Preface

Widespread international attention has recently been given to development of technologies to facilitate new ways of doing science. This book contributes to this burgeoning interest by using sociological methods and theories to explore the use of computers in scientific research. Specifically we analyze the increasingly prominent uses of information and communication technology (ICT) infrastructures for storing scientific data, performing analyzes and carrying out collaborative work, often known in the U.S. as cyberinfrastructures and in the UK as e-science. Led by data-intensive fields such as particle physics, astronomy, and genetics, new infrastructures are being designed that promise to allow scientific research to be conducted on a larger scale and with greater efficiency than previously conceivable, and to explore ever more complex questions. It remains to be seen how far this model will generalize, to what extent the models currently envisaged will translate to other disciplines and what the impacts may be for more traditional approaches. This collection looks at these innovations in the organization of scientific research, focusing on the factors which shape their inception, promote their uptake, lead to variability in application and result in the recognition of significant impacts. Our topic is, in short, the social dimensions of ICT-enabled science.

Social studies of science have established a set of approaches to the study of the scientific endeavour which explores processes of knowledge production as they emerge through particular social, spatial and material arrangements. Science and technology studies (STS) is an established field offering commentaries on scientific and technical developments, spanning the full range from top-level critique to detailed analysis and making contributions both to understanding the process of designing and to design itself. There has at times been an

uneasy relationship between STS and the practitioners of science. The attitude of STS has sometimes been construed as anti-science, or as neglecting the concerns of scientists altogether in its focus on contesting philosophical accounts of scientific enquiry. This clearly does not exhaust the possibilities for the engagement of STS with science. A more hopeful scenario could be drawn from technology design, where approaches from STS have found popular acceptance. In technical spheres the relationship has often been less confrontational: It has become common for companies developing mass market technologies to employ ethnographers, using broadly STS-informed approaches, to inform their design activities. E-sciences, and more broadly the design of new infrastructures for the conduct of science, provide the occasion to extend this approach to the design of new technologies for science, allowing for STS to develop a much more constructive engagement with scientific practice. This is an unrivalled opportunity to demonstrate the relevance of social studies of science for the conduct of science itself.

An STS-informed understanding needs to be high on the agenda for those funding and designing new infrastructures for science, in order that technical capabilities be complemented with an in-depth understanding of social processes and consequences and with a theoretical tool kit to comprehend diversity. Innovation in the organization of scientific work should benefit from an enhanced understanding of social process. Woolgar and Coopmans in the first chapter of this book lay out a scope for the possible contributions that STS could make, but they note that the engagement of STS with e-science issues is in its infancy. That observation captures the motivation in putting this book together—to collect together and make more evident the ongoing work in the field, consolidating its contribution and making it more visible both to e-science practitioners and to the science studies community, beginning the work of forging relationships between STS analysis and e-science practice.

In the process of putting together this book it has proved helpful to adopt a fairly broad and flexible notion of appropriate technologies and situations to include. The goal has been, in part, to use the STS tradition of scepticism in order to question the way in which the e-science phenomenon is constituted. Part of our job has been to pursue connections which might otherwise have remained tacit between the current constitution of the phenomenon and the prefiguring technologies and policies, and to imagine the ways in which it might have been otherwise. A more practical reason for extending the scope of the technologies that we consider is in order to acknowledge the heritage of work on distributed, collaborative work and the foundations of scholarship on disciplinary practice and disciplinary difference. For e-science practitioners, it seemed that one of the most useful things we could do would be to bring this prior work into the domain of e-science and show how it can inform current dilemmas.

In the process there have inevitably been some omissions. Much of the existing work that we have been able to include has focused on communication and the

preconditions for data sharing and other forms of distributed work. Much less focuses on computational aspects of the work of scientists, asking about changes in the knowledge production process and the ways in which findings are constituted, recognized and made communicable in e-science projects. Part of this is because e-science too is in its infancy, and thus examples to study for the whole of the knowledge production cycle are as yet rare. There are some promising signs that STS will be grappling with the import of computational e-science in emerging work on simulations and modeling (Lenhard et al., forthcoming 2006). Studying how knowledge is constituted has been the core of the sociology of scientific knowledge, and we can expect that this aspect of the e-science endeavour will be a focus in the future—all the more reason then to make sure that we are in on the ground now, exploring how the infrastructures that will support those knowledge production processes are being shaped.

