Michael Nentwich

cyberscience

Research in the Age of the Internet

Chapter 0

PROJECT OUTLINE AND BACKGROUND

Austrian Academy of Sciences Press
Vienna 2003

Submitted to the Austrian Academy of Sciences on 10 April 2003
by Gunther Tichy, member of the Academy

British Library Cataloguing in Publication data.
A Catalogue record of this book is available from the British Library.

ISBN 3-7001-3188-7

Austrian Academy of Sciences Press
Tel. +43-1-5129050-3405, Fax +43-1-51581-3400,
Postgasse 7, A-1010 Vienna
Email: verlag@oeaw.ac.at
http://hw.oeaw.ac.at/cyberscience

Layout, cover & type-setting: Manuela Kaitna, A-1080 Vienna
Printed and bound in Austria by Manz Crossmedia GmbH & Co KG, A-1051 Vienna

DETAILED LIST OF CONTENTS

0 PROJECT OUTLINE AND BACKGROUND

Since the early 1980s, the scholarly community has been witnessing a considerable increase in the use of information and communication technologies (ICT). The networked PC, E-mail, the Internet, on- and off-line databases, the World Wide Web, electronic publications, discussion lists and newsgroups, electronic conferences, digital libraries and "knowbots" are but a few of the trends which increasingly influence the daily work of the scientific community. As opposed to "traditional" science and research which has done without networked computers, the notion of "cyberscience" will in the context of this study designate the use of these ICT-based applications and services for scientific purposes (see 1.1 for a formal definition).[1]

Many have already compared these developments with the sea change triggered by Gutenberg's printing press.[2] In his often quoted 1991 article on the "Post-Gutenberg Galaxy", Stevan Harnad compared the advent of electronic communication and publishing with "three revolutions in the history of human thought": language, writing, printing (Harnad 1991). Indeed, an initial synopsis of the various developments on the path to cyberscience leads to two remarkable conclusions which lend legitimacy to the use of the label "revolutionary"[3] in this context: First, ICT affect virtually all forms of scholarly activity. Second, the many developments faced by scholars do not only accelerate communication or increase the publication output, as frequently assumed, but also have the potential to lead to qualitative changes in the research community and as regards even the content of research itself.[4] This is a remarkable and exciting point of departure for my study: *How, specifically, will technological developments change the ways research is done?*

1 If this book were written in German, a tricky problem with terminology would not have arisen. The notion "Wissenschaft" is very encompassing and does not differentiate between the natural and the social sciences, the humanities and cultural studies. All of them are "Wissenschaften" (in the plural). Therefore the notion of "Cyber-Wissenschaft" would not lead anybody to think of the "hard" sciences alone. By contrast, the English term "science", when standing alone, primarily refers to the natural sciences. My topic, however, includes all the various sciences, including the social *sciences* (sic!) and the humanities etc. In English, there seems to be no straightforward, unambiguous "shorthand" solution. "Academia" normally refers to the world inside universities. "Scholarship" is mainly used with reference to the humanities. Perhaps the notion of "research" covers most aspects of my topic. However, even "research" is often connected to the activities going on in laboratories (as in the notion of "research and development"). Nonetheless, I need a practical solution that does not force me to use a cumbersome writing style repeating over and over again that my statements refer, at least in principle, to all "families" of academic disciplines.
To make this point clear, this book's title speaks of "cyber*science*" and the subtitle of "research". In the main text, if not otherwise noted (e.g. in chapter 3), I use the terms "science" (including "cyberscience") and "scientific" *not only* as referring to the natural sciences. As much as possible the term "scientific" is, however, replaced with "scholarly" or "academic" which are treated synonymously. The same applies to "scholarship" and "academia", which should be equally read as having a broader meaning.

2 In particular Marshall McLuhan (1962; followed by e.g. Harnad 1991; 1998; 1995; Dewar 1998; Zeigler 1997; Frühwald 1998; Bolz 1993; Birkerts 1997).

3 A study commissioned by the OECD acknowledges that "the quantity, quality and speed of communication among scientists" has been enhanced, "ICT use has also had various effects on the organisation of work in science", among them that "(c)ollaboration patterns have changed, the science base has widened as more scientists are able to participate, and scientific hierarchies have sometimes been affected". The study comes nonetheless to the conclusion that "for the most part, scientific work has not been revolutionised" (OECD 1998, 195, quoting Walsh 1997).

4 See 1.1.1 for an elaboration of this theme.

These are some of the questions we shall ask: Do we have to expect that new digital products, such as „hyperbases“ will finally replace the traditional paper publications? Has the Internet positive or negative impact on the productivity of the science system? Will the academic community regain control over the academic publishing sector? How can we explain the differences between the various academic fields when it comes to the use of the new technologies? What will be the role of the academic librarian of the future? How will academia secure quality in a digital world?

This chapter sets the scene by discussing the status quo of research on cyberscience (0.1) and how this study differs from such earlier or parallel studies (0.2). The main part is devoted to a description of the empirical research strategy (0.3), followed by a short overview of the whole study (0.4).

0.1 The status quo of research on cyberscience

Cyberscience is a moving target. There is no day without a new E-journal seeing the light of day, without new "cyber-collaborations" or without a new piece of software which might change the way scholars work. For sure, many of these new developments are nothing more than the repetition of something already tested elsewhere (in particular in business), but in a new context, by other people. Cyberscience is not only expanding, but also evolving and innovating along the path. Researchers of all fields are trying to stay abreast with this moving target, but hardly succeed. If articles contain empirical evidence, they are outdated the minute they are published.

Even beyond the speed aspect, *cyberscience is an elusive subject* since research done in the area so far is fragmented and often unsystematic. While various aspects of the application of ICT in academia are dealt with in many publications separately, there is no comprehensive, comparative and synthetic study yet. This seems mainly due to the fact that the discourse on the various aspects takes place in communities that are rather isolated from each other. Scholars in communication sciences, sociology of science and law as well as researchers of knowledge organisation, Internet economics and research policy studies are all interested in specific facets of the topic. Furthermore, the librarian community and, last but not least, the more technically oriented authors such as software developers, human-machine-interface researchers etc. should be mentioned. There is thus a particularly fragmented research landscape which will be further analysed under the following sub-headings.[5] While I shall look at the relevant literature from technology studies and technology assessment in 0.1.2 and 0.1.3 respectively, the following 0.1.1 will report the various contributions of the above mentioned research streams.

