
The e-Social Science research agenda

BY PETER HALFPENNY* AND ROB PROCTER

*Manchester eResearch Centre, School of Social Sciences,
Arthur Lewis Building, University of Manchester,
Manchester M13 9PL, UK*

In this paper, we use the experience of the first 5 years of the UK Economic and Social Research Council's National Centre for e-Social Science as a basis for reflecting upon the future development of the e-Social Science research agenda.

Keywords: e-Social Science; e-infrastructure; social research methods;
social shaping; sustainability

1. Introduction

The aim of this paper is to review the lessons learned from the first 5 years (2004–2009) of the UK National Centre for e-Social Science (NCeSS)—funded by the Economic and Social Research Council (ESRC)—and outline their implications for the future development of the e-Social Science research agenda. Section 2 describes the structure of NCeSS, and §3 describes its research programme. Section 4 notes the growing diversification of e-Social Science and, briefly, how this relates to e-Science and e-Humanities. Section 5 comments on the changing technical landscape over the past 5 years. Section 6 considers the social science research community's uneven responses to the e-Social Science programme. Section 7 identifies issues that have arisen around funding, especially in relation to sustainability. Section 8 concludes with some pointers towards a more grounded programme unfolding over the next few years, while noting the danger that e-Social Science's transformational opportunities could yet be lost because of barriers emanating from the wider context within which the initiative operates.

2. The National Centre for e-Social Science

Over the period 2001–2006, the UK Government invested £213 million in an e-Science programme (Hey & Trefethen 2004). The overall aim of the programme was to invent and apply computer-enabled methods to 'facilitate distributed global collaborations over the internet, and the sharing of very large data collections, terascale computing resources and high performance visualizations' (EPSRC 2009). The funding was divided between a 'core programme',

*Author for correspondence (p.halfpenny@manchester.ac.uk).

One contribution of 16 to a Theme Issue 'e-Science: past, present and future I'.

focused on developing the generic technologies needed to integrate different resources seamlessly across computer networks, and individual Research Council programmes specific to the disciplines they support. The ESRC's allocation was £13.6 million over the 5 years.

The ESRC's programme began cautiously with the commissioning of four scoping studies, one each on quantitative and qualitative research, one on Grid technologies and one on social science studies of technological innovations (Anderson 2003; Cole *et al.* 2003; Fielding 2003; Woolgar 2003). This was followed by a £500K call to fund a set of 11 pilot demonstrator projects. These were intended to test the social science research community's interest in e-Social Science by enabling 'early adopters' to generate exemplars. The purpose of these exemplars was to demonstrate the potential of Grid technologies to advance social science and encourage other researchers to adopt the emerging technologies. The ESRC's somewhat tentative start to its programme acknowledged that initial awareness among social science researchers of the potential of e-Social Science was likely to be limited.

The next step in the ESRC's e-Social Science strategy was to scale up the research programme by commissioning NCESS, with £6 million allocated for the first phase from 2004 to 2007. In the light of positive peer reviews, this was followed by £8 million for a second phase from 2008 to 2012. The Centre was designed to have a distributed structure. Research devoted to developing innovative tools, techniques and services was primarily undertaken in 12 large 3-year projects or 'nodes', located in different universities across the UK. There were seven nodes in the first phase and there are eight in the second, of which three are extensions of first-phase nodes, one is a combination of two first-phase nodes and four are new, brought on stream to fill gaps in the first phase. Until October 2009, the nodes were coordinated by a 'hub' team, based at the University of Manchester, responsible for designing and managing the research programme and a parallel dissemination programme, creating and exploiting synergies across the components of the programmes, and strategically planning future developments. The Centre also incorporated a series of 12 'small grant' (£50K) projects, each of around a year's duration, intended to explore more speculative ideas. These too were distributed across the UK. (More details about the ESRC's e-Science strategy and NCESS's part in delivering it can be found in Halfpenny & Procter (2008) and Halfpenny *et al.* (2009).)

The NCESS research programme has two strands. The *applications strand* draws on unfolding developments in technologies, tools and services from the e-Science programme and applies them to the particular needs of the social science research community. This strand is pursued by the nodes through 'driver projects' that address substantive research questions within particular social science fields, the objective being to improve existing methods or develop new ones that enable advances in the fields that would not otherwise be possible. The *social shaping strand* falls within the social studies of science and technology tradition and aims to understand the social, economic and other influences on how e-Science technologies are being developed and used, and the technologies' implications for and impacts on scientific practice and research outcomes. The objective is to understand the genesis of technological innovations and the barriers and facilitators to their uptake within scientific research communities, and then use this knowledge in the design and delivery of awareness-raising and

capacity-building activities that extend the reach of e-Science. Note that the object of inquiry of the social shaping strand is the whole of e-Science, not just e-Social Science.

3. NCeSS research programme 2004–2012

A total of 11 nodes have been funded in the applications strand and one in the social shaping strand over the two phases of the NCeSS research programme. Nodes in the applications strand can be further subdivided into six focusing on developing tools in support of quantitative social research, one on qualitative research, two on ‘mixed methods’ at the intersection of quantitative and qualitative research, and two focusing on inter disciplinary research.

The CQeSS first-phase node (Lancaster University and Daresbury Laboratory, 2004–2007) developed middleware that would allow users to exploit distributed datasets and computational facilities from within their familiar desktop analysis tools (Kewley *et al.* 2007). Notable among the tools CQeSS developed is multiR, a parallel extension to the widely used open source statistical package R (Grose *in press*). This is particularly valuable in computationally demanding applications such as geographically weighted regression—a means of modelling relationships between variables using standard regression but which allows regression coefficients to vary spatially—where multiR has been demonstrated to reduce the computation time from tens of days to less than 10 hours (Grose *et al.* 2007; Harris *et al.* 2009).

The e-Stat second-phase node (Bristol and Southampton Universities, 2009–2012) continues the development of sophisticated quantitative analysis tools, and combines this with developing interoperability between software packages. This latter is intended to encourage the sharing and reuse of cutting edge methodological developments; if complex statistical models are accessible from within researchers’ favourite statistical package, they can be adopted without having to grapple with the unfamiliar syntax of a different package.

The MoSeS first-phase node (Leeds University, 2004–2007) developed a demographic simulation and forecasting tool for urban and regional planning based on synthetic representations of the entire UK population (Birkin *et al.* 2007, 2010; Townend *et al.* 2007). The chosen policy application drivers were healthcare, transport planning and public finance. Social science problems of this type are characterized by a requirement for extensive data integration and multiple iterations of computationally intensive scenarios.

