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

| Document Name | | | |
|---|---|---|---|
| **Document Title** | *Deliverable D1.1, Typology of Barriers Report* | | |
| **Reporting Period** | 1 April 2007 – 30 June 2009 | | |
| **Author(s) & project role** | Alex Voss (project manager)<br>Rob Procter (principal investigator)<br>Marzieh Asgari-Targhi (researcher) | | |
| **Date** | 17.07.2009 | **Filename** | D1.1TypologyReport.pdf |
| **URL** | *http://www.engage.ac.uk/e-uptake/e-uptake-deliverables* | | |
| **Access** | ☐ Project and JISC internal | X General dissemination | |

| Document History | | |
|---|---|---|
| **Version** | **Date** | **Comments** |
| 0.1 | 06.11.2007 | Initial version for project meeting |
| 0.2 | 05.06.2009 | Updated version |
| 1.0 | 17.07.2009 | Final Version for JISC |

# ENABLING UPTAKE OF E-INFRASTRUCTURE SERVICES

# TYPOLOGY OF FINDINGS REPORT

## 1 Introduction

The purpose of this report is to give an overview of the types of findings the Enabling Wider Uptake of Infrastructure Services project (e-Uptake) has produced through its empirical work. It is meant to serve as a guide and index to the more comprehensive Community Engagement report[1] (Deliverable D1.2) and the database of findings[2], which contains the most comprehensive set of descriptions of issues identified.

This report can be used either as a guide to the fieldwork material produced by e-Uptake or as a way to structure data collected as part of future community engagement and research projects.

## 2 Methodology

The methodology used to collect data is described in the Community Engagement report and for the sake of brevity we will describe only those parts of relevance to this report. The typology was initially developed from a list of inhibitors and enablers found described in existing documents as well as an initial set of pilot interviews. We conducted two phases of interviews, both of which were transcribed and analysed. During the analysis, we initially used a free coding scheme and resulting codes were then compared to the existing typology. Where possible we reconciled the codes used by applying those already used in the typology. Where this was not possible, we extended the typology, so that over time we ensured that it reflected our fieldwork data. We found that, by and large, the initial typology worked well to structure the fieldwork data and make it accessible. Where necessary it was adapted by adding or removing categories.

We should point out that there are potentially many ways to categorise findings, so the typology should be seen as a way to make the findings accessible and not as an analytical device that would allow a quantitative analysis based on aggregations. Also, where a finding related to more than one category, we have assigned it to more than one in order to facilitate effective browsing and retrieval.

As the typology contains a relatively large number of items, we have organised it hierarchically to facilitate browsing. Figure 1 shows a graphical representation of some of the top levels of the typology. The main categories in the typology are social issues, technical issues and issues to do with the management and use of digital resources. They break down into further subcategories. We have tried to limit the number of categories as it can easily grow beyond the point where a typology would be manageable and useful. Nevertheless, the wealth of findings produced from the interview material has resulted in the definition of 15 higher level categories that contain a further 105 sub-categories.

---

[1] http://www.engage.ac.uk/e-uptake/e-uptake-deliverables

[2] http://cnx.org/content/col10673/latest/


*Figure 1: Excerpt from the typology of findings*

## 3 Typology of Findings

In the following, we present the outline structure of the typology of findings and give a short one-sentence description of each category.

**1 Training, Education and Outreach**

The ways that people learn about e-Infrastructure, acquire basic competencies and specific skills are important factors that can inhibit or enable uptake of e-Infrastructure services.

1.1 Understanding Distributed Systems

It is important for researchers but also for other stakeholders to have at least a basic knowledge of distributed computing principles.

1.2 Understanding the Potential of e-Research

Researchers and other stakeholders need to understand and be able to reason about the potential and the limitations of e-Research practices.

1.3 Domain-Specific Material and Events

Providing an understanding of e-Research principles and providing relevant skills is best achieved by tailoring content to the specific audiences addressed, e.g., through domain-specific material and events.

1.4 Quality of Training Material

Training material should be complete, up to date, correct and well written with the specific intended audience in mind.

1.5 Early Engagement and Outreach
The earliest experiences with e-Infrastructure services or with the concept of e-Research can influence uptake in important ways as it is at this stage that decisions about future engagement and commitments are made that can be difficult to revise.

1.6 Success Stories
The development of success stories can help to engage researchers and provide an understanding of the potential of e-Research as well as the costs involved and the paths to successful adoption.

