objective. In the case of the eDiaMoND project, this required various workshop meetings amongst the project partners. The result from these workshop meetings was a critical success factors and high-level requirements document. This document specified a technical direction and vision for the eDiaMoND project. It is crucial to the success of e-Science projects that these different expectations are managed (Lloyd & Simpson, 2005). For a detailed account of the Management Issues please see the corresponding report Project Management in eScience.

However, the critical success factors and high-level requirements document was only the start of establishing a shared vision between project partners. The shared vision also needed to be maintained. As the eDiaMoND project progressed new challenges emerged and design decisions had to be made. An excerpt from one of our interviews illustrates this.

> "Another large discussion was around the proof-of-concept. If you were actually trying to prove grid concepts to the breast screening radiologists in this country, then you needed to use real data. So there was a lot of discussion around that because the software engineers, for example, wanted to use 'tea-pots'."

In this example, the software engineers wished to use teapots (i.e. dummy data), yet the requirements engineers wished to use real data, i.e. mammograms with patient related data. Design decisions had an impact on all project partners and therefore required negotiation. Each of the project partners would argue their case to inform the final design decision made.

> "Our argument was, if we are trying to sell this to the breast screening unit your not going to sell it using either dummy data or mocked-up data."

Negotiation between project partners informed the subsequent direction of the eDiaMoND project. However, as stated previously, e-Science projects involve many stakeholders who are geographically dispersed around the globe. This makes it particularly difficult for project participants to meet and negotiate over design decisions. Such negotiations take time, which is an added pressure when project participants already work to tight deadlines and are dependent on one another's deliverables (see section 5.3).

It is crucial for e-Science project teams to work towards a shared vision (Lloyd & Simpson, 2005). Breakdowns between project partners can result in project failures. Differing expectations and goals between project partners' need to be managed. The various project partners need to agree on what they are going to do and how they are going to do it. Such collaboration is not just necessary at the outset of a project, but needs to occur as a continued activity. Overcoming the challenge of building and maintaining a shared understanding between project partners will facilitate the success of a project (Lloyd & Simpson, 2005).

5.1.2 The Project Vision and End-Users Needs

As previously mentioned, the PI developed the vision for the eDiaMoND project (see section 3). Through the use of a federated database of mammography images and four exemplar applications, the aim of this vision was to move from film based mammography images to digital mammography images. Overall, due to their recent procurement of PACS systems, this seemed a large step for the NHS breast screening programming at this time. However, there were also a number of differences between the PI's vision and the radiologists work practices. We shall now discuss each of these differences in turn.

First, the eDiaMoND vision promised the ability for mammography images to be read by any BSU from around the country, thereby distributing the workload of radiologists equally across the UK. However, radiologists raised a number of trust issues between unknown radiologists who might read mammography images from clinics other than their own. Mammography images differ from clinic to clinic due to differences in films, inks and equipment. Therefore, a radiologist may be unfamiliar with the practices of clinics other than their own. This could in turn lead to incorrect diagnoses. The following excerpt from one of our interviews illustrates this.

> "Great idea, but this is not how we work. If I am in London, I don't know your work practices, what are your standards?"

We see here how the vision for distributed reading is constrained by the radiologists current work practice. For the eDiaMoND vision to be realised, standards would have to be consistent between clinics – a difficult process to instigate in the light of the close sociality of their working practices of various BSUs (Jirotka et al 2005). Furthermore, trust relationships between radiologists from different clinics would have to be established.

Second, the eDiaMoND vision hoped to benefit epidemiologists by providing them with a larger cohort of patients to participant in their studies. It was envisioned that a federated database of mammography images with associated meta-data would benefit both the quantity and quality of epidemiological research. For example, epidemiologists might be able to search the federated database for women with breast cancer who drank more than ten units of alcohol. Then, if they wished, they could refine the search to include women who also smoked more than ten cigarettes a day. The Grid was intended to allow epidemiologists to perform more efficient longitudinal studies. However, the epidemiologists were concerned with issues around IPR. As epidemiologists invest a lot of time collecting their data they are very protective over it. Therefore, they do not want others capitialising on their hard work.

> "They [the epidemiologists] really didn't want to share or refused to share their data and put it on a federated database."

The epidemiologists wanted a separate set of grid technologies exclusively for their own use. This undermined the vision promoted by e-Science where researchers share their data providing a greater data set for researchers to search and use in their research. The challenge here is to promote positive perceptions of sharing data, over the negative implications around ownership. Overcoming this issue furthermore involves establishing trust relationships between the epidemiologists and the people who use their data. It also requires a cultural shift toward *open science* where the tensions between intellectual property rights, organisational boundaries, use of remote instruments and the sharing of data will become irrelevant issues for researchers and research programmes (Schroeder, 2006).

A further conflict between the eDiaMoND vision and the epidemiologists work practices was to do with anonymisation. The mammography images in the BSU had a two-way anonymisation with the anonoymisation key held by the NHS. This was necessary to maintain patient confidentiality. However, the epidemiologists required a record for their cohort as they conducted long-term studies. This would require the two-way anonymisation process to be broken, which was not feasible.