The GeoVUE first-phase node (University College London, 2004–2007) developed geographical information systems (GIS) tools to enable users to easily map and visually explore spatially coded socio-economic data (Milton & Steed 2007; Hudson-Smith *et al.* 2009). Driver applications included urban planning and design.

The evident complementarity between MoSeS and GeoVue is now being exploited in the second-phase node GENeSIS (University College London and Leeds University, 2007–2010), which combines the work of the two earlier nodes in the development of tools for agent-based modelling (Birkin *et al.* 2009).

The DAMES second-phase node (Stirling University and National e-Science Centre, Glasgow, 2008–2011) is developing a suite of tools to support social scientists in data management tasks such as those involved in manipulating variables and linking of different datasets. The sheer heterogeneity of social data—not only of formats but, more importantly, of coding schemes for the same variable—makes data management a difficult and time-consuming activity, presenting an increasingly important barrier as new research challenges demand the use of more and different sources of data (Lambert *et al.* 2008; Sinnott *et al.* 2008; Tan *et al.* 2009; Warner *et al.* 2009).

The MiMeG first-phase node (Bristol University and King's College, London, 2004–2007) developed tools to support distributed, synchronous collaborative video analysis (Fraser *et al.* 2006, 2007; Tutt *et al.* 2007). Digital video has become an invaluable tool for social scientists to capture and analyse a wide range of social action and interactions. Video-based research is increasingly undertaken by research teams distributed across institutions in the UK, Europe and worldwide, but there was little existing technology to support collaborative analysis.

The DReSS first- and second-phase node (Nottingham University, 2004–2011) is developing new tools for capturing, replaying, and analysing multi-modal digital records of people's activities. The digital replay system (DRS; Brundell *et al.* 2008) enables researchers to synchronize and replay heterogeneous digital records side-by-side on their computer screens, each in its own window. This facilitates the analysis of multiple data sources simultaneously. An example is provided by multi-modal corpus linguistics, where audio, video and transcript records are brought together to more fully understand social interaction (Crabtree *et al.* 2007; Carter & Adolphs 2008).

The PolicyGrid first- and second-phase node (Aberdeen University, 2004–2011) is developing a set of semantic Grid tools (Hielkema *et al.* 2007; Pignotti *et al.* 2007) to support social scientists and policy-makers using mixed-methods research techniques (e.g. surveys and interviews, ethnographies, case studies, simulations; Edwards *et al.* 2009). The tools provide an audit trail, from data, through analysis to policy recommendations, enabling the grounds of alternative recommendations to be interrogated. The substantive research focus is rural policy development and appraisal.

The LifeGuide second-phase node (Southampton University, 2008–2011) is developing tools that will enable researchers to quickly and easily construct Web-based behavioural interventions (WBBIs) and subsequently monitor the effectiveness of their implementation. The application area is lifestyle choices that affect health, since WBBIs potentially offer a cost-effective way of changing the health-related behaviours of large numbers of people, while at the same time generating high-quality data about their activities (Yang *et al.* 2009; Webb *et al.* 2010).

The Obesity e-Laboratory second-phase node (Manchester University, 2008–2011) is developing a secure virtual environment in which epidemiologists, public health researchers and social scientists can share all the 'research objects' that they use in their obesity research. A research object might, for example, contain everything necessary to reproduce a scientific result: the data; the database query used to discover and extract them; statistical scripts for data cleaning, deriving variables and modelling; reports; presentation slides (Ainsworth & Buchan 2009; Buchan *et al.* 2009; Thew *et al.* 2010).

The Oxford e-Social Science first- and second-phase node (Oxford University, 2004–2011) is addressing the inter-related social, institutional, ethical, legal and other issues that influence, both positively and negatively, the take-up of e-Science and its routinization within research practices. For example, the node has examined confidentiality, privacy, data protection, intellectual property rights, accountability, trust, and risk in distributed collaborations (Carusi & Jirotko 2009; Dutton & Meyer 2009; Meyer & Schroeder 2009).

Individually, each node is undertaking research of strategic importance for the ESRC's priority research areas, as set out in its 2009–2014 Strategic Plan. Collectively, the nodes address all stages of the social sciences research lifecycle, from resource discovery through data analysis to the publication of results. In its coordinating role, the NCeSS hub sought to ensure the programme's synergistic potential was fully realized—both internally and externally with related ESRC investments (for example, by joint activities with the nodes of the National Centre for Research Methods) and with the wider e-Science programme.

A key part of developing intra-programme synergies was the 'e-Infrastructure for the Social Sciences' project led by the hub (Daw *et al.* 2007). In particular, this has sought to encourage the development by programme partners of common technical solutions, such as metadata and service registries, tools for user authorization and authentication, user portals and collaborative virtual research environments. The project also promoted a shared understanding of the challenges faced in resource discovery, data access, security and the usability of new tools. Its output is a set of demonstrators who illustrate the capacities of new research tools and the potential of e-Social Science.

4. The scope of e-Social Science

The ambition of the UK e-Science programme was to facilitate bigger, faster and more collaborative science, driven by a vision of researchers worldwide addressing key challenges in new ways. Technically, the initial emphasis of the programme was focused on exploiting the Grid, that is, the hardware, software ('middleware') and standards necessary to co-ordinate geographically distributed—and possibly heterogeneous—computing and data resources and deliver them to researchers over the Internet, regardless of location. This model was particularly appropriate to particle physics where, for example, it is seen as essential if the best use is to be made of the vast amounts of data generated by the Large Hadron Collider (<http://www.lhc.ac.uk/>). It was less matched to those disciplines subsequently encouraged to join the e-Science bandwagon, including the social sciences, in which a mixture of numerous quantitative and qualitative methods is used to pursue relatively small-scale research, with few generic problems requiring complicated middleware to coordinate large-scale, distributed computing and data resources.

The Grid was also appropriate to complex problems such as weather predictions and earthquake modelling. The potential of the Grid to tackle complexity, rather than volumes of data, went largely unrecognized in the e-Social Science programme, and interdisciplinary applications of complexity science have been funded by a separate Engineering and Physical Sciences Research Council (EPSRC) programme outside the investment in e-Science.

