Social Science Information

http://ssi.sagepub.com/

e-Sciences as research technologies: reconfiguring disciplines, globalizing knowledge

Ralph Schroeder

Social Science Information 2008 47: 131
DOI: 10.1177/0539018408089075

The online version of this article can be found at:
http://ssi.sagepub.com/content/47/2/131

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
SSI
Maison des Sciences de l'Homme

SSI

Ralph Schroeder

e-Sciences as research technologies: reconfiguring disciplines, globalizing knowledge

Abstract. *This article examines recent e-science initiatives through the lens of the concept of 'research technologies'. It has been argued that e-science research, which makes use of advanced computing tools to share distributed resources via networks, changes the disciplinary nature of research towards greater interdisciplinarity and paves the way for the increasing globalization of research. However, these claims need to be instantiated in concrete research practices. The essay therefore presents three examples of research projects where these two features can be demonstrated. More generally these three projects – in social science hyperlink analysis, high-energy physics, and astronomy – are examples of 'research technologies', which, it has been argued, are often a radical source of innovation. The article describes how the three projects illustrate these arguments about research technologies, but also how this concept is limited as e-science research is still ongoing. The conclusion assesses how the notion of research technologies is useful for understanding how networked computing technologies are changing the current landscape of knowledge production.*

Key words. *Disciplinarity – e-Science – Globalization – Innovation – Research technologies*

Résumé. *Cet article examine les initiatives récentes dans le domaine des e-sciences au travers du prisme du concept de 'technologies de la recherche'. Divers arguments ont été mis en avant, qui tendent à affirmer que la recherche en mode e-science, en utilisant des outils de calcul avancés pour mettre en commun des ressources distribuées par l'intermédiaire de réseaux, change la nature disciplinaire des recherches, induisant une plus grande interdisciplinarité et pave la route vers une globalisation accrue de la recherche. Cependant, ces affirmations doivent se vérifier dans les pratiques concrètes de recherche. L'article présente trois exemples de projets de recherche où ces deux*

Social Science Information © SAGE Publications 2008 (Los Angeles, London, New Delhi and Singapore), 0539–0184

DOI: 10.1177/0539018408089075 Vol 47(2), pp. 131–157; 089075

caractéristiques peuvent être mises en évidence. De façon plus générale, ces trois projets – analyse d'hyper liens en sciences sociales, physique des hautes énergies, et astronomie – sont autant d'exemples de 'technologies de la recherche' qui, on l'a souvent affirmé, sont la plupart du temps une source radicale d'innovation. L'auteur montre comment les trois projets de recherche illustrent ces arguments sur les 'technologies de la recherche', mais aussi dans quelle mesure ce concept est limité puisque la recherche en mode e-science est encore en devenir. La conclusion fait état de ce pourquoi la notion de 'technologies de la recherche' est utile pour comprendre comment les technologies de calcul mises en réseau sont en train de modifier le paysage actuel de la production de la connaissance.

Mots-clés. *Disciplinarité – e-Science – Globalisation – Innovation – Recherche*

In recent years there has been a rapid increase in the use of the Internet, the World Wide Web and other computing tools in research. One area that has become particularly prominent is 'e-science', the use of shared computing resources for research, which is often said to have profound implications for the nature of scientific practice. As John Taylor, the former head of the UK e-science programme, put it, e-science is 'about global collaboration in key areas of science and the next generation of infrastructure that will enable it' (Taylor, 2001). Another set of implications revolves around fundamental changes in disciplinary practices, both within and between disciplines (Nentwich, 2003: 447–8). Sceptics, on the other hand, point out that it is impossible to identify these changes in the day-to-day practices of researchers and that these visions of e-science are merely that (e.g. Hine, 2006a).

This article argues that the reconfiguration of disciplines and the globalization of knowledge is indeed a feature of e-science, but that sceptics have a point insofar as these changes need to be demonstrated in practice rather than merely asserted. To do this the article draws on the concept of research technologies (Shinn & Joerges, 2002) and gives an account of how e-science tools exemplify this concept. Research technologies, according to Shinn and Joerges, reconfigure disciplines and extend the reach of knowledge by diffusing across boundaries, a 'practice-based universality' (2002: 245). The article describes three examples of e-science projects and argues that they provide illustrations of some of the essential features of research technologies: namely that these instruments disembed knowledge practices and provide translations between different knowledge domains. These changing practices are evidenced in different ways in the three e-science projects discussed here, but there are also common elements among all three, such as the shift towards putting data online and the increasing manipulability of these data by means of software tools. All three projects thus demonstrate in a concrete manner how disciplines interact with each other in new ways and how knowledge is moving towards a global scale, even if, in contrast with well-established research technologies, these changes are still ongoing.

Not all e-science fits the definition of research technology, and not all e-science exemplifies the changes that are described here. Within the sociology of science and technology, however, a debate about the relation between 'local' knowledge and the universality of science is ongoing. This article argues that the research described here is evidence of the trans-locality of knowledge, even if, again, this potential universality is still in the process of being instantiated. Further, there are extensive discussions in science policy about how to organize scientific efforts of an increasingly complex scale and scope, entailing collaboration that is interdisciplinary while also being distributed among different institutions and geographical locations. e-Science is a good place to examine these changes on the research frontier.

First, the article provides some background to the emergence of e-science and debates around the relations between disciplines and the geographical scope of knowledge. Next it introduces the concept of research technologies and how this concept can be applied to e-science. The following section presents three cases of e-science projects, drawn from high-energy physics, astronomy and social science analysis of the Web (webometrics), and gives an account of how different disciplines interrelate and how knowledge is globalized in each project. The conclusion will locate these and other e-science projects within the current research landscape.[1]

Background: e-science, globalization and disciplines

Recent years have seen a number of initiatives to promote e-science. Many of these have been under the umbrella of the programmes of national funding bodies (see overviews in Gentzsch, 2006; Schroeder & Fry, 2007), but there are also individual research projects, which have not been funded or coordinated by these programmes. A number of terms have been used: e-science (including e-social science and e-humanities, in which case the term e-research is sometimes used), cyber-infrastructure, and e-infrastructures. These terms are still in flux. To avoid an overly narrow definition, but one that pinpoints the novel aspects of this type of research, e-science here stands for the use of computing tools for networking in order to share distributed digital resources in scientific or academic research.[2]

One of the visions for e-science is that it is aimed at promoting global collaboration.[3] This can be said not only for the technology that underlies the Internet and the Web, it also applies to the institutional dimension of e-science:[4] even if the content of the various forms of e-science initiatives is different, there are global similarities in structure and form, such as the setting up of national offices for the promotion of e-science (Gentzsch, 2006). These have different names and are dominated by the developed societies, but they follow a pattern whereby nation-states have set up similar institutions and/or organizations for the promotion of scientific research in the course of development. This aspect of globalization is part of a wider trend towards supranational coordination efforts among non-governmental organizations (NGOs) to address a variety of political and economic issues (Slaughter, 2004), but also scientific and technological ones (Drori et al., 2003). In short, there are some organizational initiatives within e-science that aim to span the globe organizationally, in addition to the worldwide institutional isomorphisms in the national organization of e-science.

Arguably one of the main forms of organization of e-science is therefore neither global nor 'local' (based on individual e-science projects) but national, centred on the resources that fund large-scale research initiatives. This national character of e-science, which is aimed at national capacity-building or infrastructures, follows historical precedent since 'national systems of innovation' (Edquist, 1997) have dominated research policy, and 'big science' initiatives (Galison & Hevly, 1992) and 'large technological systems' (Hughes, 1987) are often bounded by nation-states.