[5] There is, however, no space to give an extensive overview of all publications in the area. The following overview can only quote items of particular quality and of direct relevance for this study. A multitude of further studies shall be quoted in the following specialised chapters. For a more detailed account of the existing literature at the beginning of the project, see Nentwich (1999a).

0.1.1 Various research streams pertinent for the study of cyberscience

We can distinguish the following major streams or areas of research relevant to the impact of ICT on science and research: (1) publishing; (2) communication and collaboration; (3) libraries; (4) teaching; (5) policies; and (6) general research.

(1) The largest amount of contributions in the field of cyberscience has been written with respect to electronic *publishing*. There is no integrative book one could point at, but a series of papers on specific aspects by American authors in particular. The very comprehensive and constantly updated bibliography by Charles Bailey (1995 ff.) on scholarly electronic publishing contains well over 1,000 entries for the period from 1990 to date.[6] The headings deal with e.g. economic issues, general works on electronic serials, but also E-books and related legal issues. The questions raised include the pricing policies of commercial publishers with respect to print and E-journals, new alternative publishing models and intellectual property rights in the electronic world. Riehm and his team (1992) gave an early (book-long) “critical inventory of E-publishing”. There are also a number of empirical studies in this respect. In the first place, they deal with E-journal use by university researchers and students (e.g. Rusch-Feja 1999; 1999; Gomes/Meadows 1998; Woodward et al. 1998), but there are also some on E-books (e.g. Siler 2000). Recently, a study using the Delphi method has been carried out, which generated interesting convergent expert opinions regarding the future of scholarly journals (Keller 2001a).[7]

(2) With respect to changes in the academic *communication and collaboration* structure, Harasim/Winkelmans (1990) present earlier studies under the headings “sociological effects and impacts”, “social psychology and behavioural change”, “organisational impact”, “technical research and development” and “educational research”. Only a small number of rather status-quo-oriented and limited (and now outdated) larger empirical studies have been carried out.[8] A first example is Rutenfranz (1997) who studied the significance of the computer for research and teaching from the perspective of the communication sciences based on an empirical survey at one German university in 1993. With a similar approach and also based on a survey during 1992-3 in Germany, Scholl et al. (1996) examined the use of E-mail and discussion lists. The latter and their impact on academia are also discussed, for instance, in the empirically based works of Uwe Matzat (e.g. 1999; 2000; 2001; see also Gresham 1994; Hert 1997). Less elaborated are a number of individual contributions identifying changes in the scientific communication, e.g. with respect to the growing use of E-publications in the scholarly world, to the citation and hence impact of E-journals (e.g. Harter 1996b). The same is true for articles regarding groupware applications intended to support scholarly co-operation (e.g. Valkenburg 1998). Co-operation in the digital age and CMC use in general is the subject of a number of empirical studies by a team around John Walsh (e.g. Walsh/Maloney 2002; Walsh/Roselle 1999, 50ff.; Walsh/Bayma 1996). Collaboratories, in turn, are well researched by Finholt (e.g. 2001).[9] E-communication and research has already been the subject of many conferences. Some of them resulted in book publications (such as Meadows/Böcker 1999; Best et al. 1994).[10]

6 <Cyberlink=134> (see 0.3.2 for an explanation of this notation of links to the WWW).
7 All of these earlier contributions and many more will form the basis of chapter 6 to 8.
8 An overview of empirical studies is given by Zhang (1998, 242-244).
9 This stream of literature will be analysed in more detail in chapter 4 in particular.
10 See for a list of relevant workshops and conferences <Cybercategory=13>.

(3) Another major theme under the heading of cyberscience is the evolution of digital and virtual *libraries* (e.g. Lesk 1997; Bishop 1998). There are also studies on the use of online bibliographic databases in various disciplines (David/Zeitlyn 1996; Mann 1998). Naturally this discussion takes place mainly in the librarian community and is linked to a number of other issues, such as digital archiving, meta-data standardisation and, above all, economic issues. In general, the research on digital libraries does not deal with organisational and sociological issues, but mainly "concentrates on how to develop the necessary infrastructure to manipulate effectively the massive amounts of Internet information" (OECD 1998, 204). In this context, we should also list those projects dealing with the electronic book and their eventual impact on reading habits (e.g. Böhle et al. 1997; Siler 2000; Wearden 1998a). Furthermore, many librarians are prime contributors to the debate about the future role and financing of academic libraries.[11]

(4) With regard to *teaching*, there is an abundant literature on tele learning, in particular on its specific applications. However, only a few general discussions of the issue of next-generation universities exist (see e.g. the special issue of Futures, Vol. 30, no. 7, and McArthur/Lewis 1998). What is missing, in particular, is an in-depth discussion of these issues in relation to the research system as a whole.[12]

(5) As to the impact for science and research *policies*, the OECD commissioned two conferences focusing on the sciences as opposed to the humanities (OECD 1998). Worth mentioning are also the various policy papers issued by ad hoc groups in the various communities, such as the Pew Higher Education Roundtable (1998). Concerning ICT use in higher education, I should mention a RAND publication from the US (McArthur/Lewis 1998, 87ff.). As with the other subjects, however, an encompassing assessment with a view to policy-making in the area is still lacking.[13]

(6) A few individual researchers and institutions contribute or are planning to contribute in one way or the other to the overall question raised in this study, that is the impact of ICT on the inner workings and outcomes of research. The research centre NERDI (Networked Research and Digital Information) of the Netherlands Institute for Scientific Information Services (NIWI) in Amsterdam carries out related research in its programme on "Knowledge production in the new digital networks".[14] Berkowitz' (2002) ongoing research on academic research paradigms promises to be highly interesting, as will be the study of Vasileiadou (2001; 2002).[15]

According to Walsh/Roselle (1999, 71f.) who analysed the existing empirical research on cyberscience, the following is still needed (excerpt):

> "longitudinal analysis that would help untangle the direction of causality [between productivity enhancements and CMC]";
> "impact of CMC on the speed with which science advances";
> "how these differences [in CMC use across disciplines] may be affecting the advance of science in different fields";
> "good empirical work [on] the uses and effects of the Internet across countries";

[11] This stream of contributions will form the point of departure of my analysis in chapters 5 and 9.

