MANCHESTER 1824
The University of Manchester

The University of Manchester Research

Paths to Wider Adoption of e-Infrastructure Services

Link to publication record in Manchester Research Explorer

Citation for published version (APA):
Voss, A., Asgari-Targhi, M., Procter, R., Halfpenny, P., Dunn, S., Fragkouli, E., Anderson, S., Hughes, L., Fergusson, D., & Atkinson, M. (2008). Paths to Wider Adoption of e-Infrastructure Services. In *Oxford e-Research Conference*

Published in:
Oxford e-Research Conference

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

OPEN ACCESS

Download date:03. Nov. 2024

Paths to Wider Adoption of e-Infrastructure Services

Alex Voss[1], Marzieh Asgari-Targhi[1], Rob Procter[1], Peter Halfpenny[1],
Stuart Dunn[2], Elpiniki Fragkouli[2], Sheila Anderson[2], Lorna Hughes[2],
Dave Fergusson[3], Elizabeth van der Meer[3], Malcolm Atkinson[3]

[1] National Centre for e-Social Science, University of Manchester
[2] Arts & Humanities e-Science Support Centre, King's College London
[3] National e-Science Centre, University of Edinburgh
Contact author: Alex Voss (alex.voss@ncess.ac.uk)

Abstract: this paper presents work conducted as part of the e-Uptake project, which aims to widen the uptake of e-Infrastructure services for research. We will discuss our fieldwork conducted so far, give examples of the barriers and enablers identified and discuss how using the accumulated knowledge can lead to paving the way for wider adoption of e-Infrastructure Services.

Introduction

Led by the U.K. e-Science programme, investments in e-Infrastructures for research over the past seven years have helped to develop distributed, networked and interoperable computing and data resources that underpin an increasingly wide range of research activities across all disciplines. In the UK, the Research Councils and JISC have funded a number of services that provide resources to researchers building on technologies and capabilities developed by the e-Science Programme. These services provide generic compute and data resources (e.g., through the National Grid Service), more specialised research services (e.g., through the National Centre for Text Mining), support for collaboration in research (e.g., through the Access Grid Support Centre), advisory and support services (e.g., the Digital Curation Centre or the UK Grid Support Centre) as well as basic infrastructural services for identity management (through the UK Access Management Federation and the UK e-Science Certification Authority).

As technologies mature and the provision of these services becomes more routine, questions about uptake and embedding of e-Infrastructures in the day-to-day working practices of researchers come to the fore. Indeed, one may argue that if these issues are not addressed, the e-Science community will not realise its full potential and will not achieve sustainability. Consequently, funders are complementing their investments in e-Infrastructures through active programmes of outreach and community engagement (Voss *et al*. 2007) in order to ensure that service provision is informed by actual needs of researchers and that researchers are aware of and informed about the services provided.

JISC's Community Engagement Strand

JISC's e-Infrastructure programme comprises a strand of three complementary projects aimed at widening the uptake of e-Infrastructure services and e-Research practices in the UK: e-Uptake, eIUS and ENGAGE (Voss *et al*. 2007). All three are currently working to establish a body of evidence about the level of uptake of e-Infrastructure services in the UK, the barriers that researchers encounter, the enablers that might widen and deepen uptake as well as examples of usage that can demonstrate good practice or

provide inspiring illustrations of successful e-Research practices. In addition, the projects have intervention elements such as training provision or consultancy and development. Together with the initiatives by service providers and some other funded activities, the projects represent a major part of the UK's response to the problem of widening and deepening the uptake of e-Infrastructure.

To date, the projects have identified over 400 individual researchers who are using e-Infrastructure services in the UK and who would be potential candidate respondents for interviews and surveys. They are collaborating closely under a common framework of understanding, including a common consent process and a data sharing agreement, to gather evidence from research communities and to develop activities to widen the uptake of services. In the following, we wish to focus on work conducted as part of the e-Uptake project, a collaboration between the ESRC National Centre for e-Social Science (NCeSS) at the University of Manchester, the Arts & Humanities e-Science Support Centre (AHeSSC) at King's College London and the National e-Science Centre (NeSC) at the University of Edinburgh.

