ECSCW 2003 Computer Supported Scientific Collaboration (CSSC) workshop report

Helena Karasti*, Karen S. Baker**, Geoffrey C. Bowker**
*University of Oulu, **University of California, San Diego
(helena.karasti@oulu.fi, kbaker@ucsd.edu, bowker@ucsd.edu)

FOREWORD

The Eighth European Computer Supported Cooperative Work conference (ECSCW 2003) provided a venue to gather researchers interested in the study of scientific collaborations and their technology support. The organizers, Karen S. Baker, Geoffrey C. Bowker and Helena Karasti, started to work together in 2002 on a National Science Foundation funded BioDiversity and EcoInformatics (BDEI) project 'Designing an Infrastructure for Heterogeneity in Ecosystem Data, Collaborators and Organizations' (http://pal.lternet.edu/projects/02dgo/). Having encountered a multitude of challenging issues in their study with the Long-Term Ecological Research (LTER, http://lternet.edu), the ECSCW workshop offered an opportunity for a community dialogue focusing on Computer Supported Scientific Collaboration (CSSC). The suite of selected position papers at the workshop by Debra Cash & Howard Cash, Jenny Fry, Timothy Koschmann, Flemming Meier, Erja Mustonen-Ollila, Giuseppe Psaila & Davide Brugali, and Sanna Talja ([16], http://pal.lternet.edu/projects/ecscw03/) represented a wide range of work relating to scientific collaborations.

INTRODUCTION

In an introduction to Computer Supported Scientific Collaboration, Geoffrey C. Bowker provided some perspective to the emerging era of 'cyberinfrastructures' [1] by identifying three epochs in the history of science and technology: 1) development and rise of the printing press contributing to the notion of accumulation of knowledge [9, 10], 2) growth of governmentality in conjunction with technoscience and large scale data collection supportive of reporting and the garnering of statistics in large scale bureaucracies ([11]; for the effects on science see [19] and [18]), and 3) the central importance of data sharing to the growth of big science this century and last (cf. http://dataaccess.uscd.edu for a recent report on data sharing).

He argues that the challenges facing CSSC today include:

- Understanding the new landscape of publishing
  - The Ecological Society of America, for example, publishes databases to go along with papers;
  - Preprints in physics are increasing in importance in comparison to archival publications in scientific journals (see [22] for a discussion of these issues)
- Working across many disciplines
  - Producing standards and organizational forms that permit good communication across disciplinary divides;
  - Representing uncertainty in data;
  - Dealing with different motivations (see [23] on the Sequoia project)
- Evaluating and assessing
  - Providing formative evaluations for the development of scientific cyberinfrastructure;
  - The need for multimodal research (see http://www.dkrc.org) in tracking these developments.

Today scientific work is going through yet another extensive change that is closely related to technology development. This brings forward further issues, such as, data sharing and data reuse. Though data sharing is nothing new among scientists (e.g. C. Darwin's book [7] on anthropological laughter was based partly on correspondence and surveys with colleagues), scientific journals, per se, are not organized to provide full datasets nor to address the differing data needs of different disciplines. Today Internet technologies offer opportunities for data sharing in vastly more extensive scales (e.g. LTER network) but this also poses problems as data reuse is a delicate matter. Various examples show that long-term datasets exist, but that they are frequently incomplete, badly maintained and not well documented. This presents a chicken and egg dilemma: data exists but is not useable by any but the local user, requiring too much in terms of resources to provide quality assurance and quality control. Furthermore, issues of trust become important with data reuse; scientists tend to use data whose originators they know or who have good reputation (cf. [6]). Two more problems were expressed through examples: 1) cataloguing rare species in biodiversity directories creates the paradoxical situation of needing to mask their locations from the public and scientists are faced with considering how to create multiple views/awareness contexts over data sets, and 2) BIRN (http://daks.sdsc.edu/projects/kbis_birncc.html), the Biomedical Informatics Research Network, which is gathering together MRI's from across the United States, has been unable as yet to fully deal with large scale coordination issues - both technically (being able to recognize that data from a given source comes in a different format) and organizationally (trying to get standards for sharing MRI data which does not trip over the requirements of a local Human Subjects committee somewhere). In addition, there is a scaling from project focuses to scientific collaboratories (http://www.scienceofcollaboratories.com/) and digital libraries (http://www.dlib.org/projects.html) to the emerging concept of cyberinfrastructure [1].