At the same time, however, in addition to the homology between national e-science policies, these policies are also increasingly coordinated internationally. It is true that the attempts to coordinate large-scale e-science projects across borders often make visible both the technical and organizational obstacles to this process. These barriers include the silo-ing of online material (national health databases), restrictive policies for accessing computing resources within national academic 'grids',[5] different regimes for intellectual property rights, and the like. To this can be added commercial barriers such as the boundaries of secure networks in firms and subscriber-only availability of online research services such as digital libraries. e-Science thus provides a good domain for investigating the global and non-global dimensions of scientific knowledge. If the globalization of science has previously been identified in terms of the diffusion of institutions (Drori et al., 2003), it can be asked whether it is possible to add to this a worldwide network of tools and the development of internationally linked systems which also need, for example, to coordinate access and standardize procedures on a global level.

Apart from the geographical scope of e-science, this type of research also spans across disciplines. It has been argued that e-science shifts boundaries in different disciplinary domains in quite different ways (see contributions in Hine, 2006b). In one sense, however, e-science inevitably involves a number of disciplines since at a minimum there are two: computer science plus a domain-specific science. There have been many debates about the difficulties of interdisciplinary research collaboration. Cummings & Kiesler (2005) argue, however, that inter-institutional collaboration often causes greater difficulties than distributed collaboration and crossing disciplinary boundaries in the case of large-scale multi-disciplinary and multi-institutional projects that are similar to e-science. And as Nentwich (2003) has shown, the extent to which different disciplines have embraced e-science varies, and is not necessarily, as might be expected, along the natural science versus social science versus humanities divide since, for example, e-science is also pursued in many projects in humanities disciplines.

e-Science thus tends to be more complex than other scientific endeavours inasmuch as it creates more extensive worldwide and cross-disciplinary linkages than traditional research. In the past, research was often housed in individual departments, research institutes and disciplines. e-Science is not the only factor that promotes new forms of collaboration, but it contributes to this by encouraging research that links organizations across institutions, entails distributed collaboration and involves a number of disciplines.

Research technologies and the reconfiguration of knowledge

Apart from this complexity, however, there is one element that has so far been missing from social-science accounts of how e-science is reconfiguring knowledge: technology. Here the sociology of science and technology provides some useful insights and concepts, and in particular Shinn & Joerges' notion of research technologies (2002; see also Shinn, 2005). Shinn & Joerges argue that research technologies have been critical to the advance of scientific knowledge from the late 19th century onwards. But arguably it is necessary to go even further back in history and follow Collins (1994; see also Schroeder, 2007b: 23–6), who thinks that research instruments are fundamental to the whole of modern science.

Collins's argument is that novel technological artefacts drive the advance of scientific knowledge – rather than the other way around. New technological development in the modern period, he says, led to 'high-consensus rapid-discovery' science (Collins, 1994, 1998: 532–8). What is new about 'high-consensus rapid-discovery science' from about 1600 onwards in Europe, he says, is 'secure knowledge' and 'a train of new results' (Collins, 1994: 157). 'What was discovered', Collins continues:

> was a method of discovery; confidence was soon built up that techniques could be modified and recombined endlessly, with new discoveries guaranteed continually along the way. And the research technologies gave a strong sense of the objectivity of the phenomena, since they were physically demonstrable. The practical activity of perfecting each technique consisted in modifying it until it would reliably repeat the phenomena at will. (1994: 163)

Collins thus avoids an 'idealist' account of science and technology: the advance of scientific knowledge here takes place only in relation to the physical and social worlds conjointly, and never just in the realm of ideas. This meshes well with Hacking's philosophy of science whereby science is the 'adventure of the interlocking of representing and intervening' (Hacking, 1983: 146), a pragmatic and realistic account of the relation between scientific knowledge and the physical or natural worlds. In Collins's terms, this adventure is driven by technologies, and by laboratory apparatus in particular. Collins describes a process whereby there is an 'outward flow of lab technology' and this technology is 'exported into the lay world' (Collins, 1993: 315): 'One machine gives rise to another in a genealogy of succession: by modifying the past machine, or by cloning it from another in the same laboratory, or by a kind of sexual reproduction recombining parts from several existing pieces of equipment' (Collins, 1994: 164). In this way, laboratory technologies ultimately gain legitimacy by means of their diffusion into the everyday practices of consumers (Schroeder, 2007b).

At this point we can extend Collins's ideas by drawing on Shinn & Joerges' concept of research technologies. The concept of research technologies is narrower than that of Collins's laboratory technologies since these are 'generic' or 'open-ended general-purpose devices' (Shinn & Joerges, 2002: 212), instruments that can be used across a range of disciplines. Initially they are developed outside of established disciplines and institutions, and in this sense they are 'interstitial' and 'disembedded', removed from the interests of particular groups. Thus they are able to create a new language and way of representing phenomena that transcends particular disciplines and institutions.

At the same time, these all-purpose devices then become 're-embedded' in multiple local contexts, spreading this shared language and means of representing phenomena across them. Thus research technologies achieve 'a universality grounded in informed and legitimate practice … a practice-based universality' (Shinn & Joerges, 2002: 245).

> Practices are independently repeated and are multiplied in innumerable environments. This is not the objectivity born of pure reason or the experimentum cruces. Objectivization is instead built up through collective practice which is structured around effect-producing materials and procedures ... Objectivization is cumulative and practical. (p.244)

Or again, research technologies consist of 'concrete … practices' rather than abstract scientific knowledge or cultural representations of technology. These include:

> design, hands-on construction, endless tinkering and analysis to probe the deep principles of devices, adaptation to improve performance, explorations and controls to determine the extent to which a generic device can be generalized, trials and modifications to check whether the processes of generalization hold, and transferring apparatus into a local niche environment for tailoring and operation by end-users. (p.217)

In this way, as illustrated by certain tools in the past, 'specialty groups' in different disciplines 'learned to communicate and came to see aspects of their problem domains in the light of how ... [the] instrument represented and dealt with the physical world' (p.217). Once they had done this, they would have 'moved outside the research-technology nexus and into many professions and countries' (p.218). These ideas apply Collins's outward flow of laboratory technologies to how they are used to spread research practices.[6]

This concept of research technologies fits e-science very well. e-Science is being driven by technology developers whose aim is to develop tools that can be applied to a range of disciplines and purposes.[7] In the case of e-science or Grid technologies,[8] these generic devices allow data and other digitized material to be manipulated across networks. Among the characteristics these technologies share are:

- means of finding and classifying relevant data or other resources with the help of digital identifiers. These include ontologies (the semantic web), tagging, and putting data into formats that allow search and retrieval;
- means of providing access to the data and tools through mechanisms to do this securely, 'job submission' processes, and 'portals' – in other words, coordination and control of shared digitized resources;
- standardized middleware[9] and other software protocols;
- ways of distributing the manipulation, storage and communication of data and other digital research resources between different computers (from PCs to high-performance computers) via networks.

These are features of many e-science technologies, including the three that are discussed here.

Grids too, like other research technologies, are thus 'multi-level, multi-domain intelligibility devices', say Shinn & Joerges (2002: 244). They suggest that the 'meta-methodologies and meta-artefacts belonging to generic instruments, and which are re-embedded in local, narrow-niche devices, operate like passports' (p.244). e-Science technologies fit this definition insofar as they represent 'a form of instrument design that consciously takes into account maximizing the variety and number of end-users whose local technologies can incorporate key features of a research-technology template' (pp.212–3). 'Communication between institutionally and cognitively differentiated groups of end-users', they say, 'eventually develops' (p.244). Whether the technologies in e-science do this in practice is, of course, an open question, not least within the e-science community itself.

