Usability Research Challenges in e-Science

Prepared by the UK e-Science Usability Task Force

Introduction

As the technologies mature usability issues are starting to emerge as a major challenge for the e-Science community as a whole. Although offering considerable potential benefits, the complexity of e-Science technologies and the architecture of the underpinning infrastructure represent a significant barrier for scientists wishing to exploit e-Science technologies within their sphere of scientific interest. If a future e-Science infrastructure is to become part of the everyday life of scientific work it needs to take into account the needs of its future user communities or else it runs the risk of not being used or not being exploited to its full potential.

Understanding the user concerns associated with the development and deployment of a broad e-Science infrastructure requires us to adopt a broad interpretation of usability. In particular, user issues in e-Science will require us to:

- *Involve a broad range of disciplines* (e.g. psychologists, economists, social scientists and philosophers) to address these issues from a multidisciplinary perspective.
- *Consider a broad set of user perspectives* including those who undertake scientific research through an e-Science infrastructure, support staff who ensure the smooth running of an e-Science infrastructure, developers who build new e-Science facilities on top of the infrastructure, those responsible for setting and administering ethical research guidelines and, finally, the general public who may be expected to participate in e-Science research by virtue of being its subjects.

If e-Science tools and techniques are to make the transition from concept demonstrators to viable e-Science systems, supporting real scientific communities, a number of significant usability research challenges need to be addressed. The development and effective deployment of e-Science systems require us to consider a variety of distinct research challenges, including four areas in particular:

- *Global Communities.* How do we maximise the use of e-Science technologies and applications to support new forms of scientific community?
- *Trust and Ethics.* How do we handle the ethical and policy issues to emerge from the use of an e-Science infrastructure?
- *Knowledge Production.* How do we exploit an e-Science infrastructure and techniques to support scientists' expertise, new research methods and new forms of knowledge production?
- *Design, Assessment and Management.* How can we best assess e-Science technologies and applications, and use this assessment to guide the design and management of these Systems?

Tackling these research questions outlines a usability research agenda that will require significant progress to be made in terms of new methods, tools and techniques to support those who design, develop and use e-Science systems. This research will have significant benefits for the e-Science community including:

– Improving the effectiveness of tools that represent and support the work of scientists.

– Improving the ability of e-Science systems to support the cooperative work of globally distributed scientists.

– Improving the effectiveness of tools that represent and support the work of scientists.

– Increasing the level of trust we can have in a research process underpinned by e-Science systems.

– Increasing the effectiveness of e-Science researchers through the development of new interface techniques and representations of scientific information.

– Developing an understanding of the ethical issues within e-Science and the development of approaches to the management these issues.

– Developing sustainable models of adoption and use which support and encourage the uptake of e-Science facilities.

– Developing new techniques and facilities to support the configuration and management of e-Science facilities to meet the needs of scientists.

Fundamental user centred research will need to be undertaken to understand how an e-Science system may actually be used and how the technology should be designed in order to encourage use. In the rest of this document we outline four key usability research challenges that need to be tackled to ensure that the vision of e-science becomes an everyday reality for the scientists who wish to exploit the benefits it offers.

New forms of Global Scientific Community

Global e-Science systems seek to enhance and foster global collaborations among researchers. This laudable ambition however presents a series of significant user centred design challenges. Many of these challenges build upon existing research in computer-supported collaborative working (CSCW) by seeking to understand what new forms of scientific community will be supported by global e-Science systems. This is a strongly multidisciplinary endeavour requiring a close cooperation between researchers seeking to gain an understanding of scientific community and the developers of e-Science technologies.

New methods and techniques to understand community

Understanding new forms of e-Science communities and conveying these understandings to help form e-Science systems requires new methods and techniques. Research needs to be undertaken to understand globally distributed scientific communities as they exploit a distributed infrastructure. Key research questions to emerge in this domain include:

- *How do we understand the dynamics of a distributed e-Science community* when its work is manifest both physically within the lab and digitally through the infrastructure, and what are the most appropriate techniques for conveying our understanding of these communities' interactions?
- *How are e-Science facilities and tools used in practice* and how do scientists manage the relationship between scientific activities and the recording of these activities within e-Science systems? For example, what is acceptable in terms of record and how do scientists exploit the record reading and writing skills within a broader scientific community of practice?
- *How might an e-Science system capture a community's activities* in order to support e-Science and make these activities available to others? What are the methods and techniques to capture these activities in a lightweight manner?