Important issues arising:

- New kinds of publishing avenues, e.g. scientific articles published together with datasets by Ecological Society of America (ESA) and protein databank.
- Reward (and funding?) structures are lagging behind by 15-50 years. For instance, to make a database usable for interdisciplinary research is altruistic work (not rewarded).
- Challenges of interdisciplinary collaborations, both with regard to data and standards across different disciplinary contexts. In interdisciplinary contexts it is essential to preserve the context of data, for instance, to represent original circumstances and reasons for collection. Updates in standards in different disciplines can be highly asynchronous, for example, there are three major measures for radioactive potassium decay and updating to a new standard is major effort; physicists need greater precision, geologists less precise. Another example: Scandinavian countries declined to move from ninth revision of International Classification of Diseases (ICD9) to ICD10.
- We should have awareness for the inevitable situated modification and breaking apart of standards as soon as they are brought to local uses, and we should have protocols in place to bring them back together again.
- Representing uncertainty. For instance, merged GIS datasets make up clear and beautiful pictures, that are also highly uncertain.
- Communication and blending of quantitative and qualitative work.
- 'Speaking through to power', part of intervention which is highly important to CSSC endeavor, but difficult to achieve in writing analytically about the problems faced by scientific collaborations. Funders, particularly.

The challenge is: to what extent are scientists asking new questions with the new technology possibilities. It requires that the possibilities are thought through and explored collaboratively across communities of scientists and technology developers.

POSITION PAPERS

Position papers described case studies from high-energy physics, corpus-based linguistics, cell-biological laboratory, nursing science, history, literature/cultural studies, ecological environmental (laboratory) science, forensics and medical research (see [16] or http://pal.lternet.edu/ projects/ecscw03/). The diversity of case studies helped participants elicit a variety distinctive characteristics of different scientific collaborations and also to highlight commonalities.

Jenny Fry's talk focused on the cognitive and social shaping of scientific collaboration. She drew on three case studies, high-energy physics, social/cultural geography, and corpus-based linguistics [12], to discuss how the specific cultural identity of an intellectual field shapes collaborative work practices and the use of ICTs. Her main argument was that Whitley's [24] theory of the intellectual and social organisation of the sciences, the extent of 'mutual dependency' and 'task uncertainty' manifest in a field, can be applied to predict patterns of ICT usage for collaborative work.

Flemming Meier gave a presentation of an ongoing project where the aim is to study processes of organizational learning and focus on the significance of technological artifacts and systems. Ethnographic investigations were carried out in a cell-biological research laboratory. Preliminary results indicate, that (1) technology is highly integrated in the work, (2) tweaking, 'alternative' use and varying combinations of techniques and instruments are essential for the experimental and innovative work, (3) the technological artifacts are reconfigurable in many ways, and (4) many 'small' situations of collaboration and learning evolve around the use of a certain technique, instrument or machine. Further investigations and analysis will focus on interplays between various individual research projects in the laboratory and how a 'wholeness' of the various projects is be constituted.

Sanna Talja noted the lack of empirical research on collaboration in document seeking, retrieval, and filtering (dsr&f). Forms of collaborative dsr&f range from sharing accidentally encountered information to collaborative query formulation and document synthesis. She described her preliminary findings concerning variation in the criteria for document selection and corresponding variation in collaborative dsr&f practices in research teams and projects, based on a comparative qualitative study across four fields (nursing science, history, literature and cultural studies, and ecological environmental science). She identified four general types of sharing practices: strategic, paradigmatic, directive, and social [20], and discussed the specific challenges and requirements involved in designing systems for supporting these practices.

Debra Cash presented the work of Gene Codes Forensics and the challenge associated with creating an unprecedented bioinformatics tool to support the identification of the remains of the victims of the World Trade Center disaster. The system, called Mass Fatality Identification System (M-FISys, pronounced 'emphasis'), was delivered on a schedule of one-week iterations to New York City Office of Chief Medical Examiner beginning in December 2001. M-FISys had to accommodate constantly changing laboratory and analytical practices, diverse data types and incompatible networks, baroque data nomenclature, new requirements for coordination and communication with outsourced vendors (including high-throughput commercial laboratories), highly compromised and often ambiguous DNA collected from Ground Zero and, not at all least, the requirement that not a single victim be misidentified.