The e-Uptake Project

e-Uptake aims to develop a broad empirical basis for understanding barriers and enablers for the adoption of e-Infrastructure as it needs to look beyond isolated, contingent or random problems that people have encountered in employing e-Infrastructure services. Rather, we seek to identify recurring, widespread barriers that can be overcome by a set of targeted interventions. The project will make these interventions or suggest strategies that might be followed up by e-Infrastructure stakeholders. Furthermore, the study must reflect the diversity of the target population (research active members of the UK academic community), their different interests and possible uses of the services (from the Access Grid Support Centre to the National Grid Service) and the number of potential factors influencing uptake (from individual practices to organisational factors and wider research policy). It is important that we sample not just the views of early adopters but also those of people who have not yet engaged with e-Infrastructure services so we can understand the factors underlying decisions concerning when – and whether – to use e-Infrastructure. In addition, the information gathered from academic end users needs to be contrasted with the views held by service providers and technology providers as well as intermediaries such as application developers, e-Science centres and academic hosting institutions.

Research Approach

The e-Uptake project is currently in the process of conducting interviews with UK-based researchers across all discipline areas. As it is practically impossible to establish a definitive set of respondents from which to sample *a priori*, we are using an iterative approach, starting with an initial set of interviews of researchers who we were able to determine had some experience with the use of e-Infrastructures in their research. Candidate respondents were identified using a combination of web searches, use of existing databases such as the UK research councils' online databases of grants and web mining. We have found that it was relatively easy to compile long lists of candidates but that filtering them using our selection criteria (active in research and using at least one of the JISC-funded services) has involved a large amount of manual work to compile the required information from publicly available data.

Each respondent is asked to fill in a short online questionnaire providing us with some baseline information about their research, their role, level of ICT use, institutional support as well as their use of JISC-funded services. The questionnaire data is used to arrange a telephone interview[1] with the respondent that will normally be recorded, transcribed and analysed by the research team. In this, we adopt a grounded approach (*cf*. Strauss 1987) in order to ensure that our ongoing investigation of barriers and enablers is informed by our growing understanding, so that we explore issues emerging from the analysis in more depth where appropriate and develop appropriate interventions. We have discussed our research approach in more detail in Voss *et al*. (2008).

To date, we have conducted about 50 interviews with researchers from a wide range of discipline areas, yielding about 25 hours of recorded audio data, which have been fully transcribed. We have stratified our sample of respondents to ensure that we have representation from across the range of disciplines funded by the UK research councils. Figure 1 shows the distribution for the initial sample. Please note that the categorisations used here are those given by respondents in the questionnaire rather than being categories assigned by the researchers. Our sampling in following rounds will aim to address the slight imbalance evident in the first round.

In addition to the primary stratification by research disciplines, we are also aware of the fact that other dimensions will be relevant as they influence the kinds of barriers that researchers face and the ways in which they react to them. For example, researchers at different stages of their careers may have different interests, attitudes towards technological and methodological innovation and skills sets, as well as investments in standard methods and tools. Such factors may have a marked influence on decisions about adoption. Consequently, we are aiming to ensure that our sample includes respondents of different levels of seniority.

[1] We do conduct face-to-face interviews where appropriate, especially in cases where we conduct interviews in collaboration with the eIUS project since our aim is to minimise the number of approaches made while maximising the benefits of each interview.