Fostering new forms of e-Science community

The widespread use of an e-Science infrastructure provides the potential for new forms of e-Science community to emerge supported through e-Science systems. Promoting the formation of these new communities raises a number of key research questions that combine both an understanding of users and the development of new technological mechanisms. Key research issues include:

- *How do we identify and overcome the barriers* to the formation of e-Science communities be they organisational, cultural and/or technical? From these understandings, how can recommendations best be developed for those who would seek to encourage the growth of the e-Science communities?
- *What are the key opportunities offered by the formation of e-Science communities,* how might we identify these opportunities and encourage the formation of community structures that maximise the potential benefits of e-Science communities?
- *What are the flexible mechanisms and techniques needed* to support the formation of dynamic e-Science communities, and how do we interleave technological mechanisms with broader human and social techniques for the management of access to shared resources?

Strategies for embedding e-Science facilities

A key question in the formation of a global scientific community is how will e-Science systems become embedded within the working lives of scientists located within multiple sites and disciplines of scientific practice? Strategies need to be developed to promote the effective integration of new e-Science technologies with existing work practices of scientists and the needs of the organisations of which they are members, while allowing the space for their reinvention as the possibilities afforded by the technology are explored. Understanding the dynamics of this process and providing guidance on appropriate models of innovation and adoption raise a number of research questions include:

- *How will e-Science systems be presented to promote a close fit to existing scientific practice*? One solution would be to present e-Science systems in a form that is familiar, thereby exhibiting a recognisable compatibility with existing practices and routines. How might these e-Science facilities map to existing techniques and technologies and how do we promote migration of scientific practice?

– *How will e-Science systems be designed to promote innovations in scientific practice*? What are the key methodological issues for e-Science research and what are their principal drivers? How do we create and manage design and development processes that are sufficiently agile so as to be able to support emergent research practices?

– *How will e-Science systems be designed and deployed to promote a close fit with organisational goals*? How might they fit with existing IT infrastructures and facilities, and what are the organisation issues in establishing this fit? How might e-Science facilities be designed to promote ease of installation, compatibility and configurability with respect to organisational infrastructure and policies, and its maintainability in the face of ongoing organisational, and technological change and development.

Trust and ethics in e-Science

If an e-Science system is to be accepted, both by scientists and the general public, then it must be trusted by the users of the system and those whose medical or other personal data is entrusted to it. The considerable investment in security mechanisms in the development of an e-Science infrastructure needs to be complemented by a user centred understanding of the issues of trust and ethics to emerge as e-Science systems are deployed.

Support of trust across global e-Science communities

Potential users of e-Science services have expressed significant concerns with issues around trust, whether this was trust in the infrastructure and its ability to reliably deliver high levels of performance, or trust in the service to provide secure data transmission of commercially sensitive material. Trust is also a matter of developing confidence in the decisions of others. It is about trusting the information, and the knowledge and skills that one's collaborators possess, whether one is co-located with them or working remotely. It is about believing that the processes that lead to their decisions are well founded and well informed. In the collaborative context of e-Science, with its mix of human and computational systems a fundamental issue will be understanding how people develop a trust in the capabilities –and an awareness of the limitations – of humans and machines making up an eScience System Key research questions that need to be addressed to promote the development and maintenance of trust and awareness include:

– *What are the perceptions of trust and accountability across e-Science systems and what is the impact on the use of these systems*? Research into such perceptions and their effects would include seeking to understand how trust is articulated and communicated in e-Science communities and how the distributed nature of e-Science communities might impact the formation of "cultures of trust."

– *How do we evolve work practices to promote trust in e-Science Systems?* Addressing this challenge would include seeking to understand how best to promote a "culture of trust" in the activities of others and in the resources made available through an e-Science system.

– *How do we design for different levels of accountability and responsibility needed to promote trust*? This activity includes understanding how issues of trust feed this into the design of e-Science tools and guidelines.