Timothy Koschmann described the notion of an annotation data base within the Professional Competency Project. The Professional Competency Project is a multi-disciplinary, multi-institutional project designed to improve our understanding of what constitutes clinical competency in practical settings. The project revolves around a shared corpus of videotaped protocols of medical students and residents working up cases with simulated patients. Various types of studies will be undertaken within this corpus by different project teams consisting of cognitive psychologists, psycholinguists, sociologists, and Conversation Analysts. The goal is to support collaboration among these teams, not only through the sharing of primary data, but also through the sharing of intermediate findings stored as annotations within the database.

WORKSHOP ACTIVITIES

In addition to the paper presentations the workshop participants collaborated on three group activities in which each participant identified keywords or statements describing 1) central themes in Computer Supported Scientific Collaboration, 2) specific characteristics of Scientific Collaboration, and 3) essential design issues for Computer Supported Scientific Collaboration. Recording keywords on notecards and sharing them on the wall permitted viewing and manipulation of the cards as the group worked together to identify common threads and to coconstruct categories (Figure 1).

Figure 1. Group activities

CSSC Themes

They group started by identifying CSSC themes. Clusters of cards were organized into the following categories: scientific practices, communities of practice, social relations, data, and policy.

Scientific practices One index card 'scientific community <=> practice' brings attention to different approaches in the study of scientific work and collaboration [e.g. 3, 17]. In addressing commonalities in scientific collaborations there are lessons to be learned from studies of disciplinary cultures and interdisciplinary communities, of the practices of carrying out materially and technologically mediated scientific work, and of the knowledge practices and translations in and across disciplines. In the quest for knowledge management, the themes of cooperation in knowledge sharing and of the process of learning arise.

Communities of practice We talked about how scientific communities are hard to study and decipher. Discussion generalized under the umbrella of 'communities of practice' although perhaps 'organizational units' represents a broader perspective. In the study of scientific work and practices, it is not often straight-forward to identify the unit of research under consideration, e.g. domain disciplines, fields, projects, groups, labs, communities of practice, communities of interest, comparative studies or discourse communities. Furthermore, scientific communities differ in their approach to fundamental issues in collaboration identified on the notecards as 'defining a learning process' and identifying appropriate collaboration mechanisms given 'scaling issues'. The concepts of 'boundary spanning' and intermediation between communities arise in/through communication, actors, language, and memory. Inventories of communities may help in the understanding of existing range of scientific collaborations and the development of models accounting for the variation across disciplines in collaborative and information work practices.

Social relations Regardless of the research unit, any collaborative effort involving a group of people brings with it social issues, in the context of scientific collaborations, particularly motivation and trust. Time is rarely dedicated to considering the range of participants and the multiple stakeholders, to negotiating goals and timeframes or to evaluating the state of these issues which shift over time. Further, methods involving observations and technologies enabling surveillance require discussion as to use and ramifications. Initial investigations of sociotechnical aspects of scientific collaborations bring attention to definitions of the end-user as an individual, an organization or a community, and of designers as system builders, observer-participants, and/or mutual learners.

Data Discussion started with the question recorded on a subtheme notecard: 'What do we mean by data?' and opened up into consideration of the seemingly paradoxically divergent data qualities, e.g. objective/subjective, permanent/fluid or evanescent. Additional difficult issues included 'Does replicability equal veracity?', 'How is uncertainty represented?' and 'How to deal with the data explosion?' Further, there is the question as to why data should be shared as data collecting takes considerable effort and insight yet is not rewarded. With data sharing distinguished from data availability, the issues of standards and access arise. The idea of data reuse requires infrastructures, platforms and key tools (such as data mining and document retrieval systems) to be designed to support various kinds of distributed work (e.g. asynchronous and synchronous collaboration and multiple perspectives).

Policy Policy was recognized as an important topic underrepresented in the group. Issues noted explicitly included 'intellectual property' and 'social aspects of publication (especially copyright)'.

SC Characteristics

An organizing premise for the workshop was the recognition that collaborative efforts in general have characteristics in common. Research and development in CSCW focuses on some of these, but some characteristics may be considered unique to scientific collaborations.