Finally, the eDiaMoND vision intended to implement 'find one like it' searching. It was envisioned that if a radiologist were unsure about an mammography image they were reading, they would be able to search for a number of other mammography images that were similar to the one they were viewing. It was believed that this would be able to help the radiologist inform their diagnosis. While the radiologists considered the 'find one like it' mammography image search mechanism interesting, they commented that it would not be useful in practice. An excerpt from one of our interviews illustrates this point.

> "They [the clinicians] showed for example how in reading an image it was necessary in a sense for the clinician to account for every kind of possible cancer that they might look at. They'd look around and where something was suspicious they would come to a kind of opinion and judgement about whether that was suspicious or not. It wasn't really clear how displaying 10 images that the computer thought was like it necessarily helped. As a consequence you're left saying, 'I just don't know how this is useful'".

When radiologists are in doubt they will always perform a biopsy so the similar images displayed on the system would be irrelevant in practice. From these reported findings, we see some differences between the vision for the eDiaMoND project and the end-users' needs and work practices. Such differences between a project's vision and the users needs and work packages can result in project failures (Grudin, 1994). Despite the differences it should be acknowledged that the flagship eDiaMoND project was aimed towards demonstrating the use of e-Science and grid technologies.

The ethnographic work conducted by the requirements engineering team identified some tensions between the project vision and the radiologists' work practices that could not be incorporated into the design of applications that were developed as part of the eDiaMoND project. Therefore, these tensions were represented as requirements for future systems, which formed part of the project's *Blueprint* document that was meant to provide a catalogue of lessons learned for the future development of e-Science technologies and applications. Similar to an argument made by Grudin (1994) for the CSCW community, now e-Science technologies and applications are becoming more established in researchers' working lives, we as a community need to make sure we take end-users' needs and work practices into account from the outset and throughout any e-Science project. This is a necessity to successfully embedding e-Science applications in the work practices of scientists, practitioners and researchers who use these technologies.

5.2 Application Developers as Users of Infrastructure

Many communities such as the Human-Computer Interaction (HCI) community have long put an emphasis on designing for end-users (e.g. Brooke *et al*, 1990; Preece *et al*, 2002). Brooke *et al* (1990) argues for the need to focus on the user, task and domain – the foci of the HCI community. It has been strongly argued that such a focus will lead to the development of more usable systems and applications (e.g. Preece *et al*, 2002).

An emphasis of the 'Embedding e-Science Applications' project aims to move beyond the user, task and domain foci to consider the wider working environment as argued by Grudin (1994). We are no longer designing for a single user sitting behind a desktop, working on a particular task, in a given domain. This is particularly true in e-Science. First, we are no longer designing for an end-users domain; rather we are designing for individuals and their organisations. Second, we are no longer designing applications exclusively in terms of tasks to be performed; rather we are designing for collaborative sharing and research activities. Finally, we are no longer designing exclusively for end-users; rather we are designing for developers, system administrators, virtual organisations, made up of different stakeholders and organisations with differing and sometimes conflicting practices and policies. Such complexities are present not only for the researchers who use e-Science applications, they also exist for the developers of these technologies. As such, we need to move beyond designing for end-users and design for all users. An excerpt from one of our interviewees illustrates this point.

> "Usability of who? I've got a grid API against which I had to program in order to get images over and so on. Was that usable for me as a user of grid technologies, probably not. I was also a user of some software produced by [the small industrial partner], which I had to build the workstation on and they had an API I had to go to... was that API usable to me? No, it's bloody awful for various reasons."

Here we can see that e-Science infrastructures and applications are comprised of various layers. In the eDiaMoND project the developers of applications had to use the middleware and APIs developers had available to them, which in turn had an effect on the design of applications that could be developed. We can see here that issues with embedding e-Science applications stem beyond the developers of applications and into the design of infrastructure.

In order to take into account the constraints faced by applications developers we need to move beyond designing for end-users. Rather, we need to bring together infrastructure developers, application developers and end-users when designing e-Science applications, thereby forming *constituencies of infrastructure*. Infrastructure developers, application developers and end-users should be equal partners when designing e-Science applications. As such, new techniques to facilitate co-design between developers and end-users should be examined across e-Science projects, examining how developers can converge on a solution that meets the end-users' needs, while working with the constraints faced by infrastructure and application developers.

Two approaches have been considered by Büscher *et al* (forthcoming) to facilitate co-design between project participants. First, identify application developers as users of infrastructure and involve them in design activities with infrastructure developers. This activity would allow the infrastructure to more closely meet the requirements for the application design. Hence, a tailored infrastructure would allow the application developers to address more closely the needs of the end-users. Applications developers from the eDiaMoND project who were interviewed expressed that such an approach would produce far more effective and efficient technical solutions, than current practice. The second approach, namely 'use experiments' (Büscher *et al,* forthcoming), involves the testing of infrastructure within a local setting where target end-users actually perform their work. These field studies document the initial successes and failures of the technology in use, as well as documenting the emerging practices that emerge from its long-term use. The documentation from use experiments provides descriptions of interactions with the infrastructure, applications and tools. These descriptions feed back into the design process in an iterative fashion. Both these approaches foster a convergence of 'architecture design, application design and experience of use' (Büscher *et al,* forthcoming) to support an evolving co-design practice.