1.7 Use Cases
Detailed use cases can provide an understanding of real-world uses of e-Research practices and e-Infrastructures that can help to overcome the initial hurdles of making sense of technological options and ways to utilise them.

1.8 Paths to Adoption
Training, education and outreach activities need to be aligned with clearly defined paths to adoption that effectively link up different stages of engagement from initial interest to project formulation, making technical choices, acquiring specific skills, the development of appropriate socio-technical configurations and, finally, routine usage.

## 2 User Relations with Technology Supply Processes

The relationship between those involved in technology supply processes (from initial innovation, through development, deployment and support) and their potential and actual users has an important influence on the fit between technological options and the requirements of users and user communities.

2.1 Understanding Research Practice
In order to understand and formulate requirements that can adequately inform the development of technologies for a specific community or researchers, designers need to gain a working understanding of the relevant research practices.

2.2 Understanding Available Options
In order to participate in a technology supply process as well as to understand and effectively communicate their requirements, researchers need to gain an understanding of what is technically possible and what is not as well as what options can be practically implemented.

2.3 Hybrid Knowledge
People with knowledge both of technologies and relevant applications domains can serve an important function in the design process.

2.4 Social Organisation
The way that the collaboration with users is socially organised has important implications for the technology supply process and its outcomes.

2.5 Understanding Infrastructure
Both those involved in technology supply and researchers need to understand that working infrastructures cannot be ‘built’ but need to be ‘fostered’ as they are not merely technical artefacts but complex assemblages of technologies and social arrangements around their deployment and use.

2.6 Priorities
In the light of resource constraints and the need for researchers to push the

boundaries of what is possible, the definition of clear priorities becomes particularly important.

2.7 Conflicting Research Agendas and Commitments
As technology innovators are often researchers themselves (e.g., from computer science departments), their research agenda often conflicts with their commitments to produce working technologies.

2.8 Reward Structures
The wider environment in which technology innovation and development takes place defines the reward structures that influence the decisions made by both designers and users of technologies.

2.9 Software Engineering Practice
Software engineering is still beset by a number of problems that can only partially be addressed and that lead many development efforts to fail, to be only partially successful or to achieve the original aims only at large expense and after a longer period of time than anticipated.

## 3 Collaboration

e-Research often involves collaborations between individuals from different backgrounds and from different institutions.

3.1 Finding Collaborators
Funding suitable collaborators for a research project is often a first and crucial step followed by the scoping of the collaboration and the definition of roles.

3.2 Data Sharing
Collaboration in research entails the sharing of data among a defined but potentially open set of participants.

3.3 Coordination
The working division of labour within a research collaboration needs to be coordinated to achieve common as well as individual aims.

3.4 Communication
Collaborating parties need to find ways to communicate effectively and efficiently.

3.5 Size of Collaborations
The size and composition of collaborations can have important implications for the social organisation of work and for the effectiveness of common projects.

3.6 Geographical Distribution
e-Research collaborations often span large geographical distances, leading to problems associated with remote collaborations, especially where distances are large and span time zones.

3.7 Communities and Ad-Hoc Collaboration
Not all contexts for collaboration are well-defined and static – ad-hoc collaborations and collaborations within open communities are just as much features of e-Research practices as formal collaborations within funded projects.

3.8 Collaboration Readiness
e-Research collaborations are underpinned by at least a partial alignment of interest for at least a period of time as well as the motivation and ability of partners to act collaboratively.

3.9 Accountability
In collaborations it is important that people can account for the actions of others and it is necessary for partners to be accountable for their activities.

3.10 Mediation
There is sometimes the need for mediation in collaborations, when different interests collide or other problems occur.

3.11 Trust
Collaboration is based on a sufficient level of trust between partners and can be undermined by a lack of it.

## 4 Policy and Funding

The wider policy context and the funding mechanisms available to support e-Research practices play an important role in shaping uptake.

4.1 Research Assessment
The assessment of research quality, e.g., for the allocation of funding to HEIs, influences the aims and objectives researchers will choose to pursue.

4.2 Data Curation and Sharing
Decisions made at a policy level by funding organisations have an impact on the ways in which researchers create, manage, curate and share data.

4.3 Models of Innovation
Policies often implicitly or explicitly contain a model of innovation that influences how people conduct e-Research, e.g., a technology-push model is not well suited to foster engagement of potential users of technologies.

4.4 Measuring Impact
Researchers need to measure the impact of their activities and the impact of the use of e-Research technologies needs to be visible and sufficiently high to warrant the investments made.