Organizational context The heterogeneity in types of scientific collaborations and partnerships suggests the value in articulating the organizational arrangements, including both how a 'project' is defined and what infrastructures (administrative, scientific, educational, and outreach including public and policy interfaces) are supported.

As science seems to be moving toward more holistic understandings of systems, there appears a tendency toward larger scientific projects in order to gain expertise on the many research components of a particular issue. Bringing together and sustaining communication between the many layers of infrastructure and of research components is an effort requiring ongoing attention to developments, changes, and re-negotiations. The tradition of individual informal communications does not scale to meet the needs of larger groups where small changes may affect goal definitions and hence may lead to serious misalignments.

Data context The heterogeneity of data from field and laboratory, from analysis and collaborative work is marked and increases as data is accumulated over time or augmented by new instruments and technologies providing streams of previously unavailable data. Each dataset represents one view into the subject of study. Different datasets may differ in spatial and temporal scale of sampling, in disciplinary and national boundaries. Though each new dataset provides new information, it also represents a use of limited resources.

Collaborative methods context A recurrent characteristic of collaborative work is the need to balance a focus on an outcome with a consideration of the process. It is the nature of scientific endeavors to reform questions or identify new questions. The process of collaboration with brainstorming combining both data and ideas, brings even more potential for unexpected integration and innovative insight that may change expected findings and existing practices. Flexibility is required to re-negotiate goals in order to incorporate change in a project involving multiple partners. A focus on process brings opportunities for formative evaluation, learning and adaptation to change although there is today a gross underestimation of the time needed in planning and supporting larger group interactions. Fields of technology research and development such as CSCW and HCI are articulating and creating mechanisms for collaboration by developing new vocabularies, considering how to optimize competence sharing and to elicit intertwined tacit knowledge as well as how technology and groupware can enhance knowledge sharing. In addition to intellectual sharing, the seemingly straightforward task of time for sharing becomes evident when a topic must be discussed with multiple team members sequentially or a joint meeting planned with a team of colleagues with overloaded schedules and multiple partnership projects. Larger scale collaborative practices are a growing subject of research today. That is, the vocabulary and best practices for the varying types of scientific cooperation are under development as organizations focus not just on information consumption but on knowledge production.

Disciplinary – inter/multidisciplinary science context There are disciplinary traditions for warranting claims. For instance, 'topic' is an entirely different concept in different disciplines and 'systematic literature review' is an entirely different concept in different fields [21]. Consequently, careful discussion is required to understand disciplinary identities and to identify multidisciplinary or interdisciplinary views of a problem. Participants are needed who can maintain contacts and knowledge of activities of colleagues in their own field as well as with colleagues in associated fields.

Traditional - emergent approaches Communication may take the form of publishing a peer reviewed scientific paper; discussing issues with a colleague; sharing pointers and recommendations with a team of colleagues. Today these traditional methods or approaches of competence sharing can be supported by groupware applications. New forms of dissemination are emerging to address collaborative science needs. There is a growing recognition of the need to create recognition for methods and processes not just findings, and for policy infrastructures to broaden epistemological structures. One finds tentative new reward structures and career paths emerging along with new notions of sharing and power.

CSSC Design Issues

Group activities 1-3 produced interesting overlaps which were integrated into the discussion of essential design issues for Computer Supported Scientific Collaboration.

Data issues were seen at the forefront of design considerations in CSSC. Where traditionally CSCW has involved looking at the 'front end' of the collaboration, much of the important investment (and many of the far reaching social and organizational decisions) are now being made at the back end – at the point of database design. Here the role of standards setting bodies is central. The field of CSSC can provide much-needed input on the organizational work of creating and maintaining standards. For this, it is particularly important to stress that standards will always break down over time: and so an understanding is needed of the ways in which standards dissolve, and of how and whether to migrate data into new database structures. Further, data in science times out differently in different sciences (we still need to know the month and often day of publications stretching back to the eighteenth century in systematics and taxonomy – few will go back to astronomical or physics data beyond the latest generation of equipment with its higher resolutions). We have no general understanding of this issue. Digital libraries have a key role to play here [4].

Since scientific work is increasingly important to policymakers (one need only think of climate change, the hole in the ozone layer, the preservation of biodiversity), we need an understanding of the best ways to present scientific data to the intelligent lay public. Further, scientific problems like these are increasingly being addressed by convening interdisciplinary or multidisciplinary teams. We do not know how to facilitate this form of communication (see [8] for a superb discussion of this issue).