5.3 Working With Constraints

The eDiaMoND project, like all software development projects, had a number of constraints. The main constraint for eDiaMoND was time. The project had a very short timescale of two years. Such a short timescale put immense pressure on project partners to meet all their deliverables. These deliverables fell into four "buckets" as described by one of our interviewees.

1. A research bucket – the process of conducting innovative research and producing publications.
2. A demonstrator bucket – the process of showing the potential of the grid
3. A prototype bucket – the process of developing grid technologies that could be used
4. A blueprint bucket – the process of filtering lessons learnt from the other buckets to inform the future design and development of grid technologies

In addition to deliverables that had to be completed in the limited time frame, there were also a number of external pressures as described by one of our interviewees.

> "I think there was a lot of pressure to show something and they [the project partners] were getting so much publicity and press releases. There was a lot of pressure to be able to demonstrate something – from the university, the e-Science programme and of course [the large industrial partner], especially the people who were reading about it… they wanted to see something."

These constraints meant that the project partners had to work in parallel in order to meet their deliverables and the mounting expectations. Figure 1 presents a diagram drawn by one of our interviewees to illustrate the timelines of the project partners throughout the eDiaMoND project[3].

[3] Figure 1 is a simplification of project partners' activities. The following points should be acknowledged: (a) project partners could involve multiple teams. For example, the lead university consisted of a requirements engineering team and a development team; (b) the four hospitals involved in

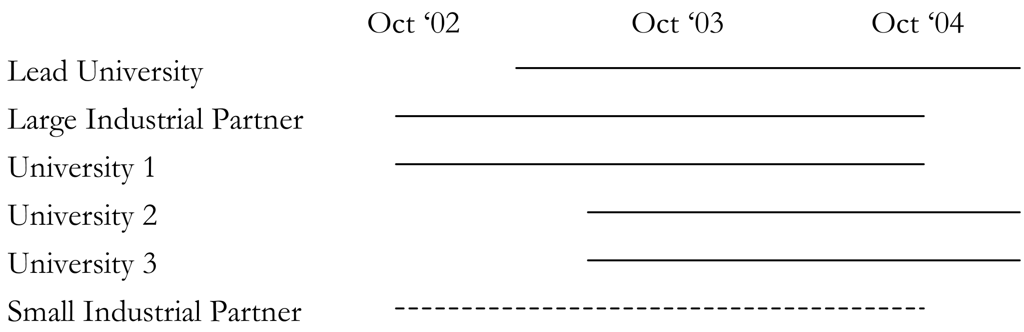

Figure 1. Project partners timelines throughout the eDiaMoND Project.

Due to time constraints and the waterfall development model used there was a strong focus on meeting the technical aims of the project – to demonstrate the potential use of e-Science and grid technologies in the medical domain of breast screening. Therefore, activities were broken down into a number of work packages. These work packages implicitly supported a waterfall software development process as illustrated from project plan summary shown in figure 2.

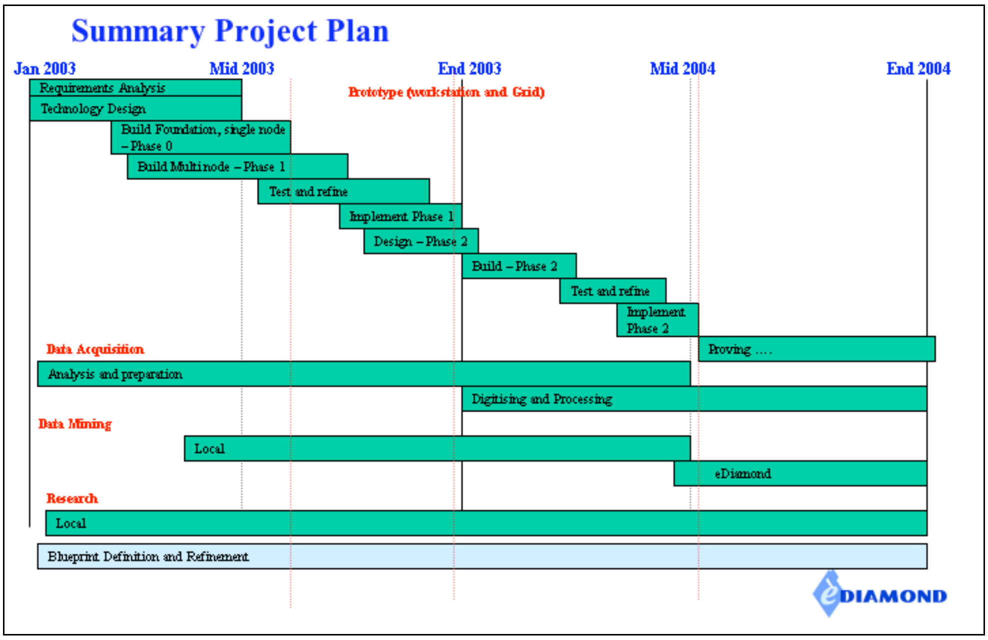

Figure 2. A summary of the eDiaMoND project plan.

The waterfall model is made up of five main stages: requirements engineering, design, implementation, verification and maintenance. The deliverables from one stage are passed onto the next and so the software development process continues, until a final solution is delivered. However, during the eDiaMoND project, some of these activities had to overlap one another. For example, the development of the grid infrastructure began before the requirements had been fully specified. An excerpt from one of our interviews highlights this.