In respect of each of the e-science projects given as examples here, we therefore need to ask these questions: To what extent, or in what ways, does the project aim to be global in scope or reach? How does the project cross existing disciplinary boundaries? In what sense does the project represent a research technology that instantiates 'practice-based universality'?

At this point, we can turn to three research projects that are developing these tools. The Virtual Observatory for the Study of Online Networks (VOSON) project has created a tool for doing social-science research on online networks; but this, as we shall see, can be applied to many domains. Enabling Grids for e-SciencE (EGEE) is also developing tools for analysing data, among other things by means of sharing computer-processing power. This entails developing a range of other software tools that are applicable to various domains. The International Virtual Observatory Alliance (IVOA) is producing a tool for managing, federating and annotating astronomical data and, again, this can be used for other types of data. As we shall see, the contributions towards extending the global scope of research, reshaping disciplines and developing research technologies are mutually reinforcing. Yet at the same time, in each case, these characteristics are emergent; they are instantiated, but there are also limits. We return to these limits after first describing the three projects.

Project 1. VOSON: a tool for studying online networks

VOSON is a research project to study online networks based at the Australian National University (http://voson.anu.edu.au/). This effort had been ongoing for several years under the leadership of Rob Ackland, but

VOSON began its formal project life only in 2005, when it was funded by the Australian Research Council as part of a special research initiative for e-science support. To date, VOSON activities have focused on the development of new social-science research methods and tools and especially 'webometric' approaches, which use data from the Web such as the hyperlinks between webpages to identify, among other things, the visibility of sites (see also Thelwall, 2006). One goal of the VOSON project is to produce software to enable these new research methods and develop tools that can be shared and used collaboratively by researchers. In this sense VOSON is a good example of e-social science – using e-science technologies to enable new forms of collaborative social-science research. So far, the project has focused mainly on methods and tool development, on one side, and on applying these methods and tools to the analysis of political parties and environmental activist networks, on the other (see the papers at http://voson.anu.edu.au/papers.html). The aim, however, is to extend this approach to a range of online social networks.

A number of the features of this project are noteworthy. One, already mentioned, is that the main thrust of the research is to build tools and data that can be used in collaboration with other researchers. The project has deliberately adopted an 'open source' approach to software development, with the aim of making the tools and data available to other researchers and encouraging them to do the same. The project has also adopted a Creative Commons approach to licensing and follows various 'open' standards (for example, for metadata).

Second, VOSON is part of a small community of researchers using novel tools in a relatively new area of research, and so there has been a highly focused research effort with a great deal of interchange between researchers. For example, VOSON has worked with several institutions worldwide that are in similar areas and with which collaborative research has been undertaken, including the UK, USA, and other parts of the world. In view of the small number of groups and the fact that all researchers have to keep in contact to remain abreast of rapidly moving developments, it is appropriate to speak of a research 'community' here. This community is reinforced by the fact that the researchers share a common approach to social science, which falls under the umbrella label of 'social-network analysis', a larger research community and broader research specialism which also goes beyond social-science disciplines. And while webometrics as an area within e-social science is relatively new, there has been a rapid increase in the number of papers and scholarly interest in this field.[10]

At the same time, there are a number of challenges in this research area: VOSON and similar research projects measure online networks and thus the visibility, for example, of activists and parties, but they do so for the online world. As with webometrics generally, one question concerning these online networks is about their significance for offline relationships: is the main influence through search (if a site has a greater number of network links, it may be easier to find) or is there a more significant way in which the structure of the online world can tell us something about the offline world? If, for example, certain environmentalist groups are more closely linked online with other such groups, what – if any – are the implications for how environmentalist groups work offline (Ackland et al., 2006a)?[11] These challenges will require integration with and testing against other social-science findings.

VOSON illustrates the global scope of knowledge insofar as its analysis of online networks covers the entire (global) Web network as its object of research. The limitation in this case comes back to the challenge mentioned a moment ago: even if the scope of social science in this area is global, how well does the global Web map onto other globalizing processes apart from these online networks? As for disciplinarity, one question is which discipline this project 'belongs' to? This is a problem, for example, for the dissemination of the work: should the work be presented and published in traditional political-science or sociology journals (including specialisms within – but sometimes also beyond – these, such as social-network analysis), or should the publications focus more on tool development and webometrics, and target journals in computer science or information science? The VOSON project has so far taken the direction of conducting social-science research using a variety of methods from other fields, and yet the output has been primarily targeted at social-science journals and conference proceedings. However, this has met with mixed success: there has been enthusiasm about novel methods and insights, but also a lack of understanding about these new methods and about the relevance of Web data to traditional social-science concerns. These obstacles are bound to be overcome in time.

One way to get around the lack of understanding of this new area is, of course, subdisciplinary specialization. Work on webometrics is well suited to new outlets such as conferences and journals for e-science. It can be foreseen that this area, like other such areas, will not only become one small domain within other areas but also form a subdiscipline or specialism of its own. One indicator for this development are the specialist journals devoted to the topic – for example, the journal *Cybermetrics*, founded in 1997 (http://www.cindoc.csic.es/cybermetrics/journal.html). It is also noteworthy that there are a number of disciplines involved, and that physicists, computer scientists, library and information scientists and others – as well as social scientists – have been a strong presence in this type of analysis. VOSON fits this multidisciplinarity, with several computer scientists involved and papers often containing detailed accounts of the advantages and disadvantages of different types of software (Ackland et al., 2006b), but also papers with political scientists and researchers from other social-science disciplines such as media studies (Ackland et al., 2006a). This illustrates the key point that Shinn & Joerges make about research technologies: that knowledge is transferred via the skills needed in the development of generic devices that need to be applicable in a range of settings.

Yet VOSON faces considerable technological hurdles: for example, although it is envisaged that high-performance computing and access to this via the Grid will be needed for VOSON research, the project has so far developed the tools in anticipation of this development but not used Grid capabilities as yet because these are still not mature enough (Ackland et al., 2006b). Nevertheless, VOSON is adapting its tools and methods towards Grid development – and it can be expected from the other side that the Grid and other e-infrastructures will make progress so that research projects can make better use of Grid resources.[12]

Project 2. High-energy physics grid computing and beyond: EGEE

EGEE is a European project, funded mainly by the European Union (EU), to provide a large-scale multidisciplinary Grid infrastructure for the European Research Area (as it is known in EU research policy). Although it started on a smaller scale, it has grown into a large-scale collaboration which provides access to some 30,000 CPUs (central processing units) and several petabytes (or one quadrillion bytes – a megabyte, which will be familiar to PC users, is 10^6 bytes and a petabyte 10^{15} bytes) of storage which are currently used by more than one hundred research groups from several scientific domains and from around the world. This makes EGEE the largest e-science collaboration worldwide in terms of scale, diversity of disciplines involved and perhaps organizational complexity. EGEE is led by CERN (the European organization for nuclear research) in Geneva, a laboratory that has been conducting fundamental research in physics since 1954, focusing on experiments with large-scale particle accelerators. The most powerful particle accelerator, the Large Hadron Collider (LHC), is due to come into operation in 2008, superseding its predecessors and generating data on a larger scale than previous experiments.