[12] While I shall not analyse the general literature on teaching in the digital age, I shall discuss the issue in chapter 5.

[13] Policy issues will be dealt with throughout this study, in particular in chapter 11.

[14] <Cyberlink=838>.

[15] Furthermore, the US National Science Foundation initiative "Implications of Information Technologies" is planning to commission a literature review about the implications of IT for, inter alia, education and science (<Cyberlink=227>).

> "longitudinal data and field experiments separating out the effects of other changes in scientific problems and institutional conditions (such as funding policies) to determine what the actual impact [of CMC] has been".

This is a committed research programme not yet carried out to date. The present study will, as outlined below, contribute to this programme by adding to all but the last point.

To sum up, while a considerable amount of research has already been done in the area, this literature is fragmented because so many disciplines and communities have an interest in the topic. Large empirical studies are the exception. Many important research questions have not yet been addressed. As this study aims at contributing both to science and technology studies (0.1.2) and technology assessment (0.1.3), the status quo of related research in these two areas is summarised in the following two sub-sections.

0.1.2 STS and cyberscience

Science and technology studies are a diverse and open field. Some of the authors quoted above would claim to belong to this colourful group of researchers. Others would perhaps not do so, but still study aspects of science and technology from a meta-perspective. However, many contributions to the study of cyberscience come from specialists in the information sciences, library studies, legal studies etc. Still others are researchers usually working in a particular field of science, but who give accounts of the ongoing developments in their discipline, often with a view to making their community aware of what is going on.

STS is often about either "S" that is science (e.g. how scientists arrive at results) or about "T" that is technology (e.g. how society reacts to a new technology or how the latter is shaped by the former). Only seldom it is about both at the same time, that is about technology (use) in science. The contributions of the STS community to my research topic are therefore either general – and hardly related to ICT use – or they treat the latter only partially. The STS literature of the first group which inspires my analysis here relates to the role or place of personal encounters in science (e.g. Traweek 1988); to informal networks (e.g. Crane 1972; also Stichweh 1989); to disciplinary differences (Becher 1989; Knorr Cetina 1999); to reputation and status (e.g. Rost 1996c; Fröhlich 1993); and to technology diffusion research (e.g. Rogers 1995; Rammert 1990). The second group is characterised by a targeted, subject-specific approach, for instance, on collaboration (e.g. Finholt/Brooks 1997; Walsh 1997), on digital libraries (Kilker/Gay 1998), or on scientometrics (e.g. Harter 1996a; Zelman 2002; Leydesdorff 2001).

What is needed in this situation is a study aiming to both bridge the gap between the general STS literature – which will inform my conceptual framework (cf. 1.2) – and ICT-related changes, and to fill the many holes in the scarce STS literature which already treat ICT aspects. In other words, we need *a study on technology use in science and research*.

0.1.3 TA and cyberscience

Cyberscience has many characteristics of an ideal subject for technology assessment: It is about emerging technologies whose impacts are already partly visible in the present and which have the potential of widespread application. It calls for an encompassing study because the various aspects are strongly interrelated. The topic needs to be treated in an interdisciplinary manner, as the impacts are both in the political, cultural, legal, economic and social sphere. Last, but not least, it is about an ongoing development reaching into the future, so it makes sense to look at it not only from an analyst's perspective, but also with a view to eventually formulating policy recommendations.

Notwithstanding this, technology assessment worldwide has only dealt with individual aspects of cyberscience, so far. For instance, the "deceased grandfather" of all technology assessment institutes, the Office of Technology Assessment (OTA) of the US Congress, did not have the time to fully appreciate cyberscience (as it was closed in September 1995). Earlier reports dealt with teaching, learning and education (OTA 1982; 1988; 1995b; 1995a), with intellectual property rights (1986), with high performance computing and networking for science (1989; 1991). The impact of ICT on science and research in particular has never been the topic of the OTA. In Europe, too, technology assessors looked only at partial subjects with no particular focus on academia, such as the ITAS in Germany on E-books and E-publishing (Riehm et al. 1992; Böhle et al. 1997) and the European Academy (Bad Neuenahr) on copyright issues (Banse/Langenbach 1999). SPRU in the UK has a related project on the introduction of advanced ICT technologies and its impact on the relationship between tacit and codified knowledge. Other TA-related research units are dealing with tele teaching and various other aspects of ICT-induced changes. However, the focus is never on academia but always much broader related to society at large. Hence, this study will be the first to treat the topic as a whole.

0.2 Innovative aspects of the project, aims and methodology

As outlined above, the research done in the area so far is fragmented, often unsystematic and in most cases not theoretically informed. Some of it is, to a certain degree, rather politically than scientifically motivated in the sense that the publications were targeted at influencing the behaviour of university administrations or the library community (in particular in the field of E-publishing). Most importantly, there is *no synthesis available* which would enable us to see the full shape of cyberscience and to draw overall conclusions.

Starting from this state of affairs, this study breaks new ground by developing an *encompassing conceptual framework* and by synthesising the knowledge available. Furthermore, this study goes beyond previous research by explicitly studying in-depth two important questions which, so far, have been addressed only occasionally and unsystematically: the differences (in type of applications, use, acceptance etc.) among the various disciplines; and the qualitative changes of structures and procedures in academia, on the path to cyberscience.