Cyberinfrastructures are developing, and will in the future, develop very quickly. We had some 400 years as a species to adjust to the printing press before it began to be superseded – slow moving institutional forms (publishing houses, traditional libraries) are out of phase with the new era of technoscience. A key CSSC challenge here is to contribute to a flexible, ecological design of new systems sensitive to the evolving technological infrastructure.

MAPPING IDENTIFIED CSSC ISSUES TO THE LTER CASE STUDY

After the productive group activities the organizers took on the challenge of mapping the identified CSSC issues against the case study of Long Term Ecological Research (LTER) program as a kind of 'ground truthing' exercise. Karen S. Baker provided an overview of the LTER network aspects which Helena Karasti linked with the identified CSSC issues.

The issues LTER deals with in relation to scientific collaboration and information technology support can be summarized as data heterogeneity, large scale distributedness (both geographic and institutional), multidisciplinarity, and long-term perspective. The LTER program was initiated in 1980 with National Science Foundation support. It was recognized at that time that much of ecological research was addressing time scales of less than a month despite the critical need for long-term research to reveal and understand protracted phenomena.

Currently U.S. LTER involves more than 1200 scientists and students from a diversity of disciplines conducting multidisciplinary investigations of ecological phenomena in a variety of biomes. Recently the LTER ecological multidisciplinarity was augmented by elements of social science with the inclusion of urban sites in the network. Research proceeds at each site independently while participants also increasingly join together for cross-site work and to contribute to the LTER Network. LTER collaborations span not only the 24 research sites of the US LTER, but International LTER consisting of over 25 countries (http://ilternet.edu) and alliances with other research networks.

From the outset LTER placed an emphasis on preserving and sharing data. LTER is one of the pioneers in public sharing of scientific data: since the mid 1990's each site is required to have data publicly available on the Internet two years after its collection. Therefore, data sharing practices and policies are frequently discussed.

Ecological data are extremely heterogeneous, resulting in datadiversity [5] and the specific challenge to LTER information management of how to maintain datasets over the long-term. This has been addressed at the LTER local sites by developing metadata (data about data) forms and traditions of use, at the ecological community level by developing an ecological metadata language standard under the auspices of the Ecological Society of America, and at an intermediate 'project level' by partnering to develop software to facilitate implementation of the standard. Recently, efforts have begun to implement metadata standards at the LTER local site level. Simultaneously, research into alternative aspects of metadata is extending exploration of organizational metadata into the realm of narratives [15].

Centralized network information systems and infrastructures have been developed in support of long-term data and its reuse. These systems invariably involve environmental data, metadata, databases, and increasingly, Internet delivery. Information managers have had a fundamental role in both developing the technological infrastructure [2] and in providing ongoing support for science and data care [14]. Located at research sites information managers have rich understandings of local science domains and expertises, and at the same time they have wide knowledge and skills of technologies, a combination which makes them well positioned as mediators for technology design between ecological scientists, computer scientists and other partners. The LTER has created an environment for information management over time with a grounding in field data, a bridging with domain science, and a sheltered community for collaborative development.

In the continued attempts to bridge between the contexts of scientific work and technology development in CSSC, it is of utmost importance to pay attention, in addition to the above mentioned rich understandings of science domains and participants capable of acting as mediators, to the ways in which interventions are carried out. As part of our action research oriented empirical case study we have presented initial findings of analyses to the community for comment, reflection and dialogue through 'modest interventions' (cf. [13]). Interventions are not only important but also delicate matter. New methods need to be created for the collaborative participation of various roles and standpoints throughout all levels of science and technology activities: from the actual sites of scientific work and collaboration to the venues of formulating science and technology strategies and policies.

CONTINUING CSSC WORK

Possible future activities and openings identified during the workshop for continuing exploration of the CSSC challenges include a special issue in the CSCW journal and a CSSC workshop in a technology venue such as San Diego Supercomputer Center (SDSC) or a participatory venue such as a digital library facility or a design laboratory.

ACKNOWLEDGEMENTS

Our collective thanks goes to the conference organizers and the student volunteers who assisted us during the workshop.