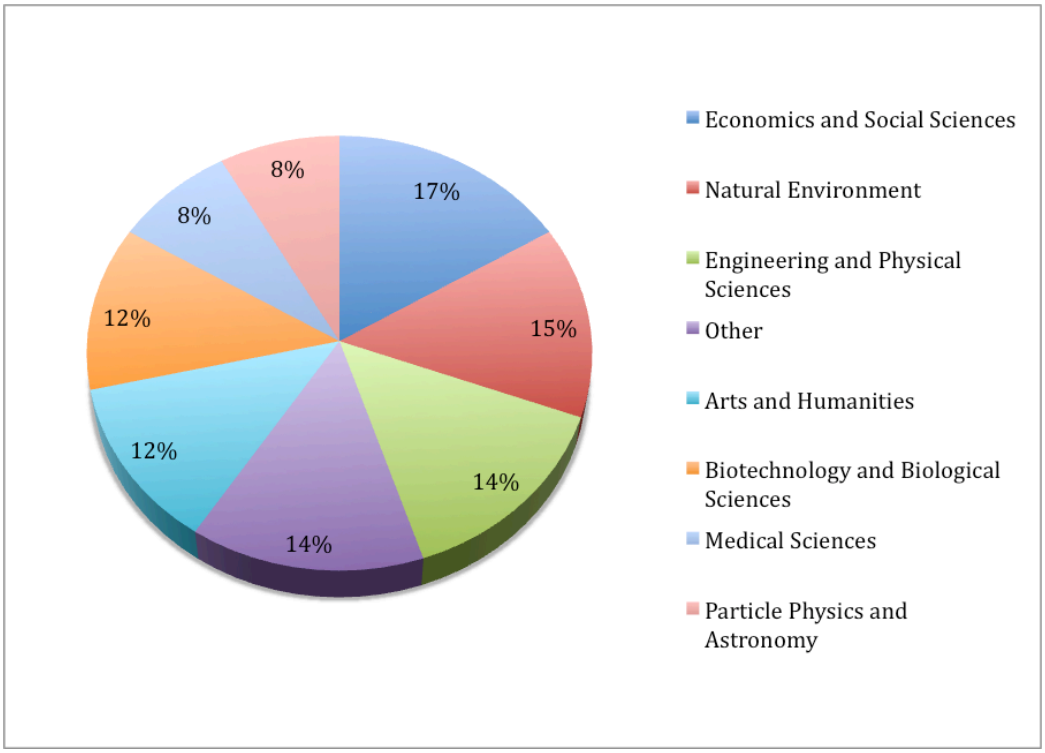

Figure 1: Representation of Research Areas in Initial Sample

Analytical Approach

The transcripts are now being systematically coded up using a coding scheme initially developed on the basis of a literature review conducted at the start of the project. As we process the data, we continue to amend this typology using an iterative, grounded approach, developing a growing understanding of the fieldwork data (cf. Strauss 1987, Charmaz 2008). The coding scheme contains some 166 different codes at the moment, arranged in a hierarchy to make them more manageable (cf. Figure 32) and to allow coding to use different levels of granularity. Details of the coding and the analysis of our interview data can be found in Voss *et al*. (2008).

The data is represented in XML formats using Relax-NG schemas (www.relaxng.org) for validation. The main purpose of using an XML representation of the data is to allow automatic processing and easy transformation into different presentation formats. An online database of findings can thus be built that can allow stakeholders to browse through or search in our corpus of findings and that will serve as a basis for a growing repository of evidence to support future research in this field. As we have documented more than a hundred individual barriers and enablers, compiling these into a long report would be of limited utility. Instead, we aim to allow stakeholders to select information of particular interest to them and to follow connections between items using hyperlinks. An initial version of the database has been built using Apache Cocoon (cocoon.apache.org), which we expect to make available for review in the Autumn of 2008.

Figure 3: Excerpt from the Coding Scheme

The system allows searches using a number of criteria such as: disciplinary affiliation of respondent ('what problems do social scientists face?') or the usage of specific services ('what problems do people face who are using service X?'). Alternatively, the typology can be used to browse the findings. The system uses a diagram representation of the typology as shown in Figure 32. Nodes are clickable and their selection either reveals a list of related barriers or expands the typology to a finer level. For example, in the diagram presented in Figure 32, clicking on 'Collaboration >>' produces the diagram shown on the right hand side.