Another reason why enthusiasm for the Grid in the social sciences was limited is that most areas—especially quantitative social science—that have generic requirements are already well served by mature products. One example is survey research, where the last three or four decades have seen increasing computer assistance transform the field. This started with computer-assisted telephone interviewing (CATI), relying on mainframe computers for random-digit dialling and the presentation of the survey questions on-screen to the telephone interviewer. It progressed to computer-assisted personal interviewing (CAPI) using laptops to replace clipboards and paper questionnaires, modems to download the questionnaires and upload the completed interviews, and survey management software to issue samples and audit progress. This is now supplemented by computer-assisted Web interviewing (CAWI), where the survey instrument is delivered directly to the respondent via a Web browser, with the potential to include multi-media. In all cases, the computerized environment allows great intricacy by incorporating automatic filtering of questions and routing of the interview according to answers, undertaken in real time but hidden from the users.

A second example is statistical analysis, where there are numerous commercial general purpose and niche software packages available, alongside well-supported open source alternatives and Web-based services. In some cases, the analysis routines are integrated with data collection software, providing joined up computer support for the whole survey research life cycle.

Some areas of qualitative social research are also well served, with established products readily available. One such area is computer-assisted qualitative data analysis software (CAQDAS), though this title is misleading because the packages aid data management, leaving the analyst to interpret and categorize the data (although there have been some recent attempts to include text-mining techniques to automate the categorization of data). In this respect, the impact of computational tools on qualitative social research has been to *support* analysts, whereas it is deployed to *do* quantitative social science, for example by running analysis routines.

There are some qualitative social researchers—those subscribing to the hermeneutic tradition—who maintain that the analysis of human action and social interaction is beyond the reach of computerization because it is reliant on interpretive understanding, an essentially human capability; they are ‘epistemological sceptics’. This hermeneutic tradition’s central tenet is the rejection of the unity of the sciences (which lies at the heart of the rival positivist tradition), instead maintaining that there is a fundamental and unbridgeable division between the human and the physical sciences.

These features of social sciences—mature analysis tools available for quantitative researchers, mature support tools available for some qualitative researchers and antipathy towards computerization by other qualitative researchers—meant that there was relatively little demand in the social sciences for the Grid model when the e-Science initiative began. However, as the NCeSS research programme unfolded and gave impetus to the application of a wide range of advanced computing technologies to social science research, what was considered to fall within the e-Social Science ambit broadened out considerably to include numerous research practices beyond Grid computing (Halfpenny & Procter 2009). These included the use of digital data harvested from the

Web to capture people's views and map their individual behaviours and their social networks; the exploration of new forms of digital data such as mobile phone logs and GPS (Crabtree *et al.* 2007); the creation and exploitation of metadata to facilitate the sharing and reuse of data (Hielkema *et al.* 2007); linking different sources of data about individuals (Warner *et al.* 2009) and the confidentiality and ethical issues that this raises (Elliot *et al.* 2005; Thomas & Walport 2008); webometrics or cybermetrics concerned with measuring the Web and uncovering its structure by tracing hyperlinks (Thelwall 2009); mapping geo-referenced data (Hudson-Smith *et al.* 2009); large-scale social simulations of various sorts (Townend *et al.* 2007); distributed, collaborative markup and analysis of video data about social interaction (Fraser *et al.* 2006); text and data mining (Ananiadou *et al.* 2007; Gibson *et al.* 2007); and tools for delivering behavioural interventions over the Internet (Yang *et al.* 2009).

Outside a few core areas, such as particle physics and biosciences, the situation is similar in the physical sciences, with a plethora of relatively small-scale applications of advanced information and communication technologies (ICTs) being pursued; e-Science is no longer restricted to the Grid model. The programmes of the recent UK e-Science All Hands Meetings demonstrate this (see <http://www.allhands.org.uk/previous-ahms>). In the humanities, scholars have been particularly enthusiastic about the digitization of their data resources and the provision of advanced search facilities because in combination these dramatically increase the efficiency with which original sources can be discovered, interrogated and combined (Goldberg 2009).

Overall, the initial characterization of e-Science as Grid-enabled, bigger, faster, more collaborative science has given way to a much more blurred picture of exploiting a wide variety of ICTs wherever they assist in the collection, analysis and presentation of digital data within existing research fields and traditions; more piecemeal evolution than radical revolution.

5. The changing technical landscape

As the notion of the Grid being at the core of e-Science has gradually given way to a more diverse view, the emphasis has switched to e-infrastructure (cyberinfrastructure in the USA). This is not simply a synonym for the Grid but marks a shift of attention to a broad range of computing tools, techniques and services that support the everyday work of research scientists, the 'digital equipment, software, tools, portals, deployments, operational teams, support services and training that provide computational services to researchers' (Research Councils UK 2009). This more inclusive view is promoted by the social shaping research strand of the e-Social Science research programme, which emphasizes that non-technical factors are critical to the successful incorporation of technical innovations into established practices. The adoption of the Grid—or any ICT—as a technical enabler cannot be considered in isolation from the institutional, social and cultural environment within which research is conducted.

One aspect of this wider environment is the take-off in popularity of the rapidly increasing number of tools that can be loosely collected together under the title of Web 2.0, whose emergence serves as a timely reminder of the

uncertain pathways of technical innovation. (Here, and in what follows, ‘tool’ is used in a broad sense to include any technology or service deployed to support research, rather than something designed and constructed for a fixed purpose.) When the UK e-Science Programme began in 2004, Grid ‘middleware’—sophisticated software that would permit large numbers of distributed and heterogeneous computer systems to inter-operate, thereby providing researchers with seamless, on-demand access to scalable processing power to handle large-scale datasets, regardless of the location of the researchers, the processing power or the data—was seen as the key to the realization of the vision. However, the Grid blueprint has subsequently proved slow to mature, sometimes difficult to deploy in practice and does not always offer the most appropriate solutions to researchers’ requirements. Meanwhile, more lightweight—though technically less powerful—technologies such as Web 2.0, designed to deliver easy-to-use services via simple protocols and familiar Web-based user interfaces have emerged to fill the void (Fox *et al.* 2007). These features considerably reduce the barriers to uptake and Web 2.0 is providing flexible solutions to at least some researchers’ needs for advanced computing tools and services. Accordingly, a benefit of Web 2.0 is that it has increased the scope for community participation in the development of research resources and facilitated a more ‘agile’ development strategy that is capable of being more responsive to changing requirements (Procter *et al.* in press). In sum, Web 2.0 is seen by many as resonating with the e-Science vision of sharing and collaboration in ways that Grid middleware—seen in contrast as being the product of a more top-down development strategy—does not.

