Fostering e-Infrastructures: from user-designer relations to community engagement

Alex Voss, Marzieh Asgari-Targhi, Peter Halfpenny, Rob Procter
National Centre for e-Social Science, University of Manchester
alex.voss@ncess.ac.uk

Sheila Anderson, Stuart Dunn, Elpiniki Fragkouli, Lorna Hughes,
Arts and Humanities e-Science Support Centre, King's College London

Malcolm Atkinson, David Fergusson, Mike Mineter
National e-Science Centre, University of Edinburgh

Tom Rodden
Department of Computer Science, University of Nottingham

Copyright is held by the author/owner(s).

Abstract
In this paper we discuss how e-Science can draw on the findings, approaches and methods developed in other disciplines to foster e-Infrastructures for research. We also discuss the issue of making user involvement in IT development scale across an open community of researchers and from single systems to distributed e-Infrastructures supporting collaborative research.

Keywords
User-designer relations, community engagement, requirements

Introduction
Existing investments in e-Science and Grid computing have helped to develop the technical capacity to build e-Infrastructures (aka. cyberinfrastructures in the US) for research: distributed, networked, interoperable computing and data resources that are available to underpin a wide range of research activities in all research disciplines. As the basic technological components and architectural styles are maturing, issues of uptake and of embedding of e-Infrastructures in day-to-day working practices come to the fore. Indeed, one may argue that if these issues are not

addressed, the e-Science community will not succeed to realise its full potential and to achieve sustainability.

Consequently, investments in the development of technologies and applications are now being complemented by active programmes of community engagement [11]. The aim is to study uptake of e-Infrastructures, to devise plans for widening and deepening adoption through targeted interventions (such as training, education and outreach programmes, influencing technology development and policy making).

Fostering e-Infrastructures

As e-Infrastructures are complex ensembles of technical components and social arrangements, the question arises whether they can be 'built' or whether we should rather speak of 'fostering' [2,14]. Interventions are required both in the technical realm where the traditional language of 'building' and 'design' can be used and in the space of social arrangements where such terms are less appropriate.

This is more than a question of semantics: changing social arrangements, individual and collective human behaviour requires different kinds of interventions that account for the fact that people have agency and are not 'tame' or 'docile' like material artefacts. They also need to cope with a constantly changing, distributed environment, changing stakeholders with only partially overlapping agendas and established traditions that resist change. While some of these challenges are not unique to the area of e-Infrastructures for research, the particular combination of factors present here challenges traditional development approaches [15].

Speaking of fostering e-Infrastructures indicates that the social arrangements sustaining them can only come about through a sustained, open-ended process of alignment of interests, negotiation, learning, alliance building, discussion and convincing, etc. The scope of the interventions is likewise not a local one but requires effort at the level of the individual, the local and organisational level as well as the wider inter-organisational, national and international level. The knowledge and skills that need to be brought to bear in this area is not readily available as it cuts across the concerns of different fields of study.

Consequently, efforts to foster e-Infrastructures need to be interdisciplinary endeavours drawing on fields as diverse as software engineering, management studies, social sciences (e.g., sociology, social anthropology, economics), workplace studies (as conducted within the computer supported cooperative work and participatory design communities), social psychology, science and technology studies and philosophy of science.

Over recent years, social scientists and researchers from the other fields mentioned above have started to engage with e-Science projects and e-Infrastructure development but the effectiveness of this involvement has been varied and has depended on a number of factors such as the point in time when the involvement started, the level of development of the infrastructure at that point, the participatory model of social science in the project and the social scientists' relations to other project stakeholders [7]. As is the case in many IT development efforts, issues such as participation in design, usability and socio-technical alignment [5] are often addressed at a late stage, experts are brought in only once problems become apparent and,

consequently, interventions are less effective because the *status quo* has become solidified.

Operationalising Lessons Learnt

The crucial challenge, then, is to bring the different concerns, analytical approaches, research methods, theories and interventions that various relevant fields provide together in a form that can be effectively *operationalised*. In particular, we need to pay attention to the need to make our approaches scale to match the scale and complexity of e-Infrastructure developments [14].

In the following sections, we will discuss some ways in which this can be achieved, either through incorporation of new elements in the work of those working on e-Infrastructures or through the development of models of involvement of researchers from various disciplines. Both elements are necessary because, on the one hand, we cannot afford the luxury to have experts on every project and on the other, we cannot assume that all relevant knowledge and skills can be learned by the relatively small group of people involved in any endeavour. What we are aiming for is a rich and consistent set of measures that can demonstrably benefit efforts to foster the development of e-Infrastructures for research. Because of space limitations, we will mention only a few of these, a longer version of this paper will go into more depth and cover more aspects.