Ethics in e-Science

Within e-Science, there is a difficult balance to strike between acquiring information regarding scientific research that is in the public interest and may improve scientific discoveries in general, and protecting ordinary citizens from the unscrupulous use of their data. If systems are designed too tightly, they may become rigid, inflexible and difficult to use at the practical level. If they are too loose, they become vulnerable to abuse and may fall afoul of ethics committees, thus ruling out their use in research. Key research issues that need to be tackled to help us strike this balance include:

– *How do we understand ethical practice in e-Science and what support can be provided for scientists*? This may include developing practical tools for researchers to help them plan and conduct research projects employing personal data based on an understanding of what constitutes good ethical practice. For example, is it possible to assess reliably the risks to individuals' confidentiality incurred from linking datasets? Equally, it may involve determining good practice guidelines for gaining ethical approval for e-Science projects.

- *What are the ethical barriers to the use of e-Science techniques*? This requires us to understand and identifying issues that may severely constrain or indeed potentially obstruct the successful completion of e-Science activities in order to bring such issues to the attention of policy makers. Do policy makers and members of research ethics committees understand the security and confidentiality provisions of e-Science systems?
- *How might scientist best manage the various ethical constraints in e-Science?* In this case, we need to understand especially how ethical constraints impact on the lived work of clinicians and scientists with a view to determining how changes in ethical or legal procedures will affect the e-Scientist. Will new ethical debates and dilemmas occur with the rise of e-Science?
- *How do we best manage the issues of* distributed access, disclosure and anonymity in large scale data repositories? The most promising areas of e-Science research involve integrating and comparing data within large-scale repositories often derived from different initial resources, and often across disciplines and organisations. This mix brings about significant scientific, organisational and technical challenges that need to be identified and addressed by the e-Science systems and approaches.
- *What kinds of awareness raising activities might be useful for educating the general public and achieving greater acceptability of e-Science*? Medical and social science research projects generally depend upon the general public's willingness to be recruited as subjects and to give permission for the use of confidential data. Researchers may also wish to make use of subject data routinely collected by public and private sector organisations.

Knowledge production and expertise in e-Science

The production of new forms of knowledge is a central feature of e-Science systems with a globally distributed team of researchers working together to undertake research leading to new knowledge assets. A key feature is the interplay between the production of knowledge and the expertise held by the e-Science community exploiting the distributed e-Science systems. A number of areas require investigation:

Understanding the knowledge lifecycle in e-Science

Given the central role of knowledge production in e-Science it is essential that we gain a user centred understanding of the knowledge lifecycle in e-Science and the role of users in the production, consumption and maintenances of this knowledge. Key research questions that need to be considered include:

- *How do we understand the production of new forms of knowledge in e-Science*? The global access to digital resources promoted by e-Science systems allows the production of new forms of knowledge that are expressed and represented digitally. It also allows provenance information to emerge as a knowledge resource in its own right. What is unclear is what sort of activities will be undertaken by users to generate these new forms of knowledge and how these may be best understood and recorded.
- *What new forms of reasoning are enabled by e-Science*? Globally distributed e-Science systems will make digital information available in an unprecedented manner. It will also promote accessibility to tools to help scientists analyse and reason about this information. How these facilities will be used, what new forms of reasoning will emerge when they are supported by these tools, and the impact for future e-Science systems all need to be investigated.

Representation of information, knowledge and expertise in e-Science

The use of e-Science systems raises many challenging issues concerning the design and usability of representations of information, knowledge and expertise, especially when distributed among large, interdisciplinary and potentially distributed scientific teams. The challenges include the design of novel visualisation techniques. These include tools to support the management of multiple forms of representation and the mappings between them, support for the evolution of new forms of symbolic notations by scientists for communication and collaboration between them. A number of questions are significant for research in this domain.

- *How do we design appropriate representations for e-Science*? Designing representations for e-Science requires us to investigate the utility of different representations across the variety of

potential users within e-Science. This exploration will include understanding how e-Science facilities and scientific information are best presented to Scientists. What are the novel forms of representation needed by e-Science given the challenges of scale involved? How do we handle interactivity and shared representations?