I began this introduction by talking about e-science as an opportunity for STS to engage with science practice and policy on a constructive level. The opportunity is, however, not confined to an altruistic mode. There is a considerable pay-off for STS. Several of the chapters show that engagement with e-science is an opportunity to examine, refine and question accepted notions in STS. New technologies often provide reflexive opportunities, as much because they are perceived as new as for any specific transformative properties they offer. The reflexive opportunity that e-science offers to STS is the chance to explore knowledge production processes in the making, finding in their apparent novelty an occasion to reinvigorate established frameworks and highlight previously taken-for-granted assumptions about the ways that knowledge production works. The balance shifts between chapters, some focusing more on practical outcomes or policy critique for e-science, and some containing a stronger element of developing and reflecting on STS concepts.

The contribution to STS is methodological as well as theoretical. While social scientists have endeavoured to study scientific activity wherever it is carried out, the classic location for study is the laboratory. In the implementation of novel and spatially distributed ways of doing science we have an opportunity to see new locations of science in the making, and to ask new questions about the ways in which knowledge production is organized as a socio-technical process. We will learn new things about the processes by which scientific cultures change, about the importance of location, disciplinarity, and collaboration in science and about the development of new technologies in knowledge production contexts. E-science provides the opportunity to build on tried and tested methodologies of ethnography, of technology assessment and of scientometric and network analysis, and make them afresh for the situation in which they find themselves.

The first section of the book focuses on organized efforts to promote the development of new infrastructures, asking what motivates these efforts, what assumptions shape and constrain them and what structures they put in place. The impetus behind these examinations is to find modes of analysis that avoid falling

into existing assumptions whilst maintaining a constructive dialogue. In the process, the chapters in this section highlight connections and limitations that might otherwise remain hidden between old and new technologies for science, and imagine alternative infrastructures that might serve the goals of different communities.

In the first chapter, Woolgar and Coopmans lay out a broad framework for the contribution that STS could make to the understanding of e-science. They argue that while STS approaches to e-science are in their infancy, there is a broad range of contributions to be made to understanding the genesis of these technologies, understanding the social and economic aspects of design, uptake and use and exploring the implications for the practice and outcomes of knowledge production. They argue that STS offers the potential for an in-depth examination of the ways in which notions of data, networks, and accountability develop in e-science contexts. For example, this makes it possible not just to follow the mobility of data but to ask how it is that data are recognized and rendered mobile and for whom. In a final section, Woolgar and Coopmans illustrate their argument by exploring notions of witnessing, with the aim of finding out how far e-science is replaying earlier debates about the adequacy of experimental reports. Notably, their argument suggests that we can use the juxtaposition of old and new versions of scientific witnessing to reflect on both forms, and, indeed, to question the distinction between old and new. E-science proves to be an opportunity to reflect on some previously taken-for-granted aspects of scientific communicative practice. Woolgar and Coopmans suggest that STS can both contribute to e-science and benefit from the engagement.

In the second chapter, Hine examines one discipline's use of information and communication technologies, exploring the dynamics which have produced a "computerization movement" within the discipline. The computerization movement framework highlights the cultural connections across diverse spheres that have been occasioned by computers, with the technology seen as a tool for beneficial social transformation. Hine argues that traces of such a movement can be discerned in the recent experience of systematics. This discipline has found itself with a high political profile thanks to its foundational role in biodiversity conservation efforts. This has produced a recent emphasis on using, and being seen to use, distributed databases to make data widely accessible. Grand visions for the transformation of the discipline have been promoted. The publicity accorded to grand visions allows systematists the opportunity to stress the progress that they are making, disaggregating and redefining the terms being used and taking care to point out that activities in this sphere need additional funding in order to succeed. Hine argues that the computerization movement in systematics, whilst not necessarily meeting with wholesale approval from those within the discipline, has nonetheless provided an opportunity for wide-ranging reflection and debate on its goals and practices. The upshot of this analysis for our understanding of e-science more broadly is that we

should not be too daunted if grand visions fail to be realized. A payoff in terms of reflection, debate and learning across a spectrum of issues concerning technical infrastructure, practices and goals may occur quite independently of the realization of grander transformative schemes.