4.5 Priorities
Given the limits imposed by resourcing, researchers need to make decisions about the use of e-Research technologies that can sometimes lead to non-adoption because of short term priorities.

4.6 Political Influences
Events in the political sphere can sometimes influence the adoption of e-Research practices, e.g., the Alder Hey scandal set back data sharing practices in biomedical research for years.

4.7 Justifying Costs of Resources or Adoption of Services
The costs of adopting e-Research practices need to be covered by research grants or other funds and need to be justified by a projected outcome.

4.8 Delays in Gaining Access to Services
Gaining access to services can sometimes take time because or peer-review processes or bureaucratic procedures involved or for other reasons.

## 5 Disciplinary Factors

The way that research disciplines are organised, their internal structure, traditions, career paths, reward structures, etc. influence the adoption of e-Research practices.

5.1 Epistemic Traditions
The ways that knowledge in a discipline is traditionally created and how it becomes accepted define whether e-Research practices lead to outcomes that are immediately recognisable as ‘good research’ or not.

5.2 Career Structures
Investments in uptake of e-Research practices need to result in a clearly defined career progression, otherwise researchers will be systematically discouraged from engaging.

5.3 Collaboration
The traditional ways in which research in a discipline is conducted individually or collectively defines what one might call the 'baseline collaboration readiness' within a discipline.

5.4 Scholarly communications
Traditional ways of publishing and otherwise communicating knowledge that are accepted in a discipline also form an important background condition for e-Research practices.

5.5 Research Programmes
e-Research is often about establishing new directions in research as well as new methods and as such can be difficult to get funded.

5.6 Multidisciplinarity
Collaborations in e-Research often involve partners from different research traditions working in fields that are relatively less established, leading to the problems associated with multidisciplinary endeavours.

5.7 Silo-ed Communities
In some research disciplines communities can be relatively separate from each other with little history of collaboration or intention to change this, thus making it difficult to establish e-Research practices.

## 6 Individual Attitudes and Behaviours

The personal background, the habits, abilities, hope, aims and fears of individuals can influence their willingness or ability to take up e-Research practices.

6.1 Resistance to Change
Established ways of working can give rise to a resistance to change that can cause researchers to not adopt e-Research practices even if there are clear benefits that would outweigh the costs of adoption.

6.2 Uncertainty and Fear
The outcome of engaging with e-Research is to some extent uncertain and as e-Research can involve quite radical departures from traditional ways of conducting research, this can give rise to fears about the outcome and possible negative impacts on researchers' careers.

6.3 Career Choices
As the adoption of e-Research practices often involves a significant investment of effort and time, it is a strategic move that needs to be considered in the light of career choices made by the individual researcher but also by people working on the technologies, on service provision or in a support role.

## 7 Institutional Issues

The institutional context in which e-Research is embedded influences individual researchers' ability to engage in e-Research as it can both give rise to barriers as well as providing crucial support.

7.1 Rate of Change
How organisations respond to technological and social changes and how they manage to provide an environment in which individual researchers can

effectively plan their work and careers can inhibit or enable uptake of e-Research practices.

7.2 Relationships
The social organisation of research within institutions, including its management structures and the nature of the relationship between research groups and research support services can inhibit or enable e-Research practices.

7.3 Financial
The financial resources that institutions commit to supporting e-Research practices and the model used to fund local services and support have an influence on the uptake of e-Research practices.

7.4 Provisions for Curation
Institutions often provide only limited support for data management and curation of research outputs as the latter is often limited to publications, hence other solutions often need to be found to ensure the sustainability of other research outputs.

7.5 Labour Market
The ability of institutions to recruit and retain key staff with the necessary expertise to effectively support or conduct research using e-Infrastructures and e-Research methods is important for the uptake and sustainability of e-Research practices.

7.6 Support
Researchers often require IT support that goes beyond the level of support routinely provided by institutions through their normal information services departments.

7.7 Access to Services
The ability to access services can be limited by lack of availability of local resources or by technical constrains imposed such as restrictive network configurations.

7.8 Institutional Context and Culture
The wider institutional context and culture influences uptake of e-Research practices in many ways, e.g., by encouraging or discouraging multidisciplinary work

## 8 Ethical and Legal Issues

The use of personal data, of resources owned by other parties and the generation of intellectual property in e-Research can raise important ethical and legal issues.

8.1 Data Protection
The need to comply with data protection and privacy legislation can cause problems that can make e-Research projects difficult or even impossible to conduct; expertise and effort are required to conduct research in ways that is legally and ethically sound and the use of advanced ICTs raises questions that may not have been adequately addressed in legal practice.