The appearance of a plethora of Web 2.0-based tools over the past 2 or 3 years has drawn in many social science researchers, who have discovered immediate benefits from using, for example, Connotea, Slideshare, SurveyMonkey, Flickr, Vimeo, YouTube, Skype, etc., all of which have shallow learning curves. In other words, Web 2.0 has generated a lively interest in deploying advanced ICTs to support research across a much wider constituency than did the promises of the Grid. Web 2.0 has also raised expectations about what might be possible in what has been dubbed ‘Science 2.0’ and ‘Research 2.0’ (Waldrop 2008). The publishing of their service application programming interfaces (APIs) by major Internet companies such as Google and Yahoo, and the consequent flood of ‘data mashup’ applications, as exemplified by MapTube (Hudson-Smith *et al.* 2009), has reinforced this trend.

The changing technical environment also includes the recent promotion of cloud computing (Buyya *et al.* 2009), the outsourcing of computing resources—hardware, software, data storage—to commercial enterprises that take care of all the technology issues and deliver services over the Internet on demand. This is already a feature of Web 2.0 applications through which people upload their personal data, email messages, photographs, videos and files onto servers owned by large companies such as Google. The attraction to both individual and institutional users is obvious: no procurement, set-up and maintenance, and pay as you go services. There are major concerns, however, about such issues as continuity of supply, privacy and security, centralized control, and lack of customization and flexibility. Whether outsourcing to the cloud is a feasible route for scientific researchers remains to be seen; its advantages are likely to be greater than the disadvantages in some fields but not in others.

6. Understanding the community

Segments of the social science research community can be differentiated in terms of their openness to technological innovations. Based on the NCeSS hub's profiling of the social science research community, we devised a threefold division to help understand how the opportunities offered by e-Social Science might unfold. We use this in preference to that of Rogers (1995) as the latter charts a unitary trajectory towards adoption (Sørensen 1996), overlooking the variety of routes taken by different social science research groups engaging to different degrees with different tools for different purposes.

'Early adopters' are, by definition, keen to experiment with innovations and to take risks. Many are inside the 'NCeSS tent', having been attracted by the calls for pilot demonstrator projects, nodes and small grant projects. It is they who are creating e-Social Science tools and demonstrating their value through driver projects. The hub's role is to encourage the adoption of shared standards and exchange of technical expertise.

The 'interested' form the test-bed for e-Social Science applications; they are drawn to demonstrators that expose them to potential benefits for their own research. Monitoring their experiences in implementing the demonstrator technology in their own fields provides valuable feedback to guide further development of the tools.

The 'unengaged', still by far the majority, include those who have had no exposure to e-Science and lack awareness of its potential, through those who are aware but remain unwilling to invest in new ways of working, to technophobes who shy away from all computer assistance. They will also include those wedded to the hermeneutic tradition, introduced above, who remain resolutely opposed.

Extending adoption of e-infrastructure to the unengaged is handicapped (as is e-research as a whole; see Voss *et al.* (2008, 2010)) by a complex of factors whose importance varies with the bases for not engaging: lack of awareness; lack of current availability or uncertainty about continued future availability; unreliability; problems translating innovation in one field to another; risk aversion; weaknesses in IT support; steep learning curves; and the lack of incentive because career goals—publications and promotions—can be achieved using familiar tools.

Even the more tractable of the unengaged will adopt new tools and services only when these are 'hardened' to production level, that is, become easy to use, stable, documented and supported; when they offer immediate benefits; and when they complement existing research practices (Hardisty *et al.* 2008). It is the ease-of-use and utility of e-infrastructure, and the predictability that it will advance their substantive social research—and thereby their careers—that will persuade them to adopt new ways of working. It is this that lies behind the enthusiasm among many social scientists for Research 2.0 despite their indifference to or scepticism about Grid enablement.

7. Funding and sustainability

Whether narrowly focused on the Grid or broadened out to Research 2.0, e-Science has innovation at its heart, the invention and application of new tools, techniques and services that enable research to advance in ways that

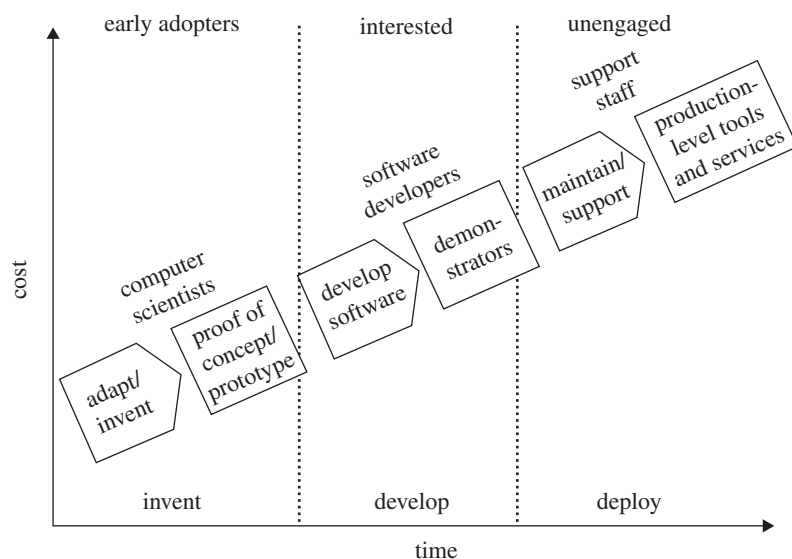


Figure 1. Innovation pipeline: time and cost.

otherwise would not have been possible. Innovation, by its very nature, is unpredictable and therefore planning its support is immensely difficult: how does one estimate when innovations will occur, whether they will be adopted by the research community, how they will be adapted as they are absorbed into everyday research practice, what support users will need and, especially, what budget is required? These questions all relate to the ‘innovation pipeline’, the route taken by innovations from their initial emergence to their mainstreaming, that is, to their widespread incorporation into the normal practice of researchers. Various different stages along the pipeline can be identified and plotted on a cost–time graph, as in figure 1. The innovation pipeline captures what is sometimes referred to the ‘hardening’ or ‘productionizing’ of software. It is the software development necessary to turn a researcher’s ‘proof of concept’ into a working demonstrator and then into a stable product that is supported and maintained. Of course, the angle of the pipeline—the rate at which innovations proceed along it and the cost of mainstreaming—and whether the pipeline is straight, curved or wavy are unknown in advance and are likely to vary from application to application. All that can be predicted is that, although there might be ups and downs along the way, the pipeline will rise inexorably towards the top right-hand corner: moving along it takes time and costs money. And yet moving innovative software along the pipeline is essential to the success of the e-Science programme. To paraphrase a statement from the NSF-CSI Report on Planning for Cyberinfrastructure Software: e-Science software is infrastructure that must be acquired or constructed, stabilized and deployed if the promise of e-Science is to be realized (Blatecky & Messerschmitt 2005, p. 4).