---

the eDiaMoND project were linked to the universities; and (c) the small industrial partner went through some organizational changes during the eDiaMoND project, which affected their involvement.

> "[The large industrial partner] started Oct '02, we all started spring '03. So, it was a little bit back to front to be honest. You've got the team developing the underlying middleware here, five to six months before those that are developing any thing else. So, you start the process of 'defining requirements'. What are the requirements for the system? By the system we mean, the screening application and the underlying middleware. And you doing that at the same time they [the large industrial partner] already got the kit in place and their starting to write code."

Due to the projects time constraints it was not feasible for all the requirements elicited by the requirements engineering team to be integrated into the development of applications. The software engineers had to deliver a working demonstrator of the Grid for an All Hands Meeting (AHM). Therefore, they had to focus on the development of the original project vision as expressed in the *Critical Success Factors* and *High Level Requirements* documents. Changes in the original requirements would have caused delays in the development process, which could have meant the demonstrator would not have been ready for the AHM. If time had been permitting, it would have been advantageous for the eDiaMoND project to engage in an iterative software development process. As discussed in section 5.1, there were a number of differences between the eDiaMoND vision and the end-users' needs and work practices. It was during the ethnographic work conducted by the requirements engineers that the radiologists' needs and work practices were identified and represented as a set of requirements. During an iterative software development lifecycle these user requirements would have had an impact on the design decisions influencing the development of specific applications. Bergman *et al* (2002) describes this as a shift of project ownership *from* the project initiators *to* the end-users. If this shift had occurred, it would have allowed the support provided by the eDiaMoND system to more accurately match radiologists' work practices and culture. However, the process of re-design and re-implementation is very time consuming. Therefore, in the eDiaMoND project, rather than user requirements being iteratively integrated into the system, they made up part of the Blueprint document informing recommendations for the design and development of future e-Science technologies and applications.

5.4 Shared Understanding Between Project Partners

During the eDiaMoND project the requirements engineers built upon and conducted a number of ethnographic studies aimed at understanding radiologists work practices, to thereby eliciting requirements. However, the requirements engineers and the developers had different views on the need for the ethnographic work. The requirements engineers believed they should be looking at the entire work practices of the breast screening units.

> "You couldn't just do the workstation and the requirements for the screening unit. You had to have embedding… fitting into the whole breast screening unit itself and the administrative work of it, because the whole thing about, how a woman comes in, the paperwork that is taken around that, how the image is taken, how the images are then put up on the light boxes, how they are stored. I think the first thing that you discover when you do any kind of ethnographic work on patient records is that typically doctors tend to follow the career of an illness. This is very important."

However, the developers did not believe the requirements engineers should be doing this kind of ethnographic work.

> "They [the developers] didn't want to know how breast screening units worked. They [the large and small industrial partners] seriously did not want to. They didn't think that we ought to be looking at that kind of ethnographic work."

From the perspective of non-developer interviewees, the developers were not interested in knowing about radiologists work practices *per se*, rather they just wanted a specification from the requirements engineers describing how the system was to be built.

> "They [the developers] just wanted to be told, 'here's the spec'. Give us the stuff and we just go and build it."

However, this is not the way the requirements engineers had set up the studies for incorporation into the development. An excerpt from an interview with one of the requirements engineers illustrates this.

> "I said at the time [during an initial project workshop] that we needed to have iterative development, because you can't use ethnographic work without doing that. That was something that was never possible."

As discussed in section 5.3 the developers had a limited time frame in which to develop a working demonstrator for the AHM meeting. An iterative development process has the advantage of allowing for refinement, but would have been time consuming.

As an iterative development process was not possible, a lot of the richness of the ethnographic work that had been done was lost when passed onto the development teams. Rather than this work being integrated into the technologies under development, the outcomes of the work were disseminated through the Blueprint document for future reference.

A major reason for the loss of richness of the work conducted by the requirements engineers was due to developers not understanding the ethnographic work and the user-centred requirements. As one of the requirements engineers mentioned.

> "I don't know why they [the developers] couldn't understand them [the requirements], its not exactly written in CSP."

This information was not at the right level of granularity for the developers. The developers did not want to get involved with the users they were designing for and their users needs. Rather, they wanted to be given a specification telling them exactly what to build. One of the reasons for this mismatch between what was delivered and what was expected to be delivered was due to a lack of communication. The distributed nature of the eDiaMoND project meant that it was hard to interact with project partners. Hence, several of our interviewees emphasised that difficulties existed when passing deliverables between teams. This problem is not isolated to e-Science projects. There have long been issues in the process of translating users needs and expectations into specifications to be used by developers (O'Neill, 2000). In the eDiaMoND project, this transition between the requirements engineering phase and the development phase required the project manager to produce seven scenarios, which specified the information the developers needed. The requirements engineers and the project manager unofficially referred to the seven scenarios as "the seven deadly scenarios". They were given this name, because the requirements engineers and the project manager felt that the seven scenarios were an oversimplification and did not reflect the richness of the ethnographic work.

The understanding of the ethnographic work would have come from the iterative development between the requirements engineers and developers.

> "I've got that background [of ethnographic work and requirements], several people have got that background and can take this on [the ethnographic work and requirements]. But it is fairly time consuming. Because these guys [developers] are never going to learn, they're trained to do a particular job… they've got a certain skill set."