The goals of the LHC include identifying the 'Higgs Boson' particle, which is critical to validity of the 'standard model' in particle physics. The experiments to do this will generate many petabytes of data per year that will require the use of shared computing resources across a number of sites. This kind of 'big-science' collaboration, involving many institutions and hundreds of physicists, has been common in particle physics since the middle of the 20th century (Galison, 1997), but the scale of the data that needs to be processed and analysed in this case is unprecedented (Hey & Trefethen, 2006, section 15.2.2).

EGEE began in 2004 as an initial two-year phase of a four-year programme which is now in the second phase (called EGEE-II), though with the anticipation that it will continue into further phases.[13] The first phase began exclusively with European members; but one reason for its inclusion among the projects which aim at a global scale is that, in its second phase (EGEE-II), it has become international, with partners from 32 countries, including the USA, Asia and other parts of the world.[14] These countries are still dominated by the developed countries of the 'Global North', and their geographical scope reflects the fact that e-science is mainly being carried out by the countries that can afford to do so. At the same time, it can be anticipated that EGEE will spread more widely over time.

Similarly, although the project was initially centred on the physics community (extended in the first phase of EGEE to biomedical research), it now includes many different disciplines, even though the main applications of the infrastructure remain high-energy physics and biomedicine (Gagliardi et al., 2005: 1741). For example, the 'user forum' that took place in May 2007 showcased projects from a wide range of disciplines (http://egee-intranet.web.cern.ch/egee-intranet/User-Forum/), and the EGEE II project description lists nine application domains (http://www.dcc.ac.uk/events/policy-2006/EGEE-II%20overview%20paper.pdf, Table 1).

EGEE, again, is a distinctive e-science project because it is the largest such project, whether this is measured by resources (funding), number of partners or networked computer-processing power. In fact, calling it a research 'project' is somewhat odd, as the longer-term aim of EGEE is to become – or be part of – a European e-infrastructure. A major challenge, therefore, which is mentioned in the documents describing the project, is to create a permanent infrastructure from short-term funding programmes and funding cycles. This will entail, according to some of the project leaders involved, finding new models for funding and for maintaining the 'project' or the 'infrastructures' over the long term among research policymakers at the European Commission (Gagliardi et al., 2005: 1741).

The demands of the physics community for high-performance computing also make this a unique project. In this area of physics there is a need for – or a dependency on – the resources that are afforded by sharing high-performance computer processing across a number of the most powerful machines in Europe. This contrasts with the other disciplines that are participating in EGEE, which do not have a need for high-performance computing and are unlikely to have such a need for the foreseeable future. It can be mentioned that in biomedicine, perhaps the next-largest computing-intensive area in EGEE, there is a need for large databases and access to these databases, but storing data entails different needs and challenges from processing large volumes of data.[15]

The relation between physics and other disciplines is thus particularly complex in this case, given the range of disciplines involved. There is bound to be a two-way influence: other disciplines will need to adapt to the 'core' discipline of this project, and the physics 'core' will need to take on board the needs of other disciplines. This may also turn out to be more of a two-way learning exercise on both sides rather than adaptation. In any event, many of the research projects that are part of EGEE are unlikely to get involved in the technical parts of EGEE concerning middleware and other parts of the computing infrastructure, although the methods of organizing large-scale research may still rub off on them.

One example of these shared practices is the long-standing experience in physics of large-scale collaboration: multi-institutional research efforts on a large scale in physics have used memoranda of understanding (MOUs) for the collaborative processing and analysis of data, wherein each institution specifies the resources (funding and computing resources) it will input, and what rights it has to publish the data in return.[16] Perhaps these will become more widely used in non-physics projects. This also applies in the other direction: to what extent will the core physics community, which initiated the project, bend to the needs, both organizational and technical, of the other disciplinary communities in EGEE, such as the contractual arrangements for the EU projects through which the disciplinary communities participate in EGEE? And, finally, this applies to requirements: to what extent will the non-physics disciplines introduce needs, apart from shared high-performance computing, again with the corresponding organizational mechanisms that go along with this?

Another example of shared practices in EGEE is that the project provides elements which are common across the project, such as middleware.[17] The main middleware for EGEE is now called gLite, and use of this tool is required for members of the EGEE consortium. To what extent less technically inclined project partners outside of physics are able to handle this tool, however, is still an open question.

EGEE has taken a multi-level approach to organizing the collaboration in view of the large number of participants. In terms of funding, EGEE is a 'consortium' with a lead partner and other partners – as per the agreement with the European Commission. But there is also a 'federated' structure, which is based on the idea of bringing together the different regional and national Grid services and integrating them within a larger European e-infrastructure (Appleton, 2006, section 6).[18] With this organization, and the regular international meetings and research activities that bring the constituent projects together, it is hoped that a more permanent community can be formed around the EGEE infrastructure.

In view of its size, and as this is a flagship European project, EGEE is bound to be a test case for worldwide e-science collaboration – or the lack thereof. This European effort will have to be compatible with or drive the efforts towards a common infrastructure (standards, data sharing, organizational links) with other major efforts such as the cyber-infrastructure projects in the USA (such as TeraGrid) and the Open Grid Forum (OGF), a worldwide body promoting software and other technical standards. Otherwise EGEE will risk being left out of major e-science developments. The reverse also applies – no major e-science project will be able to ignore the EGEE behemoth in the realm of standards and in terms of how e-science collaboration is organized, both technically and institutionally. This can be said of only a few large-scale projects at this level, where global collaboration in this sense is inescapable.[19] Note also that this gives EGEE, like other large technological systems, a momentum of its own. EGEE cannot fail since it is a flagship European project, and so will need to be kept going even if it will take on different forms.

Geopolitics thus plays a role at this level, and it involves both competition – CERN's LHC competes for status with the US high-energy physics research community – and cooperation. These geopolitics give the project a global reach. And apart from this spatial global reach, there is also a temporal global dimension: the aim of running the EGEE Grid service on an around-the-clock (24/7) basis, including the support for users, can be achieved only by including the American and Asian sites, since this allows coverage of the different time zones around the world (Appleton, 2006, section 2.4).

EGEE is an ongoing project, like the others discussed here, so the characteristics identified here are 'science-in-the-making'. But even if it is not possible to say anything final about the global and disciplinary shape of this (and the other projects), the processes of creating research technology on a large scale are clearly under way in this case. And since particle physics and biomedicine are currently at the leading edge of big-science research in terms of the organizational complexity and scale of research collaboration, these areas may serve as examples for other areas of research. In short, although driven by physics and its need for shared computing resources, the EGEE project will shape the globalizing ambition of e-science, as well as its interdisciplinary ambitions, both as a model and as practical instantiation.

Project 3. Federating astronomical data: IVOA

Astronomy has been a geographically distributed and collaborative enterprise for some time because it has been necessary to use a number of telescopes in remote locations and interpret the data elsewhere. ‘Virtual’ observatories have come into being more recently, the term ‘virtual’ in this case signifying that the data and images can be accessed online independently of using a telescope. The effort to bring the various national virtual observatories together into a single resource under the umbrella of the ‘International Virtual Observatory Alliance’ is more recent still. IVOA was formed in 2002 with the aim of developing international standards for accessing, correlating and manipulating astronomical data (see http://www.ivoa.net/). In other words, the aim was to pool data from all the national virtual observatories into a single resource and make them openly available.

The initiative began with 12 national virtual observatories (VOs) and has grown to more than 16 (in May 2007). Membership is open to new virtual observatories that would like to join this initiative, but they must fulfil certain criteria, which include being a major recognized national effort and being willing to share in the procedures and standards developed by IVOA, including open access to the data.[20] The organization is governed by an executive committee and has coordination meetings several times per year. It has also established a number of working groups and special-interest groups, such as the ‘Theory’ group (Lemson & Colberg, 2004), and groups to support the creation of standards and tools.