The third chapter turns to e-science initiatives to argue that, regardless of the broad gains to be expected from the stimulus they provide to develop technologies and explore visions, we should not adopt a laissez faire attitude towards their development nor assume that models will automatically diffuse across disciplines. Wouters and Beaulieu use the observation that science consists of diverse epistemic cultures (Knorr-Cetina, 1999) to explore how the locations in which e-science concepts are developed could be consequential for their subsequent generalizability. Indeed, Wouters and Beaulieu suggest that generalizable e-science might be an impossible dream, if we expect it to deliver infrastructures that travel across disciplines. Having established that much of the current focus of e-science initiatives is on computational work, they examine the epistemic culture of women's studies to suggest that quite different notions of data, analysis and infrastructure might prevail there. The idea of disciplinarity, and its consequences for the generalizability and deployment of new infrastructures, arises several times throughout this book. Wouters and Beaulieu provide the first of several discussions of disciplinary specificity, to be followed by Haythornthwaite et al., Merz and Fry in the second section and Nentwich and Barjak in the third.

The final chapter in this section moves to a different frame of analysis, to look at the ways in which e-science initiatives structure the labor relations between disciplines. Vann and Bowker introduce the term "epistemic IT," standing for the information technologies developed for use by scientists in knowledge production. Deliberately, this term includes both e-science and its predecessor attempts to promote the transformation of knowledge production through the development of new information technologies, since the chapter argues that preceding visions may have an unrecognized effect on the shaping of future possibilities. Vann and Bowker explore the predecessors of e-science to remind us that these visions are not new, and that they have the ability to shape goals, actions and expectations. In particular, they trace the emergence of a focus on interdisciplinary collaboration, and explore the ways in which this brought into being requirements for particular kinds of labor. Vann and Bowker show how grand visions of interdisciplinarity ask for the investment of labor, from domain scientists and computational experts, and describe one initiative where the commitment of scientists to make this investment was secured by some inventive funding arrangements. Vann and Bowker bring this section of the book to a close by exploring the "will to produce" that grand future visions promote, and the creative social and technical arrangements that may be required to enable that will to be pursued. In the next section the focus remains on the level of the experience of deploying new technical infrastructures, asking how they become embedded in and appropriate to their specific circumstances of use.

The second section of the book, on communication, disciplinarity and collaborative practice, asks what it takes for scientists to use new infrastructures for knowledge production. The focus shifts to the scientists using mediated communication for their collaborations and data exchanges, exploring the new forms of expertise that it demands and the extent to which certain disciplinary practices predispose their practitioners to use technologies in particular ways. These chapters suggest that use of new infrastructures can be a highly skilled form of work, and yet these skills may only be recognized when they break down, either in interdisciplinary initiatives or where the practical preconditions for their deployment are absent. This section focuses on the question of the specificity of new infrastructures for knowledge production as introduced in the first section by Wouters and Beaulieu, exploring in detail the dimensions of specificity and the constraints on extension of technologies beyond their contexts of production.

Merz begins the elaboration of the specificity of knowledge production infrastructures with a chapter that describes the embedding of electronic communications in disciplinary practice. She argues that the epistemic cultures of science can differ quite markedly in ways that are relevant for their adoption of electronic communications, and that the use of these technologies can develop in quite different ways in different communities. We should not, therefore, expect e-science to be a unified phenomenon. The argument is made by examination of a case study focusing on theoretical particle physics. Merz looks at the collaborative practices and preprinting conventions of this community, showing that use of electronic technologies is rooted in very specific ideas about the importance of collaboration and the nature of collaborative work. The relative freedom from physical location allows theoretical particle physics to be a particularly mobile community. Distributed collaboration making intensive use of e-mail is common, but does not render face-to-face communication redundant. Preprinting has a history which predates the availability of internet-based repositories, and is used by theoretical particle physicists in their everyday work in ways quite particular to their culture. Use of e-mail and of preprint archives is portrayed by Merz as highly cultural specific and deeply embedded in the culture of the discipline. This brings into question the extent to which models from this discipline might diffuse into other fields with quite different practices and expectations.