Consequently, the basic aims of this study are five-fold: (1) to describe and analyse the use of ICT in the academic world (status quo); (2) to explain the use based on a model;

(3) to draw a realistic and differentiated picture of probable future developments (trend extrapolation); (4) to assess the impact of ICT on various aspects of academic activity and on the substance of research (impact assessment); as well as (5) to discuss the implications for research policy and the steering mechanisms within scholarly organisations. To achieve these aims, a complex structure of this study is required. There are no simple research steps to be elaborated in a consecutive way, but the aims can only be reached in an integrative way,[16] as shown in Figure 0-1:

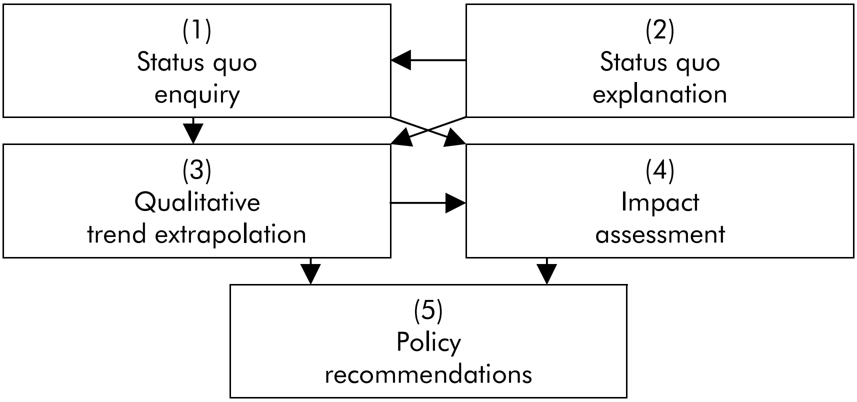

Figure 0-1: Relationship between research aims

Central to this research is the notion of "qualitative change", both in the way scholars work and in the way the research system is organised. In contrast to purely quantitative change such as in the speed of communication and in the amount of publications, I define "qualitative" as a change in kind rather than in degree. For instance, structural change concerns the organisation of academia and change with respect to the roles played by the different actors in the research system. Other areas of possible qualitative change are the characteristics of scientific communication and the way in which scholarly knowledge is processed. Last but not least, qualitative change can refer to shifts in the substance of the research or in the way the research questions are addressed.

The following *guiding questions* will be addressed in order to contribute to our understanding of the developments at stake:

- Do the new applications have the potential to substitute more traditional forms?
- What are the factors influencing the process?
- Which kinds of qualitative and quantitative change do we observe?
- How can we explain the differences in time, discipline and location?

[16] Consequently, research aim (1) is pursued in particular in the whole chapter 2, in sections 3.2 and 3.3, and passim also in chapters 4 to 8 (in particular in 4.1, 5.4.1, 6.2 and 7.1): research aim (2) is mainly dealt with in 3.4, based on the conceptual framework outlined in 1.2.3; I contribute to aim number (3), based on my framework developed in 1.2.3.5, at various places throughout this study, in particular in 6.3, 6.5, 7.3.2, 7.3.3 and 8.3, and in a comprising way in 12.1; impact assessment (aim 4) is, based on the framework in 1.2.4, the topic of most of chapters 4 to 10; finally, research aim 5 is mainly fulfilled in chapter 11, but also addressed in some of the interim conclusions throughout the study, e.g. in 9.3.

Given the great variety of aspects to be dealt with when trying to draw a well-founded and integrative picture of the state of academia in the not-so-distant future, it will be necessary to study cyberscience on the basis of an *interdisciplinary approach* (as known from technology assessment; cf. 0.1.3 and below). This study's main focus is in the area of the social studies of science and technology (cf. 0.1.2). Technology and innovation studies covering e.g. diffusion research and the formation of technologies are also of relevance. In addition, a number of further disciplines inform this study, in particular political science (policy conclusions to be drawn, research policy system), sociology (social shaping, social organisation, group behaviour, roles), communication sciences (technical aspects of scholarly communication processes), legal studies (copyright and contractual issues) and economics (Internet economy, pricing policies, market research). As far as necessary, relevant studies from these disciplines were included in my analysis (cf. 1.2).

This study stands in the tradition of *technology assessment* (cf. 0.1.3). In addition to having an interdisciplinary impetus (social sciences, law and economics), this means that it focuses on impact assessment and heads towards policy recommendations (cf. 11.3). What cyberscience will look like in a few years time is not simply given but depends (inter alia) upon political decisions of a variety of actors both inside and outside academia. This study is a typical technology-induced TA study as I start from a particular technology (ICT) and look at its impacts on the societal sub-system “academia”. It is an expert TA, as it is mainly based on a literature survey and on a series of expert interviews, but not on lay participation. It includes all the classic components of a TA study (Porter et al. 1980, 54 ff.). I shall start with a description of the technology and a foresight of the future technological development (chapter 2), and discuss the present and future societal sub-system (throughout the study). The identification, analysis and evaluation of impacts of cyberscience is contained in Part Three; the analysis of political options is in chapter 11. An executive summary as well as contributions in non-scholarly journals and the mass media serve to communicate the results in a generally accessible form.

While this study may be read as a TA study, it does at the same time contribute to an academic research field, namely *science and technology studies*. The conceptual framework developed for this study and presented in chapter 1 is not only based on STS approaches (cf. 1.2), but is also intended to contribute to a systematic understanding of both scholarly communication and the relationship between it and technology. Inter alia, it defines and systematises the crucial characteristics of scholarly communication, such as communication partners, means of communication, properties and functions of the publication system (1.2.2). This study proposes a number of classifications of key elements of academia, such as the current quality control systems (8.1.2), and a typology of academic publications (7.3.1.3). Furthermore, it gives a comparative account of thirteen academic disciplines as regards the status quo and the current developments of the use of ICT (chapter 2). It also discusses a number of classical concepts of the STS literature in the changing environment, such as the Matthew effect (5.6.2) and invisible colleges (4.3.5.1).[17]

The main thrust of this research is “hypotheses generating” (not: hypotheses testing) and problem-oriented (not: variable-oriented, see King et al. 1994). This is explorative and predominantly qualitative (rather than quantitative) research. Quantification is nevertheless used wherever reasonably appropriate, to analyse the different disciplines in chapter 3. In any case, in this report the language of hypotheses and variables is used

[17] For a more in-depth presentation of the four main circles of thought informing this study in general, and its conceptual framework in particular, see 1.2.

(similarly Walsh/Bayma 1996, 360f.). This is a deliberate choice. I hold that this language helps to structure ideas and focus our thinking when I discuss which factors influence the developments under scrutiny (similarly Ragin 1987; 2000). It is useful to distinguish between dependent and independent variables and between necessary and sufficient causes. Even if the latter seem not to exist here, searching for them and contrasting this ideal concept of mono-causality with the multi-faceted social reality in academia sharpens our arguments.