8.2 Licensing
Researchers often make use of licensed material – data or code – that come with licensing conditions that have not been devised with distributed computing environments and collaborative practices in mind, causing issues with the management of licenses.

8.3 Intellectual Property Rights

Collaborative research can give rise to problems in managing the ownership of intellectual property – either in the use of material owned by third parties or in terms of ownership of the rights where intellectual property is produced.

## 9 Project Management

The management of e-Research projects poses challenges that go beyond the problems associated with project management in other areas, partly due to the nature or research and due to the structure of the academic system.

9.1 Methodology

Project management methodologies developed in the commercial sector may or may not be suitable for e-Research projects depending on the specific circumstances.

9.2 Managing Consortia

e-Research is often conducted by consortia of different institutions that are formed for the purposes of a specific project, are therefore time limited and have to reconcile different interests to achieve an overall project aim.

9.3 Managing Expectations

Expectations at the start of an e-Research project may be shaped by the vision formulated in a grant application and need to be managed throughout the project to avoid partners being disenfranchised when benefits take longer than anticipated to materialise.

9.4 Presentation / Public Engagement

Presenting the outcomes of research to the wider public is an important aspect of academic work that is often neglected or difficult to do effectively; e-Research both poses challenges and provides opportunities in this respect.

## 10 Crossing Boundaries

Working across organisational, geographical and disciplinary boundaries poses challenges that need to be addressed; e-Research technologies have partly been developed to address these issues.

10.1 Identity Management

Establishing mechanisms to manage identities across organisational boundaries is a challenge usually addressed using a public key infrastructure with certification and registration authorities, a solution that is separate from institutions' existing identity management solutions and that many researchers find difficult to use.

10.2 VO Management

Managing the sharing of resources requires hierarchical authorisation structures to scale; this is addressed through the concept of a virtual organisation but the management and effective usage of VOs can cause problems.

10.3 Communities

While VOs allow for the management of collaborations within relatively well-defined groups, more flexible mechanisms are sometimes needed to allow for collaboration within wider communities.

10.4 Collaboration with Partners Outside Academia

Many of the solutions that make up e-Infrastructures for research are implemented in academic institutions and collaboration with partners outside academia can be hindered by factors such as licensing or funding conditions.

## 11 Infrastructures

The technical and social arrangements that make up e-Infrastructures for research are the fundamental basis for e-Research practices to scale beyond local uses of resources and their characteristics will inevitably influence their uptake.

11.1 Middleware Maturity

Grid middleware in use today have not seen the level of investment into their development that would make them robust products; they are often still under development and have been used and tested for a limited time.

11.2 Heterogeneity

Part of the idea of e-Research is to link heterogeneous resources in a way that enables them to be used through common mechanisms but the heterogeneity of resources and, crucially, of different middleware stacks are often visible to the user.

11.3 Capacity

Some e-Research projects have resource requirements beyond the capacity currently provided by the infrastructure.

11.4 Support

The provision of e-Infrastructures needs to be complemented by effective and routinely available support structures that researchers can rely on, ideally at a local as well as national level.

11.5 Operating Environment

The wider operating environment within which e-Infrastructures operate can influence the quality and level of service experienced by researchers.

11.6 Role of Computing Services

Computing services as the immediate local contacts and intermediaries need to have a defined role in supporting research e-Infrastructures and need to be integrated with other forms of support provision.

11.7 Security Restrictions

Restrictions placed on the use of resources for security reasons can sometimes inhibit use of e-Infrastructures or make their use less effective.

11.8 Digital Divide

Differences in the provision of services and support can be observed between different institutions but also within institutions between different departments.

11.9 Dependability

The infrastructure used by researchers needs to fulfil a number of dependability criteria such as availability, security, usability, etc.

11.10 Rate of Change

Changes in the infrastructure can cause a misalignment of arrangements that can disrupt usage and therefore such changes need to be carefully managed and made accountable to users and those facilitating usage.

11.11 Ease of Access

Resources should be easy to access by researchers in a way that does not require many steps or complicated mechanisms.

11.12 Scale

The scale of the e-Infrastructures developed can cause issues with their effective coordination and management.

## 12 Applications

Research applications used to conduct e-Research are of immediate interest to researchers as they form part of the research methods used.

12.1 License Management

The practical aspects of license management can complicate the usage of applications within e-Infrastructures.