Merz shows that the communication technologies that a discipline uses are thoroughly embedded in practice. In the next chapter, Elvebakk explores another aspect of embedding, looking at the preconditions for data to be shared electronically. Through an exploration of the data-sharing practices of chemists, she shows how data have first to be rendered exchangeable in the eyes of their producers and users, by establishing the equivalence of representations with the material objects that precede them, and by making appropriate arrangements for the portrayal of contextual or otherwise tacit aspects of work

with objects and data. This case study is situated within the existing literature on the relation of scientists with their objects of study, focusing on the processes through which objects are successively reconstituted through chains of representational practices. Elvebakk argues that current digital technologies are not radically different, instead continuing an existing trend diminishing the prominence of the material object.

If, as these two chapters have established, use of digital technologies takes place within distinctive disciplinary cultures, then it follows that interdisciplinary work can be particularly problematic. As Vann and Bowker discussed in the first section, interdisciplinary approaches have played an important part in the visions for new scientific infrastructures. In this section, Haythornthwaite, Lunsford, Bowker and Bruce discuss in depth the experience of distributed interdisciplinary work. They show that issues such teams face include the need to attend to work scheduling, learning practices, the nature of relations within and beyond the team and their use of technologies. While working practices in everyday discipline-based work often remain on a tacit level, working in a distributed interdisciplinary team can require that they be articulated and made explicit, so that all of those involved can share expectations and take care to head off likely pitfalls. New collaborative skills need to be developed to make distributed interdisciplinary work happen.

The final chapter in this section focuses again on disciplinary diversity, this time establishing a framework for making systematic comparisons between disciplines. Fry uses existing taxonomies of scientific disciplines to compare data on the experience of scientists from three disciplines with digital technologies. Interviews were conducted within high-energy physics, social/cultural geography, and corpus-based linguistics. Fry found that Whitley's (1984) framework accounting for organizational differences between scientific fields proved fruitful for understanding the different responses of these disciplines to the use of new technologies for collaboration and data sharing. Specifically, the framework highlights the importance of contrasting levels of interdependence between scientists and of varying levels of uncertainty around research problems, the objects of research, the techniques to be used and the evaluation of outcomes. Fry argues that we can use this framework to understand why some disciplines are more ready than others to adopt these new technologies, and to look at the way that the existing mechanisms of coordination and control of research practice may be affected by new infrastructures.

The third section of the book then moves to examination of the structures of science on a broader level, asking how new infrastructures might occasion wholesale change in the ways that science is communicated and in the inequalities between groups of scientists. The potential for change in the science system has, after all, been a feature of discussions about e-science and cyberinfrastructure, and it seems important to consider what signs are discernible at this early stage. Chapters in this section examine various aspects of the

science system and of the structures and inequalities within it, to consider how far the promise of change is being realized.

In the first chapter of this section Nentwich examines the prospects for change in the scholarly publishing system. His approach is rooted in the tradition of Technology Assessment, a practically-oriented branch of STS that focuses on analyzing the potential consequences of emerging technologies, aiming to provide an informed basis for policy decision-making. Nentwich explores the evolution of new infrastructures for scholarly publishing and their differential diffusion across disciplines. He finds that there is likely to be considerable variation between disciplines, assessing potential change in terms of the endurance of print publishing in the face of electronic alternatives, the mechanisms of quality control, the economic and legal aspects of scholarly publishing and the varying collaborative structures of disciplines. He suggests that print systems may coexist with new publishing infrastructures in the short term, but that print is ultimately unlikely to persist for most texts. There is a need, he argues, for a more organized constructive technology assessment to lead developments in desirable directions, involving representatives from the library community, from publishers, universities and, of course, researchers themselves.