The task of developing standards has been particularly promoted by ‘interoperability’ workshops, which have been held twice a year, and also by specialized working groups, which have concentrated on coordinating individual aspects of the larger drive towards standards (e.g. uniform content descriptors for catalogue entries, image-access protocols, the VOSpace collection interface). A number of tools have been developed to support the federation of existing data repositories into a virtual repository that can be accessed from the desktop. The status of working prototypes and standards specifications are in various stages of completion and can be found in the latest roadmap (see under http://www.ivoa.net/pub/info/). And although this roadmap has a timetable for the various aspects of the task, it can be envisioned that this will be an ongoing development with a continual refinement of the tools.

Astronomy is an interesting case in terms of e-science because it is a relatively small community of researchers with a clear and relatively well-defined task.[21] This task requires public funding on a large scale, and it is therefore not surprising that efforts towards international coordination have taken place. Moreover astronomy, like other disciplines, has faced a ‘data deluge’ (Hey & Trefethen, 2003; Hanish & Quinn, n.d.: 1 use ‘data avalanche’), and this lends urgency to seeking common solutions. Here, as in other disciplines, the problems have revolved around putting the data into common standards such that they can be submitted and accessed by researchers in a uniform way. This initiative can be expected to continue until the data are available on the desktop in a stable form and it is possible to analyse entire image collections using Grid-type technology.

This is also an interesting example because data are being made accessible worldwide. Tens of thousands of astronomers have already accessed the service, and since the worldwide community of researchers in this discipline numbers only in the tens of thousands, it can be envisaged that this will become a commonly used tool.[22] Moreover, the task of coping with a number of sources of data has focused the discipline on creating a common interface. Describing the data with common names and identifiers has added computer-science research (such as the searchability of the data via metadata) to the traditional tasks of astronomy, and these classification tools may also be adapted to other disciplines as well as being compatible with them, for example in terms of search.

Apart from bringing data to the desktop, one challenge in projects like IVOA is how the data are to be curated.[23] There are already initiatives to do this (see Choudhury et al., 2006), spearheaded by the National Virtual Observatory in the USA (which is a member of IVOA) to create a single virtual resource accessible via the Web, which brings together the data published in journals and in libraries. In other words there will be an end-to-end process that makes data available from capture to end-user analysis and which is available via a single Virtual Observatory portal with a common worldwide standard.[24]

Frontiers and limits of e-science

These three projects are not a representative sample of e-science. Other projects may be less global in scope, more narrowly contained within disciplinary boundaries, and they may also not fall easily within the rubric of research technologies. Still, the three projects that have been described do meet these criteria: the tools being developed within these projects are not project specific but can be built on and scaled to include more organizations and users worldwide, adapted and combined with other tools, and carried over into other disciplinary domains. They are also clear examples of research technologies for e-science because they are tools for collecting, storing and manipulating large amounts of data and other digital research materials, and making them accessible as a shared and distributed resource for research.

Nevertheless, it is also possible to see that there are a number of limits to research technologies in e-science. To recognize this, we can note, first, that e-science technologies are not synonymous with the Internet or the Web as a whole. e-Science is part of the larger technological system of electronic networks, but only a small part of the Internet and the Web is devoted to research, and an even smaller part fits the definition of e-science used here. Second, within the much smaller domain of e-science in this larger system, there is nevertheless a variety of tools on different scales, ranging from the infrastructure as a whole, to middleware, to individual domain-specific interfaces and software components. Not all of these will fit the notion of research technologies since, for example, different parts of e-science may remain limited to a specific discipline, or may be geographically bounded or specific to one of these layers. Nevertheless, it is evident from the three examples given that they will contribute as e-science technologies on all three levels: for example, all three will provide access to the Grid, include middleware and develop a variety of applications.

At the same time, being part of a larger infrastructure is problematic; it involves a lot of effort at creating and maintaining integration. One example may suffice: the UK has created a repository to make the middleware arising from UK e-science projects available from a central source, the Open Middleware Infrastructure Institute (http://www.omii.ac.uk/). However, it has been estimated that converting the software into a form that can be deposited, which involves providing documentation and describing and implementing standards and interoperability, requires five to ten times as much effort as developing the software in the first place![25] On the other hand, if it is successful, integration entails translation, since many research communities that would otherwise be isolated need to engage with one another, and the software involved in the technological infrastructure allows – or requires – them to do this.

Being part of an emerging infrastructure means that e-science, like other 'big-science' (Galison & Hevly, 1992) efforts and developments of large technological systems, has a momentum of its own (Hughes, 1994). Nevertheless, the allocation of resources devoted to research and different areas of research is limited, and so policies for research funding and research policy generally are bound to impact on this advance, pushing it further in some cases than in others. Whether e-science will continue to be supported and funded is, of course, an open question.

This limitation, the competition for resources, brings us to another: the competition for attention focused on e-science. To recognize this limited attention space, we need to step back for a moment to note the seamlessness – or 'openness' – of science. As Fuchs has argued, scientific communication is in principle 'open' because closing off communication would violate the scientific norm of an endless refinement of knowledge (Fuchs, 2001, 2002). The implication is that the cultural, disciplinary, political, economic or other boundaries that beset other social networks are transcended by scientific communication. The requirement to communicate in the sciences, to publish one's findings in order to be recognized, cuts across all the channels for scientific communication and disciplines (Becher & Trowler, 2001). And although different rules for publication exist in different media, different disciplines and different settings, being first to publish an idea or produce an innovation is thus a 'global' filter.

The flipside of this openness and universality is the limited amount of attention space and competition for resources to support research within e-science: one consequence of the universality of scientific knowledge (and this applies to technological development associated with scientific knowledge) is that there is only ever one leading edge of research for a given domain of phenomena. As Collins has argued, the main constraint on scientific communication is therefore the limited attention space (Collins, 1998: 37–40) in each domain of knowledge advance. Thus, for example, the online resources devoted to e-science – research reports, papers, project descriptions and the like – must vie with each other for attention in the larger scientific community just like other (non-e-sciences) scientific results. This limited attention space means that, in addition to intra-scientific priority (competition to be globally first to publish a result), there is an extra-scientific factor (public attention and funding for research), which constitutes the wider environment that selects, or provides a mechanism for, how research is prioritized.

This competition can be linked to the organizational differences between disciplines. As Whitley (2000) has argued, disciplines are coupled in different ways to their task (degree of 'uncertainty') and in terms of the extent to which researchers are dependent on each other ('mutual dependence'). In the case of the e-sciences, as we have seen, this coupling to tasks and mutual dependence are being reconfigured such that different efforts integrate and coordinate their online – or Grid-enabled – resources. Nevertheless, it also follows from the account of 'high-consensus rapid-discovery sciences' that the leading edge of scientific advance at any point in time is concentrated in particular research areas: namely, there where artefacts and knowledge are coupled in the most powerful and novel ways. Yet the 'migration' between different fields (Fuchs, 1992: 189) makes it difficult to identify where the most important current scientific advances are taking place.

Now e-science, as we have seen, exemplifies this novel coupling in several places. The ability to create new tools does not necessarily mean, however, that there will be a tighter linking between knowledge and technologies, on one side, and tasks and the physical and social worlds in which science represents and intervenes (Hacking, 1983), on the other. Furthermore new technologies, in research as elsewhere, tend to add to and complement, rather than replace and supersede, previous ones. Still, inasmuch as e-science intervenes more powerfully in the physical environment and can move onto new territory, it will capture the attention of researchers, research funding bodies and publics. e-Science thus provides a focus of attention, but it is only one area of knowledge production among others, again, competing for attention amidst the sciences as a whole.