Such an iterative process between the requirements engineer and developers would have allowed a transition between the requirements and development phases. This transition should be a collaborative rather than an individual process.

> "There is a sense of 'whose responsibility is this to go this far' [to transition from requirements engineering phase to development phase]. Should the person doing the requirements capture go all the way down to the bullet points for the developer or should the developers meet them halfway and say, 'Yeah, I can see how we can turn that into a bullet point'. Whose responsibility is it to do that transition work from one to another?"

As well as there being a lack of time for this iterative process to occur, there was also a lack of communication, resulting in misunderstandings between the requirements engineers and developers. The requirements engineers intended to engage in an iterative development process with the developers. Whereas the developers wished to have a design specification at the right level of granularity from which to development from.

> "We could have done scenarios. We could have done use-cases. We could have done what ever they wanted, if we had the interaction with them at the time. It wasn't clear to us. We didn't find out till much later that this was an issue. We didn't even know. We didn't see them. We didn't talk to them."

Overcoming the barriers of communication is a major challenge facing many e-Science projects.

Two factors constricted requirements engineers' efforts to involve users in the design of the system. These were the software development model, which framed design activities in sequential order and a development culture that prefers to design from *requirement specification documents* and not *design in context* using ethnographic fieldwork data and iterative co-design activities with end-users. Even though the project was positioned in this manner and requirements engineers did engage in these usability practices, their impact had a limited effect on design

Also, the framing of the project within an implicit waterfall software development model and tendency to favour a development culture privileged the project's pre-defined technical goals as specified in the *High Level Requirements* document over participatory usability practices created a situation that produced a development focused, not a user focused environment for design. Because of this, usability findings could be incorporated only if they improved upon pre-defined technical goals, thereby creating a situation in which usability practices and efforts were reduced to a passive role that provided insight for the *Blueprint* document, rather than informed design decisions for the development of the e-Science technologies and applications.

The developers and the requirements engineers also had divergent outlooks on the value of different knowledge artefacts (e.g. specification documentation, video analysis, etc). Goguen (1992) referred to these two cultures as 'the dry' - consisting of a preference for formalised abstractions that represent system functionality; and 'the wet' - consisting of a preference for understanding the day-to-day work practices of individuals and organisations to understand the social factors that will determine a system's success. Describing the differences between the two cultures and its effect on managing requirements for usability is an issue that may be solved by replacing the waterfall method with more Agile software development methods. For example, where the waterfall model suggests deliverables be passed between stakeholders, such as a requirements document being passed from a requirement engineering team to a group of developers, Agile methods suggest a closer collaboration between different stakeholders, thereby facilitating the transfer of tacit knowledge. Therefore, in e-Science projects Agile methods might promote the transfer of the requirements engineers understanding of end-users working practices and organisational factors to developers.

5.5. Sustainability

As well as the challenges faced during the design and development of e-Science technologies and applications discussed so far in this report, there is also the challenge of sustaining e-Science. If e-Science is to realise its full potential, we as developers must also understand how to support researchers interactions with e-Science technologies and applications, and understand how to embed these tools successfully into the researchers everyday working lives. This in turn calls upon the need to provide a sustainable infrastructure for users to have confidence that adopting these new tools will indeed be worth their while. For example, on the eDiaMoND project, once the project had ended the system was dismantled. This means that the end-users of the eDiaMoND project had been investing both time and effort assisting the development of the tools to be delivered to them, only to have them taken away at the end of the project. If e-Science technologies are to be successful, these tools have to be sustained, thus allowing them to become a core part of researchers' everyday working practices.

However, sustaining e-Science tools in researchers' working lives is currently inhibited. Sustainability is inhibited due to the necessary funds required to support and maintain e-Science technologies not being available. Research councils currently fund most e-Science developments, yet while the research councils fund the development and innovative research that is conducted during e-Science projects, they will not fund ongoing support and maintenance costs. Hence, once a project ends the developed e-Science technologies have to be dismantled.

While most e-Science technologies and applications are not available beyond the funding period of a project, the Open Middleware Infrastructure Institute (OMII) and National Grid Services (NGS) are two services available in the UK working towards the sustained future of the e-Science community. The remit of the OMII and NGS is to provide mechanisms for both developers and researchers to access e-Science tools and services. An avenue of future research for this project is to understand how researchers access such tools and services to support their everyday working practices and how this facilitates those technologies being embedded in their everyday working practices.

6 Conclusions and Future Work

In this report we have identified and explored some of the challenges of embedding e-Science applications. First, we set the scene by arguing the need for embedding e-Science applications into a research environment. We achieved this by building upon research findings from the CSCW community identifying the challenges of understanding end-users work practices and the need for changing software development cultures. Second, we presented our case study project – the eDiaMoND project: a flagship e-Science project that gave rise to many challenges and issues that were new to the project participants of the eDiaMoND project, the e-Science community and the research councils. Third, we presented our methodology to identify and explore challenges specific to embedding e-Science applications. The methodology we applied to the eDiaMoND project was a retrospective analysis, which consisted of a number of interviews with various stakeholders from the eDiaMoND project, the findings of which were triangulated with the project documentation. Finally, we presented and discussed five core challenges identified from our preliminary findings:

1. Recalibrating the vision
2. Application developers as users of infrastructure
3. Working with constraints
4. Shared understanding between project partners
5. Sustainability

We introduced each of the challenges, provide supporting evidence and discuss each challenge in turn, reflecting upon their implications on the eDiaMoND project and e-Science in general. We shall now summarise each of our findings below.