As noted above, the NCeSS nodes are accumulating demonstrators that illustrate the potential of advanced computing technologies to support social science research in a variety of ways. This process has increasingly revealed that

the innovation pipeline, from initial research through to stable and supported tools, includes stages that are difficult to sustain in a research project-based funding regime. The reasons for this are twofold: the portfolio of skills needed and the nature of the funding required.

It is computer scientists, working closely with social science colleagues, who initiate the innovation process by creating new tools or adapting existing ones, and their careers advance by virtue of their skills in innovating to the level of proof-of-concept or prototype. While such outputs might form the basis of demonstrators to draw in the 'interested' segment of the potential audience for e-Social Science, further development by software engineers is required to make the tools stable and platform independent, to ready them for the 'unengaged'. However, as the International Review of the UK e-Science Programme observes 'The "software engineering gap" continues to be a problem in research computing' (Atkins *et al.* 2010, p. 37). The software engineering skills essential to create new research infrastructure and tools rarely attract rewards in academia and there is no established long-term career structure within UK higher education for the developers needed to productionize research outputs. In the third stage of the innovation pipeline, if tools are to be successfully mainstreamed, the software has to be maintained and support for end-users has to be delivered, alongside the provision of documentation, bug-fixes, updates and training. Again, this type of support activity is uncommon in universities and rarely carries academic rewards in research-led institutions. In other words, although the skills and reward system to support the first stage of the innovation pipeline are well established in university computer science departments, the very different skills needed for software to progress through the second and third stages are not.

The second two stages of the innovation pipeline are also problematic from the point-of-view of the Research Councils' short-term project-oriented funding, with its focus on novelty, that is, the first stage only. Although the Councils promote technology transfer, they do not offer funding streams of sufficient magnitude to take innovations onward from the prototype stage. Unless the development of e-Social Science, or e-research more generally, is to be left entirely to the vagaries of the market, new business models are required that provide a route from research to fully productionized, maintained and supported tools, and provide some prospect of long-term sustainability. If potential users, especially the currently unengaged, have doubts about sustainability, they will shy away from new technologies, fearing that time learning how to use them will be wasted if they are not maintained or disappear. The quest for sustainability—the implementation of business models to support the cost-effective deployment of productionized tools, techniques and services—is therefore increasingly pressing, especially as the UK e-Science programme with its earmarked funding draws to a close (Knowles *et al.* 2007; Walk 2009). There is a plethora of alternative business models, including those that derive income (or input) from: volunteer effort, grant-aid, regular subscription, licensing, services, outright sales, consultancy, customization, and linked advertisements. There is no single best model; undoubtedly different parts of e-infrastructure will be most appropriately supported by different mixes of the models, with different contributions from the wide variety of agencies that have a stake in the future of UK e-Science. These include the Treasury and government departments (especially the UK Department of Business, Innovation and Skills,

now responsible for university funding), the Higher Education Funding Councils, the Research Councils, the Joint Information Systems Committee (JISC), and universities and individual users themselves, as well as the private sector.

Open source is often proposed as a solution to sustainability. It is, however, a model of software development and support rather than a business model, because it is compatible with a variety of different inputs or income streams. The open-source route is being pursued by the EPSRC and JISC-funded Open Middleware Infrastructure Institute in the UK (OMII-UK), the EPSRC-funded Software Sustainability Institute (SSI), and the National Science Foundation (NSF)-funded Software Development for Cyberinfrastructure (SDCI) Programme in the USA (which is an extension of the earlier NSF Middleware Initiative). All three are devoted to sustaining software as a critical component of e-infrastructure. OMII-UK has a Commissioned Software Programme that funds the development of software tools, in particular the productionizing of research prototypes (<http://www.omii.ac.uk/wiki/Funding>). SSI offers researchers guides, consultancy and collaborative projects ‘to ensure the sustainability of the software that is important to their research’ (<http://www.software.ac.uk/index.html>). Similarly, as the synopsis of the SDCI programme states, its purpose ‘is to develop, deploy, and sustain a set of reusable and expandable software components and systems that benefit a broad set of science and engineering applications [in]...the current expanded vision of cyberinfrastructure’ (<http://www.nsf.gov/pubs/2010/nsf10508/nsf10508.htm>). Yet none of these initiatives is sufficiently extensive to offer a universal solution. Indeed, the very unpredictability of innovation and its take-up are major challenges to any attempt at a structured solution to sustainability: how can it be decided which tools should be supported? Guessing which will have the most beneficial impact on science in the future is notoriously difficult.

8. e-Social Science of the future

The social sciences are faced with a deluge of digital data as people’s lives increasingly involve engagements with computers that leave traces, expanding their digital footprints by the day. Some of these data are user-generated as in posts on blogs, wikis and social networking sites, responses to Web surveys and crowdsourcing exercises, ratings of products and services, and photos and videos uploaded to servers, along with their tags. Other digital data comprise administrative records, and people’s computer-mediated transactions with public and private sector institutions. In addition, people’s spatial locations are tracked by GPS devices, such as satnav systems or some mobile phones. There are also surveillance data from CCTVs and phone taps and systems-generated data, such as mobile phone logs.

Proponents of e-Science are enthusiastic that the increasing availability of vast amounts of data about every facet of human activity, combined with the incorporation of immensely powerful computers and high-speed electronic networks into the research enterprise, offers the opportunity to transform the social sciences, especially if virtual research environments encourage the most talented scholars to collaborate across the world. As the report of the 2009 International Review of the UK e-Science Programme observes:

‘social science research is on the verge of being transformed through distributed global collaborations, the use of very large data collections, terascale computing resources, and high performance visualization in a way even more fundamental than research in the physical and life sciences’ (Atkins *et al.* 2010, p. 29) This enthusiasm has been amplified as the e-Social Science programme has expanded from the initial sharp focus on the Grid, which has had little take-up in the social sciences, to embrace an increasing range of innovations in digital technologies, many under the Web 2.0 umbrella.