Another limitation stems from the fact that these three projects are at an early stage, and hence it is difficult to gauge whether the technologies discussed here will have the 'malleability' to achieve 'pan-validation across disciplines, institutions, and nations' (Shinn & Joerges, 2002: 248): clearly some of the tools in each case can be applied to other disciplines and are expressly designed to work across institutions and national borders. Yet the extent to which they do so differs for each case. Take, for example, protocols: 'One impact of generic instrumentation is the stimulation of social and intellectual cohesion' which 'sometimes give rise to a brand of universality. The adoption by an end-user audience of a generic instrument entails the audience's integration of protocols which make the instrument effective' (2002: 242–3). As was noted in the case of middleware, however, making these tools compliant with protocols can be a major effort, and extending these and other e-science tools across user communities that have different computing needs and skills has been one of the most often-noted challenges in e-science.

A final limitation is that it is not clear to what extent the impact of e-science will go beyond the boundaries of research. Shinn (2005) argues that one of the reasons research technologies are so important for innovation is that they ultimately become part of consumer technologies. Lasers are a good example of a research technology that was initially a scientific curiosity but subsequently, after several decades, became widely used in consumer technologies (Rosenberg, 1992). But e-science technologies are still at an early point of development, and how they travel outside the domain of academic research and, for example, into commercial research and development remains to be seen. Further, the boundaries between research technologies and other resources may be blurring in the case of e-science: it is possible to see e-science tools such as Grid technologies as research technologies, but they might equally, at least in part, become simply the resources that scientists or researchers use – as in the case of databanks or online libraries. This distinction is still fluid in the case of the three technologies discussed here: once they settle into use, will they be used primarily as tools or as resources?

It is therefore difficult to say how the tools in these three projects, and e-science in general, will be sustained and thus what lasting effects these research technologies will have. This relates to a general feature of technology-in-the-making (or science-in-the-making) – which is that once it becomes technology-already-made (or science-already-made), it becomes difficult to distinguish as a new tool. As Shinn & Joerges point out, 'Research technologists sometimes even opt for a measure of community "invisibility"'(2002: 215) – they deliberately stay outside established research communities so as to avoid the barriers that may exist within them. This fits well with e-science, whose practitioners often stress that a separate e-science will 'disappear' once the practices of e-science have become so well integrated with domain sciences that they are no longer visible. At the same time, as with all science- and technology-in-the-making, while they are still being developed it is difficult to assess their distinctive and novel implications, and where these begin and how they will end. Research technologies thus combine disciplinary change and change towards greater worldwide scope, but they have only begun to provide generic devices and a common language. Shinn & Joerges (2002), too, acknowledge that it is difficult to identify where the process of disembedding and re-embedding starts and finishes.

Still, e-science may itself make a contribution to a better understanding of the migration of and shifts in the disciplinary landscape because, in the case of e-science, the ‘openness’ of research is to some extent online. The research thereby gains greater visibility than offline scientific modes of communication in the sense that researchers increasingly find information and resources online. This development can be tied to recent innovations in scientometrics, which make it possible to map knowledge by means of electronic tools. Thus scientific ‘output’, its volume, scope, patterns of funding and degree to which it is made use of or accessed – which, put together, may give an indication of the ‘migration’ of disciplines – can increasingly be quantified and analysed more effectively because the information is in electronic form (e.g. Shiffrin & Boerner, 2004). And within e-science, as we have already seen, it is also possible to measure inputs; for example, the shared computing resources used, by number of users, total number of hours of central processing unit usage, disciplines, institutions, files stored and the like.[26] It is no doubt possible to do this for all Grid services and to aggregate these into a global whole. From a disciplinary perspective, what stands out, for example, is that physics has the greatest need for high-performance computing.

The technological – electronic – mediation of scientific advance is thus becoming an ever-more globally visible part of scientific knowledge as a whole with an increasing shift towards the use of shared online resources and the tools to access them. This will move the focus of attention and resources towards advances that have an e-science presence or component. e-Science thus transforms the landscape of scientific research, even if this process is constrained by the competition between different e-science efforts and the prioritization of resources among different scientific areas. And if it is possible to recognize the online visibility of e-science technologies more clearly than for offline research, it may also be possible to better direct research organization and research policy. At the same time, whatever attempts are made at steering e-science, these research systems, like other large technological systems, have developed a momentum of their own.

Conclusion

As Shinn & Joerges have noted in their research, the practice-based universality of the e-research technologies that make knowledge more manipulable is bound to continue to grow:

> In decades to come, research-technology may emerge as a still more influential and assertive component in artefact and knowledge production. As the world of learning and artefacts becomes more encumbered, complex, and differentiated, transverse mechanisms [of re-embedding and disembedding (my addition)] capable of inducing intellectual order and intelligibility and of assuring social coherence become increasingly essential. (Shinn & Joerges, 2002: 245–6)

Here is a link to the large technological system of the Internet and Web and the system of research, whose increasing size, complexity and differentiation are features (also noted by Shinn & Joerges on p.243) simultaneously of the growing online system of scientific communication (in a wide sense) and of the networks (or in this case 'Grids') which enable it.

The mistake in regarding e-science (or science in general) as inescapably tied to particular (less-than-global) social contexts or as persisting within disciplinary boundaries is to think of globalization, multidisciplinarity and the universality of science achieved via research technologies as all or nothing. Instead, the contributions of e-science consist of many incremental steps by means of which the global nature of this part of scientific advance and new forms of disciplinarity emerge in different ways: in part, for example, it consists of a physical network that spans the globe; other parts consist of the resources (the content) that represent the most up-to-date available data within a particular area of scientific knowledge, and still other parts are instantiated in organizational collaboration and forms of scientific communication, whether by means of sharing computer-processing power or repositories of data in large technological systems.

Scientific knowledge cannot be global and does not create new hybrid disciplines once and for all inasmuch as its leading edge is always provisional; but this is simply part of the open and restlessly advancing nature of science- and technology- in-the-making as a whole. Thus the flipsides to all these examples, as we have seen, are the constraints on current e-science. There are further limits: do all e-science efforts contribute to globalization and reconfiguring disciplines? No, because many parts of e-science efforts will turn out not to contribute to the advance of interlocking between computing tools and the physical and natural worlds. And some e-science projects and initiatives are bound to reinforce disciplinary and geographical boundaries, and also fail to produce tools which have the characteristics of research technologies that have been highlighted here. Finally, some, especially in the light of the limited attention space, will fail to become standard tools and thus wither, decline or become absorbed in other projects or efforts. Nevertheless, there is a cumulative effect as a result of these changes, including a shift towards shared and distributed computing tools for research and a simultaneous shift towards more global and interdisciplinary forms of scientific knowledge.

e-Science will not completely transform existing disciplinary foci or globalize knowledge production at once; nor will it reconfigure research only piecemeal and for individual projects. Rather, this reconfiguring is systematic in the sense that there are common features or crystallizations among e-science efforts. Research in general is led by the development of research technologies, and manipulating knowledge by means of shared and distributed computing resources will continue to demonstrate its effectiveness and universality in practice.