In the second chapter in this section, Caldas assesses the emergent structures of science on the World Wide Web, analyzing the extent to which online structures mirror offline patterns of centrality and marginality. He draws on a previous study of the networks of communication and collaboration between European researchers of speech and language, and develops an analysis of the visible web presence of this community and the interlinkages between institutions. In scientometrics, STS has a strong heritage of exploring structures and evaluating differentiation within the science system using visible indicators such as publication and citation records. The World Wide Web provides the opportunity to develop webmetrics which deploy hyperlinks to evaluate emerging structures. Caldas finds that he can identify patterns within the web that map onto the offline structures of centrality and prestige. He also suggests that it may be possible to use the webmetric approach to identify "digital knowledge bases," emergent intensive knowledge zones that are particularly important and respected within a field. By focusing on the World Wide Web, Caldas is able to show us an emergent information infrastructure, not designed through any specific initiative but moulded out of the practices of individual institutions and researchers and their evaluations of one another. On this basis, we find that the new electronic infrastructures may mirror their more traditional counterparts.

Barjak, in the third chapter in this section, also evaluates emergent structures within the science system, and finds that there is a considerable continuity between old and new. His report is based upon a wide-ranging survey of internet use amongst European scientists, upon which multivariate analyses have been performed to explore differences between countries, disciplines, young scientists and more established researchers and between male and female. It has

often been suggested that electronic technologies will have a leveling effect, enabling previously marginalized groups to participate more fully in society. In respect of science, it has been hoped that electronic technologies will enable fuller participation for less well-funded or less-experienced researchers and a broader accessibility of scientific information. Barjak finds, however, that levels of internet use are consistently reported as greater for male researchers and those at a more senior level. Rather than overcoming previous inequalities, Barjak suggests that the internet is tending to reproduce them, in a "hybrid divide" that encompasses both analogue and digital information.

In the final chapter we move to a focus on two aspects of the science system where inequalities have been starkly apparent, with an assessment of the implications of the internet for women scientists in developing countries. Palackal, Anderson, Miller and Shrum describe an interview-based study with scientists in Kerala, India, assessing the extent to which the internet is being used by women scientists and its impact on their experience and opportunities. Women are found to be generally highly restricted in their scientific careers in this patrifocal society, which both limits their ability to travel internationally and restricts their ability to network with male scientists locally. The internet, it is suggested, is allowing women to circumvent some of these restrictions, enabling them to connect with scientists internationally and to form research relationships which are less restricted by Kerala norms of gendered interaction. The internet is not as yet overcoming the restrictions which these scientists face, but is allowing them to circumvent some of their limitations and is a part of a growing awareness of the importance of international links among Kerala women scientists. This chapter demonstrates that the open networking possibilities provided by the internet can have a positive effect on the experience of previously marginalized researchers. When developing more sophisticated infrastructures in the future it may be important to try to preserve some of these advantages of their predecessors.

In summary, this book provides a highly differentiated account of the new infrastructures for knowledge production. Rather than having wholesale effects on the science system, the consequences are likely to be slow to emerge, and to be experienced very differently in different disciplines. The resulting structures are likely to have much in common with the existing science system, although some aspects of the new infrastructures may provide for new forms of participation amongst previously marginalized groups. How these changes evolve will depend very much on how debate and technological development are organized, and which groups are able to influence the agenda. STS perspectives suggest that it matters whose voice is heard when new technologies are being designed. They also suggest that users have a key role to play in deciding whether and how to adopt new technologies and in developing working practices which make sense of them. In the science community we have a highly empowered and reflexive group of people, who will ultimately be able to shape these tech-

nologies to suit their purposes. Hopefully, the analyses presented in this book will inform ongoing reflection and debate amongst policy makers, technology developers, scientists and STS researchers, enabling an enriched effort at designing the science system of the future.

Christine Hine

References

Knorr-Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press.

Lenhard, J., Küppers, G., & Shinn, T. (Eds.). (2006). Simulation: Pragmatic constructions of reality. *Sociology of the sciences yearbook*. Dordrecht: Kluwer Academic Publishers.

Whitley, R. (1984). *The intellectual and social organization of the sciences*. Oxford: Clarendon.