Findings

In this section, we wish to present some examples of the kinds of findings our fieldwork is generating. As the project is ongoing and the number of interviews carried out relatively small and doesn't yet include the full range of targeted respondents, it is premature to draw summary conclusions from the material we have collected to date. However, we feel that by presenting some examples of barriers and enablers identified and described, we can start to draw attention to the existence of barriers even though we may not be able to say which ones are the most prevalent or which ones have the largest impact. It is also an important step in our plans for developing a dialogue with the research community which will enable us to test and validate our findings. Our examples do highlight opportunities for further development of services or new uses and provide important insights into features of use that may need to be taken into account in service development.

Cost of uptake and investment in new technologies

An important aspect of the take-up of technologies is the process of 'tinkering', of exploring technological options before fully committing and expending significant resources. In times when there is an increased emphasis on sustainability of services and developing cost models, it is important to bear in mind that resources that are free at the point of use play an important role in the exploration of technological options. As one respondent put it:

> "[…] at the moment effectively the cost of facilities are built in but I guess at some point this is going to change, e.g., currently dedicated large computer services are for free but there was a note on the [local compute facility] recently indicating that using them is going to be costed […] That's a chicken and egg thing, until we have explored the technology we don't know what technology can do for us, we are not going to buy technology, we are talking about research process here, before you get the big funding you need to do pilot studies […] if we didn't have a good open door policy [allowing free access] here we wouldn't be where we are now."

We see here a potential barrier emerging if e-Infrastructure service usage cost models are developed that prevent researchers from running pilot studies or developing their applications before submitting a grant application containing requests for significant compute resources. This need for early experimentation was also flagged by another respondent:

> "[...] it would be good to have a, you know, some pilot funding to really spend some time with the Digital Curation Centre to run a pilot experiment to see when there are problems could be addressed by some of the […] tools."

Here, the need for some seed funding is raised that would allow experience to be gained before a larger commitment is made. When we think of 'cost', we should also not limit our understanding of this to purely financial aspects but should also consider the effort involved in making the first steps. No matter how sophisticated an e-Infrastructure and its eventual use, it should be possible for researchers to make some initial steps very quickly to gain some experience and evaluate the options before investing more effort in using a new technology. From these respondents' comments it would seem that, all too often, newcomers face the full complexity at the start of their experience and before significant benefits accrue or can be evaluated (see also below).

Our respondents' comments also draw attention to two potential solutions for this problem: either the complexity of the technology needs to be reduced so that simple things can be done quickly and easily or specific support needs to be provided that enables researchers to enter a cycle of uptake, starting with local outreach events and providing adequate follow-on arrangements. As one respondent suggested:

> "I would suggest if you want to try and increase the user uptake that probably the best thing to do is to actually start going round the institutions and then to think locally through, you know, staff development services or in the case of a National e-Science Centre, centres or something like that […] because it's quite a hard sell, you don't say to somebody we've got this wonderful technology would you like to learn how to use it and then they say, yeah, great, and then nothing really happens after that if you know what I mean because the first port of call is obviously the documentation try and work through things yourself. If it gets complicated people get put off very quickly. So I would suggest a sort of travelling roadshow – give

presentations, go round different universities, you know, show them what's available, show them how it could be useful."

It is important that such outreach activities provide recognisable paths to adoption rather than being singular events with no follow-up. As our respondent suggests, the arrangements for follow-up once interest is awakened needs to be thought through at a local level. What kind of support is available at each step in the process and how easily is this support obtained? Distance plays an important role, so it is crucial to link up local support services and national services such as training provided by e-Science Centres.