Given the enormous amount of digital data that we generate as citizens in a digital world and the scale and variety of computing resources available to help us make sense of it, there are real opportunities to conduct a sort of social science that was impossible only a decade ago. For example, routinely recorded transaction and administrative data potentially frees us from reliance on relatively small sample surveys that ask people what they do; instead, we have enormous bodies of digital data about what people actually do, and when and where. Similarly, multi-function mobile phones already act as real-time sensors of their owners’ activities and their scope will continue to expand as they are used to deliver additional services. Savage & Burrows (2007) argue that only by embracing the opportunities that new forms of social data make possible can academic sociology be rescued from irrelevance, left behind by the numerous private and public sector organizations that are already exploiting this wealth of data to refine their products and services,

On the one hand, there are reasons to be optimistic, as e-Science tools being developed by the NCeSS nodes begin to deliver advances in social research. In this respect, the NCeSS programme is providing strong foundations on which to build for the future, as noted by the 2009 International Review (Atkins *et al.* 2010). The second-phase nodes form a well-integrated set with maturing synergies (both within NCeSS and with key external initiatives). Moreover, alongside the ferment of technical innovation, a generation of ‘digital natives’ (Prenkys 2001) is entering the research community, bringing with them a willingness—even eagerness—to experiment with innovative tools and services. In addition, the key challenges of the age, as identified by the Research Councils, are of the magnitude and complexity that demand for their solution collaboration across disciplines, together with huge amounts of heterogeneous data and scalable computer resources. For example, the ESRC’s Strategic Plan 2009–2014 includes ‘seven areas of strategic challenge’ (ESRC 2009). The research advances made possible by the NCeSS nodes will be critical to the achievement of the plan’s goals (table 1). Similarly, two NCeSS nodes are strongly embedded in the Research Councils UK Digital Economy Programme, the latest large-scale, cross-council research initiative (<http://www.rcukdigitaleconomy.org.uk/>). Such developments provide the e-Social Science programme with significant new pathways for widening adoption.

On the other hand, there are reasons too to be pessimistic. Radical ambitions for transforming everyday social science have been tempered in the light of growing evidence about the very real barriers slowing widespread adoption of advanced ICT tools and services across the social science community (Voss *et al.* 2008, 2010). It is unlikely that drawing instead on an unsystematic mix of technically limited Web 2.0 applications will radically transform research practice, certainly if they lack interoperability and scalability. Unfortunately, many of the

Table 1. Mapping of NCeSS nodes to ESRC strategic challenges. (The number of + signs indicates the extent of the contribution. Based on Halfpenny & Procter (2008).)

ESRC strategic challenges	NCeSS second-phase nodes									
	DAMES	DReSS	LifeGuide	PolicyGrid	Obesity e-Lab	GENeSIS	OeSS	e-Stat		
global economic performance, policy and management	++			+		+	+			+
environment, energy and resilience	+			++		++	+			+
new technology, innovation and skills	++	+		+		++	++			+
understanding individual behaviour	++	++	++	++	++	++	+			++
social diversity and population dynamics	++	++	+	++	+	++	++			++
security, conflict and justice	+	+				+	+			+
health and wellbeing	++	+	++	+	++	++	+			++

incentives that might persuade social scientists to embrace the more technically innovative methods that would enable them to fully exploit the new and extensive data sources lie outside the control of the e-Social Science programme itself. These include revisions to the academic reward system to promote data sharing and collaboration within and across disciplines; a career structure for software developers and support staff to work alongside teams of computer and social scientists; a break from the short-termism of project grants and, instead, a willingness to offer secure funding for cumulative research over the long term; coordinated implementation by a broad range of stake-holders of a mix of business models that can drive innovations along the full length of the innovation pipeline towards sustainability, and much more focused technical support at the elbow of social science researchers. Without a concerted effort by government departments, through research councils to local research support services to overhaul the context within which e-Social Science operates, the programme may well dissipate into a few isolated remnants, like so many initiatives in the past.

Looking ahead, as the technical landscape continues to evolve and the global social research challenges become more pressing, it is not sufficient to ensure only that the current e-Social Science programme achieves its potential. As the International Review of the UK e-Science Programme stresses: ‘the [e-Science] Programme should not be viewed as a one-shot [...] project, but rather as a long-term strategy that needs to be continuously refined and carried forward as a more permanent crosscutting Programme with real authority and resources’ (Atkins *et al.* 2010, p. vii). We argue that, for its part, it is essential that UK social science continues to invest in its capacity to experiment with, adopt and exploit technical innovations if it is to remain world-leading.

References

- Ainsworth, J. & Buchan, I. 2009 e-Labs and work objects: towards digital health economies. *Proc. 1st Int. ICST Conf. on Communications Infrastructure, Systems and Applications in Europe, LNICST*. London, UK: Springer.
- Ananiadou, S., Procter, R., Rea, B., Sasaki, Y. & Thomas, J. 2007 Supporting systematic reviews using text mining. *Proc. of the 3rd Int. Conf. on e-Social Science*, Ann Arbor, MI. See <http://www.merc.ac.uk/sites/default/files/events/conference/2007/papers/paper208.pdf>.
- Anderson, A. 2003 *Human-centred design and Grid technologies*. Swindon, UK: ESRC.
- Atkins, D. *et al.* 2010 *Building a UK foundation for the transformative enhancement of research and innovation: report of the international panel for the 2009 review of the UK research councils e-science programme*. Swindon, UK: RCUK. See <http://www.epsrc.ac.uk/SiteCollectionDocuments/Publications/reports/RCUKe-ScienceReviewReport.pdf>.
- Birkin, M., Turner, A., Wu, B., Townend, P. & Xu, J. 2007 An architecture for social simulation models to support spatial planning. In *Proc. of the 3rd Int. Conf. on e-Social Science*, Ann Arbor, MI. See <http://www.merc.ac.uk/sites/default/files/events/conference/2007/papers/paper214.pdf>.
- Birkin, M., Turner, A., Wu, B., Townend, P., Arshad, J. & Xu, J. 2009 MoSeS: a Grid-enabled spatial decision support system. Special Issue on e-Social Science, *Soc. Sci. Comput. Rev.* **27**, 493–508. (doi:10.1177/0894439309332295)
- Birkin, M., Procter, R., Allan, R., Beckhofer, S., Buchan, I., Goble, C., Hudson-Smith, A., Lambert, P., De Roure, D. & Sinnott, R. 2010 Elements of a computational infrastructure for social simulation. *Phil. Trans. R. Soc. A* **368**, 3797–3812. (doi:10.1098/rsta.2010.0145)
- Blatecky, A. & Messerschmitt, D. 2005 Planning for cyberinfrastructure software. NSF-SCI Workshop Report. See http://www.nsf.gov/od/oci/ci_workshop/workshopreport_final_rev2a.pdf.