0.3 Outline of the empirical research

What will be the future of research in the age of cyberscience? With a view to analysing the status quo, the trends and the impact of ICT use in science and research, the empirical research is based on four pillars:

- First, I gathered knowledge about the forms of ICT use in the various disciplines (see 0.3.1) through an intensive *Internet enquiry*. A side product of this enquiry is the Internet link collection CYBERLINKS (see below 0.3.2).
- Second, I took into account the empirical evidence dispersed in the *research literature* (see above 0.1 for a brief guiding overview and classifications).
- Third, I carried out a series of 50 *in-depth interviews* to gather information about the actual use of the new tools in the various (sub-)disciplines (see below 0.3.3).
- Finally, I carried out several *practical tests* to explore some of the innovative tools at the disposal of cyber-scientists today and possibly in the future (see below 0.3.4).

To explain the differences between the various (sub-)disciplines, the interviews included a series of *general questions* about each field. These questions are intended to substantiate hypotheses derived from my basic theoretical model (see chapter 1) on the factors influencing whether and how a particular sub-discipline readily adopts the new media.

Finally, with a view to exploring the possible impact of ICT (in its present form and in the future) on the substance of research (chapter 10), I asked my interviewees about their *opinion* on a series of statements. Each statement concerns a change in technology, on the one hand, and a possible (direct or, in most cases, indirect) change of the content of research, on the other.

0.3.1 The selection of disciplines

The main aim of the study is to explore the research subject in a way that leads to as much information as possible about the ways in which the new technologies are used and how they influence the daily work of the scholars and their research. Given this explorative impetus of the study, I chose to include a *large variety of academic disciplines* (13). Among them are both the sciences (applied: molecular biology, medicine, parts of physics; and pure: high-energy physics), the humanities (parts of language studies, philosophy, papyrology), the social sciences (political science, sociology, economics), as well as cultural studies (parts of language studies, ethnology) and two fields which can hardly be attributed to any of the former categories (mathematics, law). See Table 3-1 (on p. 108) for all disciplines and sub-disciplines included in the comparative part of this study.

On the basis of earlier studies (e.g. Becher 1989) I expected, and it actually turned out quickly in the course of the interviews, that each expert named a very precise field as his/her speciality, often even an intersection of *specialities*. Therefore, not two of the experts actually belong to the very same speciality. They rather come from related fields. Note that this is not a problem of sample design, but reflects present-day scholarship that is ever more specialised. On the one hand, this makes it more difficult to compare directly between the experts interviewed from every one discipline. To counter-balance, cross-checking has been done by other means as well (e.g. Internet search and literature). On the other hand, the resulting picture of each discipline is more realistic and “colourful”. Given the above-mentioned explorative aim, this has to be considered a virtue. In terms of the explanatory aim, this variation makes it certainly more difficult to aggregate the data for whole sub-disciplines or even disciplines. However, it would not have made sense to hide the considerable intra-disciplinary differences by (over-)generalising from some coherent, but therefore extremely narrow speciality. In the present study, any explanatory attempt has to take this great empirical variation into account.

0.3.2 The Internet enquiry

Being the main arena of innovative ICT-based communication among academics, I had to explore how researchers actually use the Internet. I visited hundreds of academic websites looking for both standard and advanced homepages, and applications such as databases, virtual institutes, digital libraries and the like. Furthermore, the WWW offered answers to many of the technical questions raised in chapter 2 on the new technologies, tools and applications.

The result of this Internet inquiry is stored in an online database named CYBERLINKS, which can be found at the following web address:

http://www.oeaw.ac.at/ita/cyberlinks.htm

As I will discuss in more depth in other sections (in particular in 6.4.1.3), one of the most pressing problems with the WWW is its fluid character. That is, the addresses – URLs – are not persistent. They often change without notice. In consequence, quoting Internet addresses is a risky business as it may well be that by the time of reading, the quoted URL is no longer available. Therefore, no direct URLs are given in this study, but indirect quotes of entries in the CYBERLINKS database. The latter is constantly being updated so that changes of URLs will be reflected in the database. The identification number of a particular link in the CYBERLINKS database, however, will not be changed. Therefore, the quotation of a “cyberlink” in a footnote of this report remains valid (for this link database will be regularly updated). Hence, the footnotes pointing at a “<Cyberlink=XX>” whereby “XX” is a number, should be read as a link to the following URL:

http://www.oeaw.ac.at/cgi-usr/ita/cyber.pl?cmd=search&link=XX

Annex VI gives a complete list of all URLs in the database as of summer 2003.

The CYBERLINKS database is structured along so-called “categories”, that is groups of related links. In some cases, it made sense not to quote a particular link, but a whole category. This is indicated by the term “<Cybercategory=XX>” whereby “XX” is again a number and should be read as a link to the following URL:

http://www.oeaw.ac.at/cgi-usr/ita/cyber.pl?cmd=get&cat=XX

In the online version of this report, the “Cyberlink” and “Cybercategory” footnotes are active hyperlinks leading directly to the respective entry in the CYBERLINKS database.

0.3.3 The interviews

For the empirical part of this study, many interviews have been conducted. They were of two different kinds. On the one hand, I carried out in-depth interviews on the basis of a semi-structured, open questionnaire with researchers in some selected disciplines during winter 2001/02 (see below). On the other hand, a number of further people have been interviewed for this study on a less formal basis. In particular, I talked to researchers from a variety of other disciplines, to computer experts, to librarians and to publishers at several occasions all along the project period, that is from autumn 1998 to autumn 2002.

0.3.3.1 In-depth interviews versus anonymous questionnaire

The function of the empirical part of this study is not quantitative measurement, but exploration. Given the wide range of problems at stake and hence the relatively large number of issues raised (over 70 questions), it seemed even futile to expect a sufficiently high return rate if an anonymous questionnaire were sent out. Many of the questions are of an informational value only (e.g. if I asked about the existence of E-journals or of multimedia applications in the field). With a view to this type of question, the answers of all experts in one field should be roughly the same. So it would not have made sense to bother too many researchers with such a large questionnaire. Furthermore, many of the questions are of a qualitative nature. They often needed additional explanation and even discussion – something undeliverable through the instrument of a questionnaire. With a view to exploration, I therefore settled for a survey among experts. Due to the above mentioned qualitative nature of the questions and the need to allow counter-questions and open discussion, I concluded that in-depth expert interviews based on a half-standardised questionnaire (see Annex III) were the best option.