Like many research projects, a project starts with a project vision – usually the principle investigators' vision. However, there are a number of challenges turning this vision into a reality. First, due to the ambiguity of a project's vision, it can be open to interpretation. Hence, project partners with differing drivers for their involvement in the project establish differing expectations. Such expectations need to be managed and project partners need to negotiate to establish a shared vision for the project. Furthermore, this shared vision needs to be maintained throughout the project. The establishment and maintenance of this shared vision is essential to the success of the project (Lloyd & Simpson, 2005). However, it is particularly challenging in e-Science projects to establish and maintain a shared vision due to the large-scale, global distribution and multi-disciplines of project participants on e-Science projects. For example, the large-scale and distributed nature of e-Science projects often results in difficulties setting up and coordinating meetings between project participants. Second, there can be mismatches between the project vision and end-users' needs and expectations. It is important that the project vision evolves to meet end-users' needs and expectations by allowing a shift in project ownership to occur from project initiators to the end-users (Bergman *et al*, 2002) who will in the end, be the one's using the technology on a daily basis. As argued by Grudin (1994), in order for a technology to be successful integrated into end-users working lives, it is essential to take their working practices into account from the outset of a project.

However, a further challenge for e-Science projects is that we are no longer developing exclusively for end-users. Application developers are also users of infrastructure, as they are constrained by the middleware, APIs and other technologies. Therefore, the challenge of embedding e-Science applications not only involves understanding end-users' work practices, it also involves developing an understanding of the constraints faced by those developing the e-Science technologies and applications. Supporting the widening of the design community to include infrastructure developers, application developers and end-users, and conducting use experiments, will foster the convergence of 'architecture design, application design and experience of use' (Büscher *et al,* forthcoming) to support an evolving co-design practice in e-Science.

Constraints play a prominent challenge in e-Science projects. A major constraint is time. e-Science projects normally have a relatively short period of time to design and develop their technical solutions. This often requires project participants to work in parallel. However, this is particularly difficult in e-Science projects due to the large-scale and distributed nature of e-Science projects, which makes communication and coordination very difficult. Furthermore, the pressures of time are exacerbated do to the volume of media attention e-Science projects received.

There is the challenge of facilitating effective collaboration between project partners. Project participants are from different institutions and disciplines and each has their own culture. This can result in collaborative breakdowns between project partners (Fischer, 1999). To resolve this challenge there needs to be effective communication between project partners. However, as previously mentioned, this is a challenge for e-Science projects in its own right, were communication between project partners can be particular challenging due to the scale and distributed nature of e-Science projects.

Finally, there is the challenge of sustainability. If e-Science applications are not sustained they will not be adopted by their end-users and embedded in their working practices. Therefore, e-Science projects are faced with the challenge of how to sustain their e-Science applications. The sustainability of e-Science applications usually falls beyond the remit of the research councils. Therefore, we as a community must consider how e-Science applications can surpass the lifetime of a project.

Based on the exploration of these five challenges, below we suggest a number of preliminary recommendations for e-Science projects in order to facilitate embedding e-Science applications.

1. The project vision needs to evolve to meet end-users' needs and their organisation's policies. This can be achieved through an iterative process between project stakeholders and end-users.
2. Infrastructure developers, application developers and end-users need to co-design e-Science applications, identifying how to meet the needs of end-users, whilst working within the constraints imposed upon the project.
3. Agile methods can be used to spread the delivery of milestones over time, thereby reducing time pressures. Furthermore, media policies can make sure e-Science project team give a clear and consistent message, thereby avoiding false expectations of e-Science applications to be delivered.
4. Encourage project participants to go beyond their normal sphere of activity collaborating closely with other project participants to facilitate the transfer of tacit knowledge between project partners.

5. Mechanisms need to be put in place to ensure the sustainability, use and maintenance of e-Science applications.

In order to build upon these challenges and recommendations we have identified a number of potential future studies. These should not be viewed as a finite list of studies, but rather a list of possibilities that shall evolve to meet the needs of the e-Science community and our research findings. These studies include:

1. We aim to expand upon our retrospective analysis of the eDiaMoND project performing a number of concurrent analyses of ongoing e-Science projects. The aim of these case studies is to validate and expand upon our preliminary findings and recommendations. More specifically we wish to understand how best to harmonise the introduction of new systems into pre-existing work, cultural and organisational practices in order to lessen disruption and increase the potential for uptake. To facilitate this research, the software development process and in particular, requirements and design activities will be examined to understand how they support an understanding of a domain's context including the researchers work and cultural practices, and the organisations internal and external policies. In some cases we wish to make recommendations to improve the development process and observe the impact of our recommendations.
2. We wish to perform a diary study of the use of communication/collaborative technologies. It is hoped this diary study will increase our understanding of how project participants communicate and what technologies they use and for what purpose.
3. We are conducting a number of in-depth studies into the use of tools to support distributed collaboration. Such tools include Polycom PVX supporting multi-participant desktop collaboration; and AccessGrid supporting multi-participant room based collaboration.
4. Furthermore, moving beyond interactions with technologies to support explicit collaboration between project participants, we wish to study implicit collaborations through e-Science infrastructures, such as national grid services (NGS). With regards to the NGS, we would also like to explore the challenges surrounding the sustainability of e-Science applications.