- Brundell, P. *et al.* 2008 The experience of using the Digital Replay System for social science research. *Proc. of the 4th Int. e-Social Science Conf.*, June, Manchester. See <http://www.merc.ac.uk/sites/default/files/events/conference/2008/programme/thurs/1cbrundell.pdf>.
- Buchan, I., Sufi, S., Dunlop, I., Hiroeh, U., Conoy, D., Ainsworth, J., Dale, A. & Goble, C. 2009 Obesity e-lab: connecting social science via research objects. *Proc. of the 5th Int. Conf. e-Social Science*, Cologne, June. See <http://www.merc.ac.uk/sites/default/files/events/conference/2009/papers/Buchan.pdf>.
- Buyya, R., Yeo, C. S., Venugopal, S., Bronberg, J. & Brandic, I. 2009 Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Comput. Syst.* **5**, 599–616. (doi:10.1016/j.future.2008.12.001)
- Carter, R. & Adolphs, S. 2008 Linking the verbal and visual: new directions for corpus linguistics. *Lang. Comput.* **64**, 275–291. (doi:10.1177/1468794109105032)
- Carusi, A.-M. & Jirotko, M. 2009 From data archives to ethical labyrinths. *Qual. Res.* **9**, 285–298.
- Cole, K., Schurer, K., Beedham, H. & Hewitt, T. 2003 Grid enabling quantitative social science datasets—a scoping study. Swindon, UK: ESRC.
- Crabtree, A., French, A., Greenhalgh, C., Rodden, T. & Benford, S. 2007 Digital records: synchronizing visual data with diverse resources, *Presented at the Computer Assisted Qualitative Data Analysis (CAQDAS) Conf.*, April 18–20, Royal Holloway, UK: University of London.
- Daw, M. *et al.* 2007 Developing an e-infrastructure for social science. *Proc. of the 3rd e-Social Science Conf.*, Ann Arbor, MI, October. See <http://www.merc.ac.uk/sites/default/files/events/conference/2007/papers/paper110.pdf>.
- Dutton, W. H. & Meyer, E. T. 2009 Experience with new tools and infrastructures of research: an exploratory study of distance from, and attitudes toward, e-research. *Prometheus* **27**, 223–238. (doi:10.1080/08109020903127802)
- Edwards, P. *et al.* 2009 e-Social science and evidence-based policy assessment: challenges and solutions. Special Issue on e-Social Science, *Soc. Sci. Comput. Rev.* **27**, 553–568. (doi:10.1177/0894439309332305)
- Elliot, M., Pickles, S., Purdam, K. & Smith, D. 2005 Disclosure risk and Grid computing. *4th UK e-Science All Hands Meeting*, See <http://www.allhands.org.uk/2005/proceedings/papers/485.pdf>.
- EPSRC 2009 Introduction to e-science programme. See <http://www.epsrc.ac.uk/about/progs/rii/escience/Pages/intro.aspx>.
- ESRC 2009 ESRC Strategic Plan. 2009–2014 Swindon, UK: ESRC. See http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Images/Strategic%20Plan%20FINAL_tcm6-32704.pdf.
- Fielding, N. 2003 *Qualitative research and e-social science: appraising the potential*. Swindon, UK: ESRC.
- Fox, G. C., Guha, R., McMullen, D. F., Mustacoglu, A. M., Pierce, M. E., Topcu, A. E. & Wild, D. J. 2007 Web 2.0 for Grids and e-Science. In *Proc. 2nd INGRID-instrumenting, Santa Margherita Ligure, Portofino, Italy, 16–18 April*.
- Fraser, M., Hindmarsh, J., Best, K., Heath, C., Biegel, G., Greenhalgh, C. & Reeves, S. 2006 Remote collaboration over video data: towards real-time e-social science. Special Issue on Collaboration in e-Research (eds M. Jirotko, R. Procter, T. Rodden & G. Bowker). *J. Comput. Supported Cooperative Work* **15**, 257–279.
- Fraser, M., McCarthy, M. R., Shaukat, M. & Smith, P. 2007 Seconds matter: improving distributed coordination by tracking and visualizing display trajectories. *Proc. of SIGCHI Conf. on Human Factors in Computing Systems*, pp. 1303–1312. New York, NY: ACM Press.
- Gibson, J., Tekiner, F., Halfpenny, P., Nazroo, J., Fagan, C., Procter, R. & Lin, Y. 2007 Data mining for social scientists. *Proc. of the 3rd Int. Conf. on e-Social Science*, Ann Arbor, MI. See <http://www.merc.ac.uk/sites/default/files/events/conference/2007/papers/paper210.pdf>.
- Goldberg, D. T. 2009 Networking knowledge. Keynote address. In *Proc. 5th Int. Conf. on e-Social Science, Cologne, Germany, 24–26 June 2009*.
- Grose, D. In press. High throughput distributed computing using R: the multiR package. *J. Statist. Softw.*