It could have been argued that the results of a relatively small number of in-depth interviews could more likely be *biased due to individual preferences*, in particular with a view to the last group of questions (those asking for opinions about the impact of ICT). Since too small a number of interviews per (sub-)discipline could allow hazard to play a role, it seemed advisable to settle in the middle ground and stick to the concept of in-depth interviews, but make as many interviews as reasonably possible while still including a broad range of interview partners (all in all, 50 interviews). Despite the obvious restrictions in time and resources, at least three interviews per discipline have been conducted. Furthermore, the empirical evidence given by the interviewees was cross-checked through the other empirical pillars, i.e. the Internet search as well as the evidence found in the literature.[18] By this token, I trust that the results from my interviews together with this additional information do reflect what is actually going on in the various fields.

[18] A similar approach has been taken, for instance, by Walsh/Bayma (1996, 346) who supplemented their data generated through the interviews "with archival data to get broader information on changes in each discipline. This use of archival data also provides a check against some of the problems of self-reported data."

0.3.3.2 The interviewees

All interviewees were carefully selected in order to discuss the issues at stake with people of an *expert status*. The main criterion of selection was whether or not the people addressed appeared to have an above-average experience with both the new technologies and their own discipline. Only scholars with significant experience could be considered to be in a position to provide both the empirical information requested and a more than ad hoc assessment of the possible impact of the technologies. At the beginning of each interview I asked whether the interviewee considered him/herself experienced with the new media. The majority answered with being an experienced user. At least one person in each field (with the exception of sociology) was also technically well informed and able to actively shape the Internet (e.g. writing in HTML or even programming databases). I furthermore tested each interviewee's particular experience with international co-operation, international conferences, association membership, video-conferencing, E-journals, groupware and knowledge of some telling technical terms. Taken together, the vast majority of the interviewees may be called experts with regard to the use of the Internet in their field.[19]

While the expert status of the interviewees was particularly helpful with regard to the factual questions of the questionnaire, there is no need to worry about a potential bias. It turned out that the experts could not be labelled "technology freaks" with regard to their expectations of the future. They were not solely affirmative in their evaluation. Quite to the contrary, they often showed a critical approach based on their above-average use of ICT.

The *choice of interview partners* started with personal acquaintances and colleagues in the relevant disciplines (see above). Further respondents were chosen on the basis of both proposals given by the first interviewees (snowball system, similar Walsh/Bayma 1996, 346) and investigations of the homepages of various research institutes in the proximity of the offices of the author (Vienna and Cologne).

Further to their status as experts, the process of selection of the experts was guided by the principle of variance, tempered by practical considerations. First, I included researchers from various backgrounds, both university (36) and non-university research (14). Second, both more and less senior researchers were interviewed: 18 heads of research institutions, four group directors, 26 senior scientists; two interviewees turned out to be more junior than expected, in general terms, but were still valuable experts on specific aspects of relevance for this study.[20] Grouped differently, I counted at least ten among the interviewees that could be rated as "leading" researchers in their field or country.[21] Third, researchers from both sexes were included, but less successfully (45:5). This is due to the empirical gender imbalance among senior researchers, not to research design. Fourth, the interviews were not only carried out in one but in five cities and in three countries. However, as indicated above, the interviewees were addressed as experts in their field, not as representatives of a sub-population.

[19] The small minority of those that turned out to be less Internet-literate after all (5 out of 50) were still useful for serving as a possible corrective from a less technically advanced point of view and for establishing the general co-ordinates of the sub-discipline. Note that, on a whole, my experts were not overly optimistic, in particular as regards the future of ICT use in their respective fields, but rather sober and pragmatic. This is different from findings in other technology-related studies (cf. Tichy 2002).

[20] See column "Status II" in Table A-5.

[21] See column "Status I" in Table A-5.

As to a *possible geographical bias*, the interviews did not show any sign of noteworthy difference in that respect. The interviewees were based in three Western European countries – Austria (32), Germany (14), and Netherlands (4).[22] With a view to avoid any local effect, I selected interviewees to be experienced in cross-border co-operation. These are often on a worldwide scale. Additionally, almost 100 % of the interviewees regularly participate in international conferences and workshops. Depending on the number of conferences on offer in each sub-discipline, the number of conference trips ranged from once every three years to 15 per year, with an average of 2-3 per year. A large number of scholars in my group spent some time overseas. Therefore, the research communities to which the interviewees belong are obviously of a worldwide character. The question "Is your field internationally oriented?" was never answered negatively (but that is only one of several indicators to that effect). I may conclude that at the beginning of the 21st century, science and research is a highly international endeavour where we can expect national differences to play a minor role (at least in the so-called developed world[23]). Consequently, my experts were in a good position to provide information about their field as a whole, i.e. not only about their local perspective, and it seemed unnecessary to extend the list of interviewees to further countries.

See Table A-5 in Annex II for an overview of the characteristics of all interview partners discussed above.