REFERENCES

[1] Atkins, D. E., K. K. Droegemeier, S. I. Feldman, H. Garcia-Molina, M. L. Klein, D. G. Messerschmitt, P. Messina, J. P. Ostriker & M. H. Wright (2003) Revolutionizing Science and Engineering through Cyber-Infrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. January 2003. http://www.communitytechnology.org/ nsf_ci_report/.

[2] Baker, K. S., B. J. Benson, D. L. Henshaw, D. Blodgett, J. H. Porter & S. G. Stafford (2000) Evolution of a multisite network information system: the LTER information management paradigm, Bioscience, 50(11): 963-978.

[3] Becher, T. (1989) Academic tribes and territories: Intellectual enquiry and the culture of disciplines. Buckingham: SRHE & Open University Press.

[4] Bishop, A. P., N. A. Van House & B. P. Buttenfield (2003) Digital library use: Social practice in design and evaluation. Cambridge, Mass., MIT Press.

[5] Bowker, G. C. (2001) Biodiversity Datadiversity. Social Studies of Science, 30(5): 643-684.

[6] Crane, D. (1988) Invisible colleges: Diffusion of knowledge in scientific communities. Chicago: University of Chicago Press.

[7] Darwin, C. (1873) The expression of the emotions in man and animals. New York, D. Appleton and company.

[8] Duncker-Gassen, E. (1998) Multidisciplinary Research at the University of Twente. Enschede, Twente University Press.

[9] Eisenstein, E. L. (1979) The printing press as an agent of change: Communications and cultural transformations in early-modern Europe (volumes 1 and 2). Cambridge: Cambridge University Press.

[10] Eisenstein, E. L. (1983) The printing revolution in early modern Europe. Cambridge: Cambridge University Press.

[11] Foucault, M. (1991) Governmentality. In G. Burchill, C. Gordon & P. Miller: The Foucault Effect: Studies in Governmentality. Chicago, University of Chicago Press.

[12] Fry, J. (2003) The cultural shaping of scholarly communication within academic specialisms. Ph.D. thesis, University of Brighton.

[13] Heath, D. (1997) Bodies, Antibodies, and Modest Interventions. In Downey, G. L. & J. Dumit (eds.) Cyborgs and Citadels: Anthropological Interventions in Emerging Sciences and Technologies. School of American Research, Advanced Seminar Series, SAR press. pp. 67-82.

[14] Karasti, H. & K. S. Baker (2004) Infrastructuring for the Long-Term: Ecological Information Management. Hawaii International Conference on System Sciences 2004 (HICSS'37). January 5-8 2004, Hawaii, USA. (accepted)

[15] Karasti, H., K. S. Baker & G. C. Bowker (2002) Ecological Storytelling and Collaborative Scientific Activities. SIGGROUP Bulletin 23(2): 29-30.

[16] Karasti, H., K. Baker & G. C. Bowker (eds.) (2003) ECSCW 2003, Proceedings of the Computer Supported Scientific Collaboration workshop, Eighth European Conference on Computer Supported Cooperative Work, Helsinki, Finland, 14 September 2003. University of Oulu, Department of Information Processing Science, Research Papers Series: A34. ISBN: 951-42-7121-1.

[17] Latour, B. & S. Woolgar (1986) Laboratory Life: The Construction of Scientific Facts. Princeton, University Press.

[18] Luke, T. W. (1999) Environmentality as Green Governmentality. In E. Darier (ed): Discourses of the Environment. Oxford, UK; Malden, Mass., Blackwell, pp. 121-151.

[19] Serres, M. (1990) Le Contrat Naturel. Paris, F. Bourin.

[20] Talja, S. (2002) Information sharing in academic communities: Types and levels of collaboration in information seeking and use. New Review of Information Behavior Research, 3: 143-159.

[21] Talja, S. & H. Maula (2003) Reasons for the use and non-use of electronic journals and databases: A domain analytic study in four scholarly disciplines. Journal of Documentation, 59(6): 673-691.

[22] Walsh, J. P. (1998) Scientific Communication and Scientific Work: A Survey of Four Disciplines. Internet Research 8(4): 363-366.

[23] Weedman, J. (1998) The Structure of Incentive: Design and Client Roles in Application-Oriented Research. Science, Technology, & Human Values 23(3): 315-345.

[24] Whitley, R. (1984) The intellectual and social organisation of the sciences. Oxford, Clarendon Press.