– *How do we understand the use of representations in e-Science?* Addressing this question includes investigations into how the emerging representational forms such as ontological structures are used in practice, and what of the implications are for adopting these emerging structures. We will also need to explore the various ways in which these structures are interpreted by users, and especially how these representations mediate the process of scientific discovery.

Understanding the future role of records in e-Science

The emergence of e-Science systems will have a significant impact on the production of the scientific record and how it is used to support the scientific process. Many more elements of scientific practice will be recordable and this record may be accessed remotely across a distributed infrastructure. Understanding the implications of transforming the scientific record will require us to understand issues of sharing and disclosure, the extent to which scientific activities can be captured, and the ways in which scientific records are read and understood by the scientific communities they support.

Design, Assessment and Management in Global e-Science Systems

A significant set of usability challenges are centred on the design assessment and eventual management of e-Science systems. Essentially new methods and techniques are needed to promote a user centred approach to the design and development of technologies in e-Science systems. Key areas of investigation include:

Design of e-Science systems

The design of e-Science systems such that they meet the demands of the user communities they support is a multidisciplinary endeavour that needs to be understood and supported in order to ensure that future e-Science systems emerge that are informed by the needs of users. Key research issues that need to be tackled include:

– *What are the most appropriate design methods, approaches and techniques for e-Science systems*? How do we best gain an understanding of the needs of users and convey these to design and development teams when these development teams come from a range of different technical and scientific backgrounds. How do we undertake this design in a timely and costly manner and what are the best strategies for designing these systems?

– *What methods are appropriate for understanding scientific practices and what are the best ways of presenting these understandings to designers*? What level of user involvement in the design and development of e-Science systems is appropriate and achievable in practice? Do approaches such as ethnography, which emphasise detailed investigation, scale up? Is there a need for 'hybrids', i.e., people with both research and technical skills? Is so, how can this be encouraged and rewarded professionally?

– *What sort of tools can be used to support the design of e-Science systems?* The development of tools to support designers will allow us to significantly reduce the cost of design and deployment of e-Science systems. These tools might range from frameworks that aid in the process of designing e-Science systems to computational facilities that support particular design activities. These tools need to be developed in close partnership with those responsible for the realisation of e-Science facilities.

Assessment of e-Science Systems

The assessment of e-Science systems represents a major challenge for existing evaluation techniques. It is not clear that our existing approaches are up to the task of assessing globally distributed large scale e-Science systems. Key research challenges that need to be addressed include:

– *What are the most appropriate evaluation methods, approaches and techniques for e-Science systems*? How do we develop approaches to assess large-scale, distributed e-Science systems that will help guide their development? How might we evaluate the effectiveness of collaboration and the collaborative technologies to support it? How do we handle the long timescales underpinning

many scientific investigations? How do we handle the social and cultural aspects involved in assessing e-Science systems that support distributed communities of users?

– *How might we understand and assess interaction with e-Science artefacts*? In addition to understanding how best to assess the broad utility of e-Science systems we need to consider how best to assess the interaction users will have with the various artefacts made available via e-Science systems. How do we make this interaction as transparent as possible and encourage the use of these artefacts to be as productive as possible?

– *What tools can be developed to support the assessment of e-Science systems?* These tools may range form frameworks and guidelines to support the process of assessment to tools such as tracking and logging tools that can be used to help capture the use of the system for subsequent analysis.

Management of e-Science systems

The scale and complexity of e-Science systems represent a major management challenge. This challenge is manifest both in terms of managing the technological infrastructure underpinning the e-Science systems and in managing the activities of research teams supported through the e-Science systems. Two key research questions dominate our consideration of user centred views of management:

- *How might we support the management of e-Science infrastructures*? The current generation of e-Science infrastructures are already causing significant management challenges for those responsible for them. How might we understand the best ways to manage these infrastructures and provide tools and techniques to aid the systems managers responsible for ensure the correct operation of e-Science infrastructures?.
- *How might we manage large scales collaborative research, and how can we support this management*? How do we allocate effort and responsibility and support auditing; accountability? How might the work involved in a scientific investigation best be decomposed into operational subcomponents (who does it, how is it done? How to we manage projects at crucial points in their development such as start/finish? How can we achieve this without producing huge problems of information management?