0.3.3.3 Practical aspects and transcription

The 50 interviews took place from November 2001 to February 2002. All interviewees were initially contacted by E-mail (some responded by phone, not by E-mail)[24] and received a one-page information about the overall content and themes in advance of the interview. Among other things, this handout promised full anonymity (see Annex IV). At the beginning of each interview, information on the context of the interview, i.e. the whole project, was given upon request. The interviews lasted between one and a half, and three hours, with an average of two hours. Three interviews were given in English, the rest in German. The same interviewer, i.e. the author of the present study, conducted all interviews. While a small minority took place at the interviewee's home (4), and one via phone, the majority were carried out in the interviewee's office (32) or in a nearby café (13). The choice of location was the interviewees'. The interviews were not taped in order to allow for an open-minded exchange of thoughts. The interviewer took notes on the basis of a semi-structured, open questionnaire (see Annex III). The order of the issues addressed was adapted to the flow of the conversation. Direct quotes were occasionally noted. Immediately after the interview (with a time lag of no more than three days), the notes were transcribed by filling a Microsoft EXCEL spreadsheet. In those cases where the course of the interview departed significantly from the questionnaire, the interviewer filled in the answers by himself on the basis of the notes taken, while indicating with [square brackets] that these answers were not given as a direct reaction on the specific question of the questionnaire, but given in the course of a more general discussion. In

22 See 1.2.3.4 explaining why, in this study, I do not focus on country-specific differences.

23 This is the primary focus of this study on the impact of advanced technology.

24 Contacting them by E-mail was already a first test whether the selected interviewees were indeed experts in terms of Internet use. Note that given the practically universal use of E-mail in academia, this does not lead to a selection bias, as usually criticised in studies targeting the entire population.

case no specific answer could be derived from the notes a slash ("/") indicates the missing answer in the interview database. However, the very large majority of interviews led to concrete and specific answers to each question.

The large majority of the empirical material collected during the interviews formed the basis for chapter 3. A first draft of this chapter was sent out to all interviewees by end of May 2002. In the accompanying letter, they were asked to read at least those sections that summarised the findings for their own field and to give feedback if necessary. Altogether 17 experts replied and at least one expert in each of the 13 disciplines gave feedback. The overall result of this "*member check*" was affirmative. Nevertheless it led to a number of corrections and amendments of the chapter so that we can be confident that the transformation of the statements made in the interviews into chapter 3 was successful. The latter is therefore likely to appropriately represent the status quo of ICT use in those disciplines in 2002.

0.3.4 Practical tests of E-tools

Given my initial research question ("What difference does ICT make in academia?"), the ideal methodology would be to design comparative experiments of the "before/after" type. For instance, one could ask one group of researchers to use the Internet as a research tool and another one to do the same research the traditional way. Both the observation of the research process under these two circumstances and a comparison of the respective results could tell us about the influence of ICT use. Indeed, a few such experiments have been carried out (e.g. O'Hara/Sellen 1997 compared reading online and paper documents; similarly Riehm 1996; Harasim/Winkelmans 1990 studied computer conferencing). Riehm makes a few useful remarks on methodological problems with regard to the empirical comparison of media: (1) preservation of the identity of the contents in both versions (e.g. hypertext and print); (2) task suitability of media; (3) measurement problems; (4) habituation. The most obvious problem with such tests is comparability. On the one hand, while some tasks like, for instance, finding out a particular date of an event, are not dependent on who is doing it (there is only one result), some research depends on the scholar (in particular in the humanities and often in the social sciences). What questions are being asked, how exactly a research question is framed and in what context something (for instance a literary text) is interpreted or understood, is often contingent upon the approach of the individual.[25] Some academic work, even if the topic is identical, would therefore not lead to the same result when done by two different researchers (one with, the other without ICT). On the other hand, the same is even true for a researcher doing the same thing twice, but with different means. Discarding the obvious difficulties of convincing a researcher to do something already done another time, the main problem is that the first time will certainly influence the second time via learning effects.[26]

[25] I am aware of a forceful trend in the humanities towards structuralistic analysis as opposed to the tradition of hermeneutic interpretation. While the latter is more dependent on the subject of the researcher, the former claims to produce inter-subjectively comparable results. By contrast, the constructivistic approach argues that the researcher has an important influence on the outcome of the research.

[26] The final alternative being comparisons across time (how research was carried out a few years ago and today) has the problem that the research subject is not identical. Therefore comparability is equally difficult. Nonetheless, in an indirect manner, this will be my approach in chapter 3. How-

While I therefore opted not to carry out such potentially dubious experiments, I nevertheless tested several innovative tools. The first-hand experience I gathered in these practical tests help me assessing whether or not they have a potential of becoming part of everyday practice in science and research. It should be stressed at this point that these practical tests only amend the other empirical sources described above (accounts in the literature, Internet enquiry and expert interviews). By no means, does the personal experience of the present author in these tests overrule these other sources. However, in those cases where practical experience of innovative applications is still rather limited in the research communities (and consequently among the interviewees), it seems very helpful to add personal experiences gathered with the explicit aim to test the limits and promises of these new technologies.

The following five categories of innovative tools shall be discussed below: E-co-operation through interactive link collections; hypertext writing; E-conferencing; digital brainstorming at distance and chatting; E-publishing and E-archiving.

0.3.4.1 Interactive link collections

Using the Internet for research typically is, above all, searching for data, texts and documents of all sorts as primary or secondary sources. As the Internet has no top-down structure (like, for instance, an encyclopaedia), various tools have been developed with a view to easing access to this wealth of digital information. Among them are link collections. Most of them are rather simple lists of Internet addresses, edited by a single researcher or a tiny group. Others are genuine databases, developed and published by an institute or organisation. Still others are databases as well, but maintained by the research community itself. This author has initiated one such link collection of the latter type in the field of European integration research[27]. The software to administer this link collection has been developed initially for this collection, but has meanwhile been applied in other contexts, too. The software is called EUROLINK[28] and provides for a number of easy-to-use, web-based administrative tools to add and delete links, to structure them in categories and to organise a collective process of updating.

The experience with using this software and the organisation of a bottom-up, community-driven collaboration forms the background of my analysis in chapter 4.

0.3.4.2 Hypertext writing

Hypertexts are not widespread in the academic world. One of the initial hypotheses of this project is, however, that the potential of hypertext writing is excellent and that its use would make a difference in many respects (see chapter 6). To substantiate this claim, the writing of hypertexts in practice was an indispensable (though not sufficient, for sure) experience.

With regard to my topic "knowledge representation", the Internet enquiry (see above) revealed that there is no suitable hypertext authoring software on the market yet. The available tools still seriously distract the writer from the creative process by forcing him/her to simultaneously perform both writing and rather technical activities (such as cod-

ever, I will not present the results from experiments or direct comparisons, but the general assessment of researchers who compare how their discipline worked then and now.

27 <Cyberlink=723>.