Through these and similar studies we aim to understand how we can best support the embedding of e-Science applications.

Acknowledgements

We would like to thank all our interviewees who participated in this study. Furthermore, we would also like to thank the EPSRC for funding this research (Grant No: EP/D049733/1).

Further Reading

de la Flor, G., Lloyd, S., Jirotka, M. & Warr, A. (2007). Embedding e-Science Applications: Designing and Managing for Usability. Proc. of e-Social Science.

Warr, A., Lloyd, S., Jirotka, M., de la Flor, G., Schroeder, R. & Rahman, M. (2007). Project Management in e-Science. Technical Report, Oxford University.

Warr, A., de la Flor, G., Jirotka, M. & Lloyd, S. (2007). Usability in e-Science: The eDiaMoND Case Study, In the workshop, 'Increasing the impact of usability work in software development' at CHI 2007, 28 April - 03 May 2007, San Jose, USA.

References

Ackroyd, S., Hughes, J. (1992). Data Collection in Context, 2nd edition. Longman Group.

Bergman, M., King, J. L. and Lyytinen, K. (2002). Large Scale Requirements Analysis as Heterogeneous Engineering. Scandinavian Journal of Information Systems, 14, 37-55.

Brooke, J., Bevan, N., Brigham, F., Harker, S. and Youmans, D. (1990). Usability statements and standardisation - work in progress in ISO. Proc. of INTERACT'90, 357-361.

Büscher, M., Christensen, M., Hansen, K., Mogensen, P. & Shapiro, D. (Forthcoming) Bottom-up, top-down? Connecting software architecture design with use.

Fischer, G. (1999). Symmetry of Ignorance, Social Creativity, and Meta-Design. Proc. of Creativity and Cognition (C&C), 116-123.

Foster, I., Kesselman, C. & Tuecke, S. (2001). The Anatomy of the Grid: Enabling Scalable Virtual Organizations. International Journal of Supercomputer Applications, 15(3), 200-222.

Goguen, J. (1992). The Dry and the Wet. Information Systems Concepts: Improving the Understanding. Proceedings of IFIP Working Group 8.1 Conference. Alexandra, Egypt. Elsevier Holland: 1-17.

Grudin, J. (1994). Groupware and Social Dynamics: Eight Challenges for Developers, Communications of the ACM, 37 (1), 92-105.

Gutwin, C. & Greenberg, S. (2000). The Mechanics of Collaboration: Developing Low Cost Usability Evaluation Methods for Shared Workspaces. Proc. of IEEE WETICE: Infrastructures for Collaborative Enterprises.

Hey, T. & Trefethen, A. E. (2002). The UK e-Science Core Program and the Grid. Future Generation Computer Systems, 18, 1017-1031.

Hey, T. & Trefethen, A. (2003). e-Science and its Implications. Philosophical Transactions of the Royal Society, 361, 1809-1825.

Jirotka, M., Procter, R., Hartswood, M., Slack, R., Simpson, A., Coopmans, C., Hinds, C. & Voss, A. (2005). Collaboration and Trust in Healthcare Innovation: The eDiaMoND Case Study. Proc. of Computer Supported Cooperative Work (CSCW), 14, 369-398.

Lloyd, S. & Simpson, A. (2005) Project Management in Multi-Disciplinary Collaborative Research. Proc. of IEEE International Professional Communications Conference, 602-611.

O'Neill, E. J. (2000). User-developer cooperation in software development: building common ground and usable systems. Springer Verlag, London.

Preece, J., Rogers, Y. & Sharp, H. (2002). Interaction Design: Beyond Human-Computer Interaction. Wiley.

Schroeder, R. (2006). e-Research infrastructures and open science: a paradigm shift in knowledge production? Oxford Internet Institute. Oxford.

Appendix

The semi-structure interview framework was as follows:

1) Describe the e-Science project?
   a. Organisations involved?
   b. People involved?
   c. The product under development?
   d. The software development process?

2) Did you adopt usability work in the e-Science project?
   a. If yes:
      i. What were the decision factors for adopting usability?
      ii. What usability practices did you adopt?
      iii. What problems were addressed?
      iv. What results were intended?
      v. What was the impact of the usability work in the e-Science project?
   b. If no:
      i. Why not?

3) What was the overall plan for usability work?
   a. What actions were taken?
   b. Where did these actions occur in the software development process?
   c. What were the outcomes of these actions?
   d. How did you define measurable targets for usability?
   e. How were these targets measured?
   f. Who were the key people influencing the usability work?
      i. What were their roles and contributions?
      ii. How did other people experience the outcome of the usability work?
   g. What was the realised impact from the usability work on the software development process?
   h. What were the key challenges?
      i. How were they resolved?
      ii. What unresolved problems remained?