Understanding the benefits

Before researchers will even begin to weigh up the costs of uptake, they need to be able to grasp the benefits that e-Infrastructure might provide for their specific research challenges. Comments from several of our respondents suggested that understanding the relevance of e-Infrastructure was not a straightforward matter:

> "The grid stuff took considerable time to be attractive to your man in the street and certainly we're still finding trouble […] It's still seen as big tools to tackle big science problems and I'm knocking on the door as someone who's operating not necessarily such a large operation if you see what I mean. Grid is very much a large scale solution for large scale facilities and problems in many respects, only in science, historically that's the ways it's been but we're slowly breaking that down and therefore, I think there were a number of barriers and that's mainly due to the stigma attached to the word Grid more than anything, as a guy operating my little lab, it was very difficult in the early days for me to see what Grid could do for me kind of thing, and therefore […] in the early days certainly didn't help reduce that barrier or that mindset that I had. It's getting a lot better now, I must say but initially there was, for a very selected community, large scale facilities and problems and it wasn't really for me and it was quite difficult to get over that barrier."

This issue points to the potential importance of supporting a continued and concerted effort via outreach programmes to highlight ways in which e-Infrastructure can be employed to help answer a *wide range* of research questions, at both small and large scales and so counter a perception that it is a tool of use only for tackling the 'grand challenges' of research (Atkinson *et al*. 2008). This could be done by promoting concrete examples, as exemplars, of e-Infrastructure in action that relate not only to 'big science' but also to small-scale questions and facilities. We should also note that e-Infrastructures enable types of research that are not traditionally 'large scale' to benefit from opportunities to scale up in important dimensions.

Professional languages

It is important not only to match examples of the benefits of e-Infrastructure to the research agendas of different audiences but also to talk their language. Some of our respondents reported problems in understanding the 'language of e-Science':

> "I'm not an e-Scientist and it's one of the things that's sort of continually frustrating in the field is the assumed terminology if you know what I mean? That there's a lot of terminology that's come over from computing science which is never designed for the rest of us who actually do the science […]"

These comments suggest that there is a need for outreach materials to be better tailored to their target audiences. Those responsible for the development of outreach programmes must be more aware of the fact that terminology may be unfamiliar and must be willing to deal with the consequences of this. Our fieldwork also suggests that effective outreach requires the commitment of early adopters who, as researchers themselves, are naturally more able to understand how to communicate with their peers. Finally, in pointing to the problem of communication across disciplinary boundaries, it raises the question of to what extent future generations of researchers should be trained to be fluent in a new language of 'digital systems judgement' (Atkinson *et al*. 2008).

Embedding e-Research in education

In order to achieve wider uptake, teaching that equips researchers to effectively use e-Infrastructures must become part of the normal post-graduate and under-graduate courses in Higher Education. Achieving this requires arrangements at a number of different levels. At the institutional level, course content needs to be established and become recognised and integrated in teaching programmes. Lecturers and other teaching staff need to acquire the necessary skills to teach e-Research and training infrastructures must be assembled to allow young researchers to gain hands-on experience.

One barrier that has emerged in the context of our fieldwork relates to the problem that current e-Infrastructures are often not suitable for teaching. For example, licensing arrangements and access control mechanisms may not easily scale to classroom size. As one of our respondents put it:

> "[…] the data is free to download but you have to have signed an agreement or made some sort of declaration about your usage and notify the Data Archive of what you are using the data for. So that is an individual level of arrangement there. If you are talking about groups of people or groups of students accessing data then there maybe requirements for them all individually to make that arrangement. There are intermediate resources, there are teaching data sets that they provide at the Data Archive […] there are conditions in the contract that allow you to release data subsets to people but if you want to do anything particularly complex, the arrangement is between the individual researcher and the archivers."

Here, the problem is that arrangements around the use of datasets are made at an individual level and this places a significant burden on the use of these materials in teaching, as each individual student has to declare their usage of data independently. While the data archives may provide example datasets for teaching that fall under different licenses, these do not necessarily serve the purposes of those who want to teach advanced practices such as those involved in e-Research.