28 <Cyberlink=724>.

ing hyperlinks, naming and storing files etc.). In order to be able to nevertheless assess how hypertext writing might influence thinking and outcomes (see 10.4), I co-operated with a programmer trying to emulate future software developments. The result was a simple and easy-to-use hypertext authoring tool with which I was able to test to what extent scientific texts may be modularised. This hypertext editor allows one to generate hypertexts without any knowledge of the involved technology. Basically, it is a set of so-called "macros" which use the Microsoft Word and Windows NT environment to generate text files in the format of the WWW, that is HTML. The little programme has been named "Diskurs".[29] It also allows for co-operative hypertext writing at distance.

Second, this author collaborated with a web-designer to turn one of his early texts on cyberscience (Nentwich 1999a) into a hypertextual multimedia feature (Nentwich/Gagliardi 1999). The former was a traditional linear text, the latter a sophisticated hypertext with animated graphics, video-clips and sound.

The experience with both turning a linear text into a hypertext and with writing a scholarly article in hypertext format with the help of "Diskurs" is documented in 6.4.2.1 and in a conference paper on hypertext writing which is itself written as a hypertext (Nentwich 2000b).

0.3.4.3 E-conferencing

Communicating with other researchers through the Internet may not only take the form of E-mails but can also be a synchronous experience (see 2.4.1). It may well be that the telephone will be replaced or amended with E-conferencing tools based on "web-cams" (tiny cameras connected to a PC) and headsets (microphone and headphones). So far, only very few researchers use these tools. With a view to being able to assess the potential usefulness of these communication media for research, I used Microsoft's well-established NetMeeting[30] software to test how discussing a research issue bilaterally at distance feels. Inter alia, NetMeeting allows not only viewing the other person "live" in a small window on the screen, but also sharing documents and working on them simultaneously. One of the communication partners loads the document in a window on his/her computer screen and can then do one of two things: either s/he lets the other partner(s) view this window (who can then observe what is being done in it on their own screen) or even gives the other partner(s) the opportunity to take control over the remote document and programme. Such "netmeetings" have become a regular event in the weekdays of this author, both for informal conversation and genuine academic discussion – not least about the present text.

Furthermore, the author took part in a panel of a large international conference (Nentwich 2000a)[31], which had several preparatory meetings and three sessions. Two of them took place in advance of the main conference and were held as Internet conferences (where the papers were presented in turn) and one was at the conference proper (in order to discuss the experiences made during the Internet sessions). The preparatory meetings and the Internet panel sessions were realised with the Internet conference host called CentraNow[32]. In this case, we did not use web-cams, as the technical infrastructure of most participants did not allow for it. The CentraNow client provided for audio trans-

29 <Cyberlink=32>.
30 <Cyberlink=725>.
31 <Cyberlink=726>.
32 <Cyberlink=1>.

mission, application sharing, whiteboard and text chat. In this case the paper-giver shared his/her presentation slides with all other participants. Furthermore, the chair was able to "give the floor" and the auditorium to "raise hands" by simple mouse-clicks. The timing of the sessions was carefully selected to bring together, at the same time, researchers from both sides of the Atlantic and from both sides of the North-American continent. The mixed experiences with both tools are discussed in chapter 2.

0.3.4.4 Digital brainstorming at distance and chatting

Many visually oriented researchers use "mind mapping"[33] to structure their first thoughts on a new issue. Traditionally, this is done on paper. Recently, a digital version called MINDMANAGER[34] makes it possible to produce "mind maps" on the screen (with the advantage that the results can be stored, re-used and printed). The latest version of MINDMANAGER allows several people connected via the Internet to brainstorm collaboratively. Furthermore, Internet Relay Chatting (IRC), that is the synchronous exchange of short textual messages, has become widespread outside of the academic world. It may have some future in academia, too. My colleagues and I explored this potential with the software ICQ.[35] My experiences with digital brainstorming at distance and chatting are documented in chapter 2, too.

0.3.4.5 E-publishing and E-archiving

Electronic publishing is at the heart of cyberscience. As discussed in chapters 7 and 8, both in the realm of the economics and of the management of quality control of E-journals, the opinions are split. This author serves as the initiator and editor of an E-journal (in the field of European integration research)[36]. Since 1997, I have gathered invaluable experiences with running an E-journal and organising peer-review over the Net. Furthermore, this journal was a funding member of an international online archive plus search-engine for a growing number of research paper series.[37] Both the development of the software running this archive, the co-ordination of its members, and the management of the co-operation with similar archives add to the practical experience needed for a thorough assessment of the cyberscience developments discussed here. These experiences will inform, inter alia, my analysis in chapters 7 and 8 on E-publishing and quality control.

To summarise, the following instruments (each of them explained in turn in the preceding introduction) *are used in this study of cyberscience*: 50 in-depth interviews in three countries and 13 disciplines, an intensive Internet enquiry leading to a link collection database of 800+ selected entries, a survey of the empirical literature and, finally, a number of practical tests of innovative applications.

[33] Typical end products of such processes are "mind maps" with a structure consisting of a core with several branches and many twigs representing the various facets and aspects of a topic.

[34] <Cyberlink=727>.

[35] <Cyberlink=6>.

[36] <Cyberlink=699>.

[37] <Cyberlink=215>.

0.4 Outline of the book

This report has four main parts:

- Part one: METHODS AND THEORY (Chapters 0 and 1)
  gives an overview of the whole study and the methodological approach (0) and outlines the conceptual framework of this study (1);
- Part two: TECHNOLOGICAL PERSPECTIVES AND STATUS QUO (Chapters 2-3)
  sets the scene by initially presenting the array of new ICT applications relevant to the research communities (2); it then studies how the various (sub-)disciplines deal with the new opportunities and analyses the present differences among the academic disciplines (3);
- Part three: IMPACT ASSESSMENT (Chapter 4-10)
  discusses the actual and potential consequences of ICT use in academia in six main areas: how it impacts on the spatial dimension of academia (4), on roles (5), on knowledge representation (6), on publishing in general (7), on quality control (8); on the economic and legal sphere (9); and finally on the content of research (10);
- Part four: CONCLUSIONS (Chapters 11-12)
  discusses policy options (11) and draws overall conclusions (12).

The remainder of the report includes the bibliography, a list of abbreviations, a glossary of key terms and a number of annexes.