12.2 Availability

Applications codes suitable to be used in distributed environments need to exist and be available for researchers.

12.3 Cost of Implementation

Developing application codes, ensuring they are robust and deliver adequate and correct results requires significant effort and therefore funding.

12.4 Common Platforms

In order to distribute the costs of implementing research applications and to ensure their technical robustness, it is often desirable to use common platforms as a basis on top of which the specific functionality required can be built.

12.5 Domain Standards

One way to produce common platforms and to ensure that different applications can interoperate is to define application standards, which will normally be domain specific.

12.6 Quality

Many application codes are developed by academics, often PhD students and are used within the context of an individual researcher's practice or in a small group; many codes lack the quality attributes that would make them usable in other contexts.

12.7 Usability

Usability is an important non-functional requirement for the wider uptake of application codes.

12.8 Scoping / Functionality

When developing applications, important decisions need to be made about their scope and the functionality to be included; often, developing functionality over time yields better results than a 'grand design' but this is not necessarily so in all circumstances.

12.9 Compatibility

Being able to combine different applications to achieve a new overall effect requires that the codes used are compatible at some level and can be woven into an overall scientific workflow.

## 13 Standardisation and Alignment

The use of and alignment with standards has the potential to realise economies of scale and synergies as well as providing a degree of assurance of technical maturity.

13.1 Reverse Salients

Individual aspects of an evolving infrastructure can hold up its overall development; these are often called 'reverse salients'.

13.2 Lock-In

Standards can be widely adopted and continue to be used despite being superseded by a superior successor or alternative; this is called a 'lock in' situation.

13.3 Adoption
Standards have to achieve a critical level of adoption to become effective and to realise the benefits of network effects, where the utility of a standards increases exponentially with its adoption.

13.4 Compliance
Standards are often implemented only partially or imperfectly, causing problems with interoperability and reducing the quality of implementations.

13.5 Certification
In order to assure potential users of technologies that they comply with relevant standards, e.g., interoperability or quality standards, certification mechanisms are often used.

13.6 Timeliness
Standardisation can take a long time and there is a danger that by the time a standard is released and gains adoption, the state of the art has moved on and the standard becomes obsolete; on the other hand, a standard can be finalised too early before important developments have taken place.

13.7 Quality
The technical and editorial quality of a standard influences the degree of adoption and compliance and therefore has a direct impact on technological development.

13.8 Evolution
Standards often evolve through a number of revisions; the management of such transitions needs to be carefully managed to reduce the costs of moving from one version to the next.

13.9 Identifying Generic Functionality
Standards are most useful when they are widely applicable but identifying functionality that is generic across a number of domains and can be implemented to still serve specific needs is difficult.

## 14 State of the Art

In some areas, the state of the art has to catch up with the problems that e-Research poses, i.e., the solutions to problems require substantive research.

14.1 Algorithms
The development of algorithms that scale well to a large number of processing cores and work well on distributed systems is still a significant challenge.

14.2 Managing Distributed Systems
The challenges of managing distributed systems need to be met by the development of effective mechanisms to visualise and investigate distributed interacting processes.

14.3 Software Engineering
Distributed systems engineering is still a relatively young field and the development of such systems in a dependable manner is taxing the methods developed within the field of software engineering.

## 15 Digital Resources

In order for distributed resources to be reusable, they need to be produced and curated in a way that addresses a number of key concerns.

15.1 Discovery
Making digital resources discoverable requires the development of

appropriate metadata, the development of repositories, of federation mechanisms and of appropriate user interfaces for searching and browsing.

15.2 Metadata
The metadata associated with data needs to be based on agreed definitions, needs to be of sufficiently high quality and complete.

15.3 Curation
Digital resources need to be managed over time through curation processes that take into consideration aspects such as appraisal, selection, ingestion, preservation actions, storage, access, use, transformation into new forms as well as disposal.

15.4 Sustainability
Digital resources become less useful over time if they are not actively curated and maintained, so sustainability does not involve mere storage but has to be active if the value is to be maintained.

15.5 Quality
Digital resources can be of widely differing quality and completeness; in some research areas resources tend to be well-defined and of demonstrable quality while in others the question of their quality is more pertinent.

15.6 Access
Re-use of resources can only thrive if they are easily accessible through mechanisms that are routinely used by researchers.

15.7 Semantics
Without an understanding of the semantics of digital resources, re-use is impossible; often, the semantics are implicit, limiting re-use to those researchers involved in the creation of a resource.