Another barrier relates to the skills required to teach e-Research. Here we can see that our fieldwork uncovers opportunities and enablers as much as barriers. One of our respondents flagged up an opportunity to bring e-Research content into teaching in a systematic way by focusing on courses provided for newly appointed lecturers:

> "[At that point] people are thinking about their teaching and how they can improve [it …] at the universities the people who teach those […] courses I think would be a good target audience because then they will disseminate that to all the

new academic staff who are doing their training so and […] they're responsible for [improving] teaching quality and giving academics new ideas and things."

Problems with institutional support

Issues related to a perceived lack of adequate support for e-Infrastructure services from their host institutions were raised by a number of respondents as a barrier to use:

> "[…] we had hoped when [Access Grid] started that it would develop and it become something that we could just have it on our desktop, in fact we use it much less now, we have switched to WebEx, because it's so simple, and we also use Access Grid only with those that we know there is a very good Access Grid support, so it's wonderful for our collaboration with Southampton and with Edinburgh […] but all of those centres have very good support […]"

> "[…] our barriers are probably because we're not quite set up like a university is so we don't sometimes the things that the universities can access such as access grid we are not able to do so because of our technical infrastructure and the level of support internally […]"

> "I don't see that being fixed unless […] as many people use multicast as used, you know, unicast […], I'm sure that the network administrators would find a way of making it work and fixing it when it breaks but my impression is that there's still [only] a [few] people use multicast, nobody notices when it breaks […]. You are the person who finds out it failed more often because you're one of the few people who use it."

Such comments underline the importance of investigating the barriers to e-Infrastructure adoption as experienced by local service intermediaries. They also suggest a lack of dialogue between researchers and host institution IT services. We will pursue both of these issues in our new phase of fieldwork.

Conclusions

We have presented work conducted to date in the e-Uptake project, which, together with eIUS and ENGAGE, forms the community engagement strand in JISC's e-Infrastructure programme. Our collection of fieldwork data gives us an opportunity to build up a comprehensive and detailed understanding of the issues involved in widening uptake of e-Infrastructure services. While it is too early to draw general conclusions from this work, we can see the value of the material collected emerging.

The examples provided also demonstrate that the issues uncovered can be discipline specific or generic, they can be related to particular contexts of use such as teaching or be of wider relevance. When we consider the typology developed on the basis of our earlier literature review, we find that there is a meaningful relationship between it and the emerging findings but we also find reason to extend the typology. This suggests that we are uncovering evidence that points to issues not discussed in the existing literature. We hope this will come out even more strongly in future publications based on the complete corpus of data at the end of the project as well as in the dynamic database of findings that we are building.

Evidence collected to date is biased towards the views of early adopters and thus represents the issues they have faced that they may or may not have overcome. In our

work over the coming months we will need to make sure we also capture issues faced by those people who have not chosen to invest in these new technologies as well as the views of service providers and intermediaries such as trainers, providers of support and policy makers. That is, we need to broaden the range of views represented in our fieldwork. At the same time, we need to establish the scope and relative importance of issues uncovered and the promise of enablers identified. These requirements present interesting logical and logistical problems for our future work that we will need to seek practical solutions to in the light of the resources and timeframes available to us.

Acknowledgments

The research reported in this paper has been funded by the UK's JISC under the community engagement strand of its e-Infrastructure programme. We would like to thank all those researchers who have given generously of their time and participated in our interviews.

References

Atkinson, M.P, Britton, D., De Roure, DE., Garnett, N., Geddes, N., Gurney, R., Haines, K., Hughes, L., Jeffreys, P., Perrott, R., Procter, R.N. and Trefethen, A.E. (2008). *Century-of-Information Research – a Strategy for Research and Innovation in the Century of Information (CIR3)*, March 2008. http://wikis.nesc.ac.uk/_escienvoy/files/d/d3/CIR3ShortStrategy-v4a-20080311MPA-1.pdf

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

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

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

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