- Grose, D. J., Harris, R., Brundson, C. & Kilham, D. 2007 Grid enabling geographically weighted regression. *Proc. of the 3rd Int. Conf. on e-Social Science*, Ann Arbor, MI. See <http://www.merc.ac.uk/sites/default/files/events/conference/2007/papers/paper147.pdf>.
- Halfpenny, P. & Procter, R. 2008 *The e-Social Science Roadmap 2009–2012*. University of Manchester. See <http://www.merc.ac.uk/sites/default/files/roadmap.pdf>.
- Halfpenny, P. & Procter, R. 2009 Editorial. Special Issue on e-Social Science. *Soc. Sci. Comput. Rev.* **27**, 459–466. (doi:10.1177/0894439309332662)
- Halfpenny, P., Procter, R., Lin, Y. & Voss, A. 2009 Developing the UK e-social science research programme. In *e-Research, transformation in scholarly practice* (ed. N. Jankowski). New York, NY: Routledge.
- Hardisty, A., Chue Hong, N., Procter, R., Hughes, L., Voss, A. & Fraser, M. 2008 Birds of a Feather (BoF) Session on ‘e-Infrastructure: tool for the elite or tool for everybody?’ *8th UK e-Science All Hands Meeting*, University of Edinburgh. See http://www.nesc.ac.uk/technical_papers/UKeS-2009-01.pdf.
- Harris, R., Singleton, A., Grose, D., Brundson, C. & Longley, P. 2009 Grid-enabling geographically weighted regression: a case study of participation in higher education in England. *Trans. GIS* **14**, 43–61. (doi:10.1111/j.1467-9671.2009.01181.x)
- Hey, T. & Trefethen, A. 2004 UK e-Science programme: next generation Grid applications. *Int. J. High Performance Comput. Appl.* **18**, 285–291. (doi:10.1177/1094342004046054)
- Hielkema, F., Edwards, P., Mellish, C. & Farrington, J. 2007 A flexible interface to community-driven metadata. *Proc. of the 3rd Int. Conf. on e-Social Science*, Ann Arbor, MI. See <http://www.merc.ac.uk/sites/default/files/events/conference/2007/papers/paper144.pdf>.
- Hudson-Smith, A., Batty, M., Crooks, A. & Milton, R. 2009 Mapping for the masses: accessing Web 2.0 through crowdsourcing. Special Issue on e-Social Science (eds P. Halfpenny & R. Procter). *Soc. Sci. Comput. Rev.* **27**, 524–538.
- Kewley, J., Braimah, A. L., Allan, R. J., Hayes, M., Brorsson, P., Crouchley, R., Grose, D. & van Ark, T. 2007 GROWL scripts: lightweight access to Grid resources. *6th UK e-Science All Hands Meeting*, Nottingham, UK. See <http://www.allhands.org.uk/2007/programme/download4b33.html?id=871&p=paper>.
- Knowles, A., Chue Hong, N., Goble, C., Austin, J., Sudlow, A., Lloyd, S., Lock, Z., Linde, T. & Nicholson, J. 2007 UK e-Science core programme: business models for sustainability. *Models of Sustainability Workshop*, Oxford eResearch Centre. See http://www.jisc.ac.uk/media/documents/programmes/einfrastructure/day2_breakoutbusinessmodels.pdf.
- Lambert, P. S., Tan, K. L. L., Gayle, V., Prandy, K. & Bergman, M. M. 2008 The importance of specificity in occupation-based social classifications. *Int. J. Sociol. Soc. Policy* **28**, 179–192. (doi:10.1108/01443330810881231)
- Meyer, E. T. & Schroeder, R. 2009 Untangling the Web of e-research: towards a sociology of online knowledge. *J. Informetrics* **3**, 246–260. (doi:10.1016/j.joi.2009.03.006)
- Milton, R. & Steed, A. 2007 Mapping carbon monoxide using GPS tracked sensors. *Environ. Monitor. Assess.* **124**, 1–19. (doi:10.1007/s10661-006-9488-y)
- Pignotti, E., Edwards, P. & Preece, A. 2007 Semantic workflow management for e-social science. *Proc. of the 3rd Int. Conf. on e-Social Science*, Ann Arbor, MI. See <http://www.merc.ac.uk/sites/default/files/events/conference/2007/papers/paper174.pdf>.
- Prenksy, M. 2001 Digital natives, digital immigrants. *On the Horizon*. **9**, 1–6. See <http://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf>.
- Procter, R., Rouncefield, M., Lin, Y-W, Poschen, M. & Voss, A. In press. Agile management: strategies for developing a virtual research environment. Special issue on embedding e-Science. *CSCW J.*
- Research Councils UK 2009 *Definitions*. See <http://www.rcuk.ac.uk/escience/definitions.htm>.
- Rogers, E. M. 1995 *Diffusion of innovations*. New York, NY: Free Press.
- Savage, M. & Burrows, R. 2007 The coming crisis of empirical sociology. *Sociology*. **41**, 885–899. (doi:10.1177/0038038507080443)
- Sinnott, R.O., Stell, A. J. & Ajayi, O. 2008 Supporting grid based clinical trials in Scotland. Special Issue on Integrated Health Records. *Health Informatics J.* **14**, 79–93.

- Sørensen, K. H. 1996 Learning technology, constructing culture: socio-technical change as social learning. *STS Working Paper* No. 18/96, Centre for Technology and Society, Trondheim, Norway.
- Tan, L., Lambert, P., Turner, K. J., Blum, J., Gayle, V., Jones, S., Sinnott, R. & Warner, G. 2009 Enabling quantitative data analysis through e-Infrastructure. Special Issue on e-Social Science. *Soc. Sci. Comput. Rev.* **27**, 539–552. (doi:10.1177/0894439309332647)
- Thelwall, M. 2009 *Introduction to webometrics: quantitative Web research for the social sciences*. San Rafael, CA: Morgan and Claypool.
- Thew, S., Jarvis, P., Ainsworth, J. & Buchan, I. 2010 Obesity atlas and methodbox: towards an open framework for sharing public health intelligence workflows. *Proc. MedInfo 2010 Conf.*, Cape Town, South Africa.
- Thomas, R. & Walport, M. 2008 Data sharing review report. See <http://www.justice.gov.uk/reviews/docs/data-sharing-review-report.pdf>.
- Townend, P., Xu, J., Birkin, M., Turner, A. & Wu, B. 2007 Modelling and simulation for e-social science: current progress. In *Proc. 6th UK e-Science All Hands Meeting, Nottingham, UK, 10–13 September 2007*. See <http://www.allhands.org.uk/2007/proceedings/papers/889.pdf>.
- Tutt, D., Hindmarsh, J., Shaukat, M. & Fraser, M. 2007 The distributed work of local action: interaction amongst virtually collocated research teams. *Proceedings of the European Conference on Computer-Supported Cooperative Work*, pp. 199–218. Limerick, Ireland. Kluwer.
- 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. *Oxford e-Research Conference*, Oxford, UK. See <http://ora.ouls.ox.ac.uk/objects/uuid%3A06b234c6-7eec-43ab-9ce7-b970f43d4b32/datastreams/ATTACHMENT01>.
- Voss, A., Asgari-Targhi, M., Procter, R. & Fergusson, D. 2010 Adoption of e-Infrastructure services: findings, issues and opportunities. *Proc. R. Soc. Phil Trans A*. This volume.
- Waldrop, M. 2008 Science 2.0: great new tool, or great risk? *Scient. Am.* See <http://www.scientificamerican.com/article.cfm?id=science-2-point-0>.
- Walk, P. 2009 *NSF Workshop on Cyberinfrastructure Software Sustainability*. Indiana University. See <http://www.ariadne.ac.uk/issue59/nsf-2009-03-rpt/>.
- Warner, G., Blum, J., Jones, S., Lambert, P., Turner, K., Tan, L., Dawson, A. & Bell, D. 2009 A data management infrastructure for e-social science. In *Proc. 8th UK e-Science All Hands Meeting, Oxford, UK, 7–9 December 2009*.
- Webb, J. Joseph, J., Yardley, L. & Michie, S. 2010 Using the Internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J. Med. Internet Res.* **12**.
- Woolgar, S. 2003 *Social shaping perspectives on e-science and e-social science*. Swindon, UK: ESRC.
- Yang, Y., Osmond, A., Weal, M., Wills, G., De Roure, D., Joseph, J. & Yardley, L. 2009 LifeGuide: an infrastructure for empowering behavioural intervention research. *8th UK e-Science All Hands Meeting*.