Ralph Schroeder is James Martin Research Fellow at the Oxford Internet Institute, Oxford University. He was formerly a professor in the School of Technology Management and Economics at Chalmers University in Gothenburg (Sweden). Among his book-length publications are: *Possible worlds: the social dynamic of virtual reality technology* (Boulder, CO: Westview, 1996) and *Rethinking science, technology and social change* (Stanford: Stanford University Press, 2007). Recent articles include: (2002) R.Schroeder, 'The consumption of technology in everyday life: car, telephone, television in Sweden and America in a comparative-historical perspective', *Sociological research online* 7(4); R.Schroeder & R.Swedberg, 'Weberian perspectives on science, technology and the economy', *The British journal of sociology* 53: 383–401; (2003) A.-S. Axelsson, A.Abelin & R.Schroeder, 'Anyone speak Spanish? Language encounters in multi-user virtual environments and the influence of technology', *New media and society* 5: 475–98; (2007) R.Schroeder & J.Fry, 'Social science approaches to e-science: Framing an agenda', *Journal of computer-mediated communication* 12(2). He is currently working on shared virtual environments and on the social implications of e-science. *Author's address*: Ralph Schroeder, James Martin Research Fellow, Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford OX1 3JS, UK. [email: ralph.schroeder@oii.ox.ac.uk]

Notes

1. I would like to thank Rob Ackland, Owen Appleton and Reagan Moore for very helpful suggestions on parts of this paper. The work for this article has been supported by ESRC grant RES-149–25–1022 and is part of the Oxford e-Social Science (OeSS) project.

2. This is narrower than the term collaboratories, which includes any form of distributed collaboration, see Finholt (2003). e-Science, in contrast, means either sharing computer-processing power via networks or sharing the tools to manipulate digital resources.

3. In the social sciences there have been a number of debates about globalization (Guillen, 2001). In relation to knowledge and research, these have often focused on concepts such as 'information', 'network' or 'knowledge society'. More recently a number of research issues have come to be regarded as global challenges to scientific research, such as climate change, infectious diseases and poverty. For the globalization of scientific institutions, see Drori et al. (2003).

4. There is a prima facie case to be made for the global nature of e-science: The technology that supports e-science is based on a worldwide system, the Internet/Web, which, like other communications infrastructures, spans the globe. Within this system, the factors that contribute to its global nature include the worldwide searchability of online sources and the emergence of supra-national standards for accessing and storing online material (such as those being developed for the World Wide Web consortium, http://www.w3.org/). And within e-science, there are a growing number of worldwide organizations devoted to promoting this kind of coordination and standardization, including conferences and non-governmental bodies (such as the Open Grid Forum, http://www.ogf.org/). Nevertheless these general features of e-science do not obviate the need to instantiate the global features of specific e-science research projects and their practices.

5. These are also known as National Research and Education Networks, or NRENS.

6. Apart from their uses in research, scientific instruments also yield benefits for innovation and economic growth (see Rosenberg, 1992). This is discussed further in the conclusion.

7. As we saw in the previous section, e-science has been driven by a variety of factors, including national policy initiatives (see Schroeder & Fry, 2007) and the needs of different scientific domains to cope with the 'data deluge' in the sciences (Hey & Trefethen, 2003) or for specific tools (such as visualization). It is thus also driven by the development of infrastructures (see Schroeder, 2007a). The main thrust, however, is tool development, though we shall come back to how this fits into larger developments, and the limitations of applying the concept of research technologies to e-science.

8. Sometimes the term 'Grid' is used instead of e-science technologies, but 'Grid' is still used in different ways. Here 'Grid' can be used as shorthand for shared computing resources across high-speed networks, which overlaps with the definition of e-science given earlier.

9. Middleware is the software that sits between the Grid or the computing infrastructure and individual applications or projects or institutions, and allows the resources in one to share or access the other. In other words, middleware translates between the shared resources on the network and the individual application.

10. One indication of the rapid increase of interest in the area of the Web and its networks is the funding by the Office of the Cyberinfrastructure of the National Science Foundation of a $2 million project at Cornell University (http://www.nsf.gov/news/news_summ.jsp?cntn_id=104477). VOSON shares some of the aims of the Cornell project and has had some collaboration with it.

11. Another possibility is that environmental activists have only an online presence; they are individuals who have common organization only via their online links and are focused on pushing a particular agenda online. But this merely pushes the problem one step further; namely, what is the importance of these relations in the websphere for the offline world generally

12. One example of how the Grid is moving towards being more usable by the research community outside of computer science – or user-friendly generally – is by developing 'Web services', means of accessing the Grid or accessing shared data and other resources via standard Web interfaces that researchers are familiar with.

13. EGEE builds on several previous European efforts, including GEANT, the European academic computing network, the European DataGrid project and the LHC Computing Grid (LCG) (see http://lcg.web.cern.ch/LCG/), which shared computing resources before 2004. The organization of large-scale collaborative research organization in physics is analysed in Shrum, Genuth & Chompalov (2007).

14. Twenty-six projects had entered into contractual agreements with EGEE as of May 2007, though there were more than 100 partners overall associated with EGEE.

15. In fact the picture is perhaps even more complex inasmuch as processing and storage are involved both in physics and in biomedicine, but to different degrees. A third way of using Grids is as tools for distributed collaboration, and this use applies much more widely to all the disciplines involved in EGEE.

16. It is interesting that the precise contribution of each partner in terms of computing power (e.g. CPU deployment, Gagliardi, 2005: 200) can be specified and measured.

17. One indication of the collaborative nature of EGEE is that it promotes open-source software and open-access standards.

18. EGEE thus also builds on a number of existing national and regional Grids, such as the UK e-science Grid (for a discussion of GridPP, the UK particle physics e-science Grid, see Zheng, Venters & Cornford, 2007), NorduGrid for the Nordic countries, and the like.

19. There is a parallel here with the recent agreement between IBM and Microsoft on software standards for e-science: after years of trying to become the dominant standard, with bases of support in the US and Europe, the two sides decided that more was to be gained from collaboration.

20. The guidelines can be found at http://www.ivoa.net/Documents/latest/IVOA Participation.html. Note that open access to data is subject to being 'commensurate with national or project-imposed proprietary periods'.

21. This puts astronomy into the category of sciences with 'low uncertainty' (Whitley, 2000) or 'batch production' (Fuchs, 1992: 112–13).

22. According to Nic Walton, who is one of the UK IVOA members, there are 20,000 worldwide users (figure given at a presentation at the UK e-science All Hands Meeting, 18–21 September 2006).

23. One feature that makes the integration of IVOA data easier than other types of data is that it has no commercial value.

24. In this, IVOA and its members will need to coordinate not only among themselves but also with other parallel or partly related efforts, such as the earlier Sloan Digital Sky Survey (see http://www.sdss.org/).

25. Comments made by Steve Newhouse and Jennifer Schopf at the UK e-science All Hands Meeting, 18–21 September 2006.

26. For example, for the UK National Grid service (http://www.nesc.ac.uk/talks/mpa/Grid 2006BarcelonaKeynote20060919MalcolmAtkinsonV2.pdf, slide 41), a keynote talk by Malcolm Atkinson, 'e-science: Foundations for the European Citizen', GRID 2006, Barcelona, Spain (see also Gentzsch, 2006).

References

Ackland, R., O'Neil, M., Bimber B., Gibson, R. & Ward, S. (2006a) 'New methods for studying online environmental-activist networks', paper presented to 26th International Sunbelt social network conference, 24–30 April, Vancouver.

Ackland, R., O'Neil, M., Standish, R. & Buchhorn, M. (2006b) 'VOSON: A Web services approach for facilitating research into online networks', paper presented at the Second international conference on e-social science, 28–30 June 2006, University of Manchester.