Author Bibliographies

Andrew Warr is a researcher with interests in Human-Computer Interaction (HCI), Computer Support Cooperative Work (CSCW), Design and Creativity. Andrew is currently a Usability Researcher Fellow at the Oxford e-Research Centre (OeRC), University of Oxford. He is working with Dr. Marina Jirotka (PI) on the EPSRC funded 'Embedding e-Science Applications - Designing and Managing for Usability' project. Prior to this project, Andrew studied for his PhD in the HCI Group at the University of Bath, supervised by Dr. Eamonn O'Neill. His PhD research focused on 'Understanding and Supporting Creativity in Design'. During this time Andrew was also a Visiting Researcher at the Centre for LifeLong Learning and Design (L3D) at the University of Colorado, Boulder working with Prof. Gerhard Fischer and his colleagues. In addition to these research activities, Andrew also has strong interests in designing, developing and evaluating mobile and pervasive technologies for social settings.

Grace de la Flor is currently a DPhil student at the Oxford University Computing Laboratory studying the ways in which e-Science can change and improve the working practices of researchers and scientists. Her current research will assess how best to design and evaluate e-Science systems in support of the complex ways in which science happens; from 'in silico' experimentation, data sharing and visualization to new knowledge generation.

Marina Jirotka is Director of the Oxford Centre for Requirements and Foundations, Lecturer in Requirements Engineering at Oxford University Computing Laboratory and Fellow of St Cross, Associate Director of the Oxford e-Research Centre (OeRC), and James Martin Research Fellow at the e-Horizons Institute. She has degrees in Social Anthropology and Psychology and Computer Science and Artificial Intelligence. She received her DPhil in Context Driven Requirements Capture from the University of Oxford Computing Laboratory in 2000. Her main areas of research have been developing novel methods and techniques for requirements capture drawn from the social sciences, focussing on workplace studies, practice driven requirements and design elicitation. Her research interests have long been concerned with developing systems that support everyday work and interaction, primarily to bring a richer comprehension of socially-organised work practice into the process of engineering technological systems. She developed the Requirements Engineering module that forms part of the MSc in Software Engineering and the MSc in Computer Science at Oxford University. In recent years her research has focussed on requirements for e-Science applications, particularly e-Health and she has worked on various industrially funded research projects, DTI, ESRC and EPSRC projects. She has been the requirements analyst on flagship e-Science projects such as eDiaMoND and is advisor on requirements methods and techniques for e-Science projects such as Integrative Biology and Virtual Research Environments for Humanities. She is a member of the UK e-Science Usability Task Force, Co-Director of the Oxford e-Social Science node that is investigating the Ethical, Legal and Institutional Dynamics of Grid Enabled e-Sciences and is Principal Investigator of the EPSRC project Embedding e-Science Applications: Designing and Managing for Usability.

Sharon Lloyd is a project manager and research facilitator at Oxford University. Sharon was project manager for the 2-year £4 million eDiaMoND project led out of Oxford in collaboration with IBM and 10 partner sites. She is now project manager for both Integrative Biology and NeuroGrid projects, which are led from Oxford, working with over 120 people over 25 sites. Before joining the University in February 2003, she worked in industry for 17 years where she gained extensive experience in project management in a commercial environment as well as people management through her management of IT teams. She has experience of both scientific and commercial IT development projects both in the UK and overseas.

Ralph Schroeder is a James Martin Research Fellow at the Oxford Internet Institute (OII). Before coming to the OII, Ralph Schroeder was Professor in the School of Technology Management and Economics at Chalmers University in Gothenburg. He completed his PhD at the LSE in 1988. He has written extensively about virtual environments and recently completed a book manuscript 'Rethinking Science, Technology and Social Change'. At the OII he is continuing his research about how people interact in shared virtual environments. He is writing a book on this topic that will link virtual reality technology to other types of computer-mediated communication and new media.

Mustafizur Rahman interests focus on Information and Communications Technologies (ICT), particularly in translating and realising non-technical social aspects of Requirements Engineering into technical realities, for applications in advanced collaborative technologies in distributed computing and Grid technologies. Working closely with the Oxford e-Research Centre and the Centre for Requirements and Foundations (CRF), he has worked on several multidisciplinary projects, including CyberSEM (remote microscopy) and IMaGE (Law/IPR and e-Health), and has been involved in the requirements aspect of the eDiaMoND breast cancer screening research project as a Principal Researcher at the CRF. As the ICT Programme Manager the Oxford University Begbroke Science Park, he is tasked with fostering collaborative university-industry applied R&D, training and education and technology transfer activities.

Mr Rahman received his MSc in Software Engineering at Oxford University (UK) after completing his BSc in Computer Systems Engineering at the University of Massachusetts at Amherst (USA). He has spent a number of years working on research projects at universities before working at multinational companies such as Digital, Compaq and IBM. He was also one of the core members at a start-up company specialising in distributed systems management. He worked for over two years on a DARPA sponsored research project on Quality of Service before joining Oxford University in March 2003.

Authors Contact Information

For general enquires please or write us following the details below:

The Embedding e-Science Applications Project
Oxford e-Research Centre
7 Keble Road
Oxford
OX1 3QG
UK

Telephone: +44 (0) 1865 610600

Fax: +44 (0) 1865 610612

E-mail: oerc-usability@mailinglist.ox.ac.uk

Website: http://www.oesc.ox.ac.uk/usability