Capturing Knowledge and Sensitising to Issues

Fostering e-Infrastructures for research is not a new activity but an endeavour many people have been involved in over the past decade (e.g., [6]). However, the scale, complexity, ambition and dynamism of these efforts has increased over recent years. Consequently, there are people in the community who have experience with the issues involved in fostering e-Infrastructures for research. However, this knowledge is not routinely shared and spread throughout the wider community [8], leading to poor utilisation of existing knowledge and the danger that costly mistakes are repeated and opportunities missed.

Two steps need to be taken to remedy this situation: existing knowledge needs to be captured, documented and made sharable in a form that is readily accessible for practitioners. The eIUS project [4] is currently producing use cases produced from and linked to a series of experience reports that seek to capture successful and inspiring uses of JISC-funded e-Infrastructure services. Secondly, practitioners need to be sensitised to the issues involved in fostering e-Infrastructures. A number of studies exist that describe particular cases (e.g., [3]) or focus on particular aspects such as privacy and confidentiality (e.g., [1]). However, relatively few contributions have appeared in the mainstream e-Science literature and there are as yet no collections and distillations of these studies that would make them easily accessible to practitioners. The lack of agreed curricula in e-Science has been noted and efforts are currently underway to establish agreement on their content (e.g., in the OGF Education and Training Community Group). We would argue that an e-Science curriculum should not be restricted to technical topics but should include elements that address the wider issues of fostering e-Infrastructures.

Developing a curriculum necessarily requires a certain degree of closure and agreement on accepted terms.

Unfortunately, the dynamism of the field and an unfortunate tendency to re-invent terms and concepts has led to a good deal of confusion, especially around terms such as grid computing, cloud computing and web 2.0. If there is a lack of convergence about the key underlying technologies, there is certainly a lack of conceptual vocabulary to describe e-Infrastructures as we a larger whole [14]. Our view is that filling this gap should be a key short-term priority for researchers and funders.

Fostering Communities

An important aspect of fostering communities is to ensure that relevant baseline information is available about the size and structure of the community, its main features (such as funding arrangements, career structures or research culture) and the level of uptake of e-Infrastructures. e-Science tools can play an important role here. For example, we are currently exploring the use of text mining technologies (such as key term extraction) and social network analysis tools to study uptake of e-Infrastructure services within the e-Uptake project [10]. Repeated runs of the same process using data gathered so far for seeding will allow us to track developments in the communities and in terms of uptake of services, providing us with indicators of the success of our interventions in the longer term.

Of course, this passive collection of evidence is the first step involved and a means to monitor progress. In addition to this, projects have a need to actively involve communities in a number of activities such as formulating requirements, organising funding, influencing decision- and policy makers and sharing their experiences with peers. The last point is of particular importance to widen the uptake from the early adopters to the interested, to get the disengaged interested and to convince the sceptical.

Coordination

Initiatives to foster community engagement are frequent – most service providers have developed ways to communicate with and learn from their user communities. Surveys are frequently used, either online using mailshots to the whole community or at events in the form of exit-questionaires. There is a real danger that the frequency with which researchers are approached and the lack of clear traceability to changes in service provision will lead to *respondent fatigue* and consequently to poor return rates.

Within the JISC Community Engagement Strand – consisting of the e-Uptake and eIUS projects and the ENGAGE initiative – we have worked up a data sharing and coordination agreement that seeks to minimise the number of times researchers are approached while maximising the coverage achieved across the three projects. Prior to the approach, we are sharing lists of candidate respondents[1] and, on approach, we use a common consenting process that allows interviewees to opt in to the data sharing between the projects or to agree to follow-up interviews.

Social Research Approaches and Methods

Many IT development projects make use of methods developed in the social sciences to elicit information about (potential) users and uses of e-Infrastructures. However, there is a lack of understanding of the principles of social research and the difference, for

[1] These lists contain only information that is in the public domain.

example, between research design and research method. Research design deals with the *logical* problem of ensuring "that the evidence obtained enables [the researcher] to answer the initial question as unambiguously as possible" [9], p. 9), whereas methods deal with the *logistical* problem of obtaining and processing data ([13], cited in [9], p.9). Practitioners all too often simply pick a convenient method (such as using a questionnaire to run a survey) without going through the phase of considering the fundamental question of the research design.