Appleton, O. (2006) EGEE II 'Overview paper' (available at: http://www.dcc.ac.uk/events/policy-2006/EGEE-II%20overview%20paper.pdf).

Becher, T. & Trowler, P. (2001) *Academic tribes and territories: intellectual inquiry and the culture of disciplines*, 2nd edn. Milton Keynes: Open University Press.

Choudhury, S., Dilauro, T., Szalay, A., Steffen, J., Milkey, R., Ehling, T., Plante, R. & Hanish, R. (2006) 'Digital preservation in astronomy', in 'Proceedings of the second international digital curation conference', Glasgow 21–22 November. (Powerpoint presentation available at http://www.dcc.ac.uk/events/dcc-2006/programme/)

Collins, R. (1993) 'Ethical controversies of science and society: a relation between two spheres of social conflict', in T. Brante, S. Fuller & W. Lynch (eds) *Controversial science: From content to contention*, pp. 301–17. Albany: State University of New York Press.

Collins, R. (1994) 'Why the social sciences won't become high-consensus, rapid-discovery science', *Sociological forum* 9(2): 155–77.

Collins, R. (1998) *The sociology of philosophies: a global theory of intellectual change*. Cambridge, MA: Harvard University Press.

Cummings, J. & Kiesler, S. (2005) 'Collaborative research across disciplinary and institutional boundaries', *Social studies of science* 35(5): 703–22.

Drori, G., Meyer, J., Ramirez, F. & Schofer, E. (2003) *Science in the modern world polity: institutionalization and globalization*. Stanford, CA: Stanford University Press.

Edquist, C. (1997) 'Systems of innovation approaches – their emergence and characteristics', in C.Edquist (ed.) *Systems of innovation: technologies, institutions and organizations*, pp. 1–35. London & Washington, DC: Pinter.

Finholt, T.A. (2003) 'Collaboratories as a new form of scientific organization', *Economics of innovation and new technology* 12: 5–25.

Fuchs, S. (1992) *The professional quest for truth: a social theory of science and knowledge*. Albany: State University of New York Press.

Fuchs, S. (2001) *Against essentialism: a theory of culture and society*. Cambridge, MA: Harvard University Press.

Fuchs, S. (2002) 'What makes sciences scientific?', in J. Turner (ed.) *Handbook of sociological theory*, pp. 21–35. New York: Kluwer Academic/Plenum Publishers.

Gagliardi, F. (2005) 'The EGEE European grid infrastructure project', in M. Dayde, J.J. Dongarra, V. Hernandez & J.M.L.M. Palma (eds) *VEPCAR 2004, LNCS 3402*, pp. 194–203. Berlin: Springer.

Gagliardi, F., Jones, B., Grey, F., Begin, M.-E. & Heikkurinen, M. (2005) 'Building an infrastructure for scientific grid computing: status and goals of the EGEE project', *Philosophical transactions of the Royal Society A* 363: 1729–42.

Galison, P. (1997) *Image and logic: a material culture of microphysics*. Chicago: University of Chicago Press.

Galison, P. & Hevly, B., eds (1992) *Big science: the growth of large-scale research*. Stanford, CA: Stanford University Press.

Gentzsch, W. (2006) 'National Grid initiatives: lessons learned and recommendations', Keynote speech at the Second international conference on e-science and Grid computing, Amsterdam, 4–6 December 2006.

Guillen, M. (2001) 'Is globalization civilizing, destructive or feeble? A critique of five key debates in the social science literature', *Annual review of sociology* 27: 235–60.

Hacking, I. (1983) *Representing and intervening*. Cambridge: Cambridge University Press.

Hanish, R. & Quinn, P. (n.d.) 'The international virtual observatory'. (available at: http://www.ivoa.net/pub/info/; last accessed 2 June 2007)

Hey, A. & Trefethen, A. (2003) 'The data deluge: an e-science perspective', in F. Berman, G.C. Fox & A. Hey (eds) *Grid computing: making the global infrastructure a reality*, pp. 809–24. Chichester: John Wiley.

Hey, T. & Trefethen, A. (2006) 'Collaborative physics, e-science, and the Grid – realizing Licklider's dream', in G. Fraser (ed.) *The new physics for the twenty-first century*, pp. 370–402. Cambridge: Cambridge University Press.

Hine, C. (2006a) 'Computerization movements and scientific disciplines: the reflexive potential of new technologies', in C. Hine (ed.) *New infrastructures for knowledge production: understanding e-science*, pp. 26–47. Hershey, PA: Information Science Publishing.

Hine, C., ed. (2006b) *New infrastructures for knowledge production: understanding e-science*. Hershey, PA: Information Science Publishing.

Hughes, T. (1987) 'The evolution of large technological systems', in W. Bijker, T. Hughes & T. Pinch (eds) *The social construction of technological systems*, pp. 51–82. Cambridge, MA: MIT Press.

Hughes, T. (1994) 'Technological momentum', in L. Marx & M.R. Smith (eds) *Does technology drive history? The dilemma of technological determinism*, pp. 101–13. Cambridge, MA: MIT Press.

Lemson, G. & Colberg, J. (2004) 'Theory in the VO'. (available at: http://www.ivoa.net/twiki/bin/view/IVOA/IvoaTheory#Papers; last accessed 2 June 2007)

Nentwich, M. (2003) *Cyberscience: research in the age of the Internet*. Vienna: Austrian Academy of Sciences Press.

Rosenberg, N. (1992) 'Scientific instrumentation and university research', in N. Rosenberg (ed.) *Exploring the black box: technology, economics, and history*, pp. 250–63. Cambridge: Cambridge University Press.

Schroeder, R. (2007a) 'e-Research infrastructures and open science: towards a new system of knowledge production?', *Prometheus* 25(1): 1–17

Schroeder, R. (2007b) *Rethinking science, technology and social change*. Stanford, CA: Stanford University Press.

Schroeder, R. & Fry, J. 2007 'Social science perspectives on e-sciences', *Journal of computer-mediated communication* 12(2). (available at: http://jcmc.indiana.edu/vol12/issue2/schroeder.html)

Shiffrin, R. & Boerner, K. (2004) 'Introduction: mapping knowledge domains', *Proceedings of the National Academy of Sciences of the United States of America* 101(Supplement 1): 5183–5.

Shinn, T. (2005) 'New sources of radical innovation: research technologies, transversality and distributed learning in a post-industrial order', *Social science information sur les sciences sociales* 44(4): 731–64.

Shinn, T. & Joerges, B. (2002) 'The transverse science and technology culture: dynamics and roles of research technology', *Social science information sur les sciences sociales* 41(2): 207–51

Shrum, W., Genuth, J. & Chompalov, I. (2007) *The structures of scientific collaboration*. Cambridge, MA: MIT Press.

Slaughter, A.-M. (2004) *A new world order*. Princeton, NJ: Princeton University Press.

Taylor, J. (2001) 'News from the e-Science Programme, first phase'. (available at: [http://www.rcuk.ac.uk/escience/news/firstphase.htm; last accessed 2 June 2007)

Thelwall, M. (2006) 'Interpreting social science link analysis research: a theoretical framework', *Journal of the American Society for information science and technology* 57(1): 60–8.

Whitley, R. (2000) *The intellectual and social organization of the sciences*, 2nd edn. Oxford: Oxford University Press.

Zheng, Y., Venters, W. & Cornford, T. (2007) 'Distributed development and scaled agility: improvising a Grid for particle physics', London School of Economics working paper Series 163. (available at: http://is2.lse.ac.uk/wp/pdf/wp163.pdf)