We do not mean to suggest that practitioners should delve into the depths of social science analytical and methodological debates but what is needed is a basic orientation to the basic principles of social scientific research that can be acquired in a reasonable amount of time. At the same time, researchers need to learn to accept and embrace the fact that social science research does not always deal with hard facts and conclusions derived in a straightforward way from statistics. Rather, there is often an element of interpretation involved – good social science research makes this transparent rather than trying to avoid it. It is perhaps here that social scientists (or researchers from other disciplines) may contribute most usefully.

Conclusions

We have outlined why we think that those interested in fostering e-Infrastructures for research face challenges far beyond the scope of any single discipline and therefore necessarily need to develop interdisciplinary approaches. Findings, analytical approaches, methods and interventions that can contribute to solving these problems are in principle available as a range of disciplines exist that have studied particular relevant aspects. However, these elements need to be made available to e-Science practitioners in a readily accessible form so they can be effectively operationalised. We have discussed four areas in some more depth and have made suggestions for concrete approaches and actions to be taken.

Acknowledgements

The work presented in this paper has been funded by the EPSRC through the e-Science Institute Theme on Adoption and Sustainability of e-Research Technologies and by JISC through the Widening Uptake of e-Infrastructure Services project.

References

[1]Carusi, A. and Jirotka, M. From data archive to ethical labyrinth. *Proceedings of e-Social Science '07*. Ann Arbor, October 2007.

[2]Edwards, P.N., Jackson, S.J., Bowker, G.C. and Knobel, C.P. *Understanding Infrastructures: Dynamics, Tensions and Design*. Report of a Workshop on "History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures. National Science Foundation.

[3]Lee, C.P., Dourish, P. and Mark, G. The Human Infrastructure of Cyberinfrastructure. *Proceedings of the Conference on Computer Supported Cooperative Work*. Banff, Alberta, Canada, November 2006. pp. 483-492.

[4]Mascord, M., Argüello Casteleiro, M., Fraser, M., Voss, A., Procter, R., Halfpenny, P. and Jirotka, M. *Scoping e-Infrastructure Usage: Interim Report*. November 2007. Available at: http://www.eius.ac.uk (accessed 21.02.2008)

[5]Molina, A. Insights into the nature of technology diffusion and implementation: the perspective of sociotechnical alignment. *Technovation* 17(11/12), 2006. pp. 601-626

[6]Olson, G.M., Atkins, D.E., Clauer, R, Finholt, T.A., Jahanian, F., Killeen, T.L., Prakash, A., & Weymouth, T. The Upper Atmospheric Research Collaboratory. *Interactions*, 5(3), 1998. pp. 48-55.

[7]Ribes, D. and Baker, K. S. Elements of social science engagement in information infrastructure design. *Proceedings of the International Conference on Digital Government Research*. San Diego, California, May, 2006. pp. 450-451.

[8]Schopf, J.M., Coleman, I., Procter, R. and Voss, A. *Report of the User Requirements and Web Based Access for eResearch Workshop*. UK e-Science Technical Report UKeS 2006-07, November 2006.

[9]De Vaus, D. *Research Design in Social Research*. Sage Publications Ltd., 2001.

[10]Voss, A., Asgari-Targhi, M., Procter, R., Halfpenny, P., Dunn, S., Fragkouli, E., Anderson, S., Hughes, L., Mineter, M., Fergusson, D. and Atkinson, M. Widening Uptake of e-Infrastructure Services. *Draft paper submitted to the 4th International Conference on e-Social Science*. Manchester, June 2008.

[11]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. e-Infrastructure Development and Community Engagement. *Proceedings of e-Social Science '07*. Ann Arbor, October 2007.

[12]Voss, A., Procter, R., Jirotka, M. and Rodden, T. *Managing User-Designer Relations in e-Research Projects*. Position statement for the Workshop on Realising e-Research Endeavours, e-Science Institute, Edinburgh, March 2007.

[13]Yin, R.K. *Case Study Research: Design and Methods*. Sage Publications, 1989.

[14]Zimmerman, A. A socio-technical framework for cyberinfrastructure design. Proceedings of e-Social Science '07. Ann Arbor, Michigan, US, October 2007

[15]Zimmerman, A. and Nardi, B. Whither or Whether HCI: Requirements Analysis for Multi-Sited, Multi-User Cyberinfrastructures. Workshop on Usability Research Challenges for Cyberinfrastructure and Tools, *ACM Conference on Computer Human Interaction*, Montréal, Québec, Canada, April 2006. pp. 1601-1606.