HYPERMEDIA ONLINE PUBLISHING: THE TRANSFORMATION OF THE SCHOLARLY JOURNAL

Andrew Edward Treloar

Bachelor of Arts (Honours) (University of Melbourne)
Master of Arts by thesis (University of Melbourne)
Graduate Diploma in Computing (University of Melbourne)

Thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy
School of Information Management and Systems
Monash University

December, 1998

I Overview & Theory

1 Overview

List of Figures

List of Tables

Abstract

This thesis looks at the impact of the technologies of networking and hypermedia on the scholarly journal. It does so in five main sections. The first section, Overview and Theory, begins by outlining the aims of the study and examining prior related work. Next it defines the three main theoretical perspectives that inform the research (a constructuralist ecology of communication, punctuated equilibrium, and a genre-based framework for new media) as well as considering and rejecting a number of alternatives. The second section, Publishing and Technology, first places the scholarly journal in its historical context and then identifies the stakeholders in the scholarly journal ecology. It then looks at the range of technology developments over the last twenty years that have the potential to be applied to scholarly communication. The third section, Potentials and Responses, looks at the ways in which both publishing functions and stakeholder roles could be transformed and at some of the pressures for such a transformation. It then considers some of the responses that have developed because of these pressures and the potentials of the available technologies. The fourth section, Surveys and Case Studies, presents evidence gathered in this thesis project about users and libraries as key stakeholders. The survey is designed to gather evidence from users about their access to technology, use of electronic publishing, and attitudes to electronic journals. The library case studies look at leading edge examples of libraries who are actively facilitating electronic publishing. The final section, Interpretations and Conclusions, takes the results of all the research activities and discusses them in the context of possible transformations of the roles and practices of stakeholders and the form and function of journals. Evidence from each of the theoretical perspectives, research literature, survey and case studies is brought to bear on each transformation. The concluding chapter discusses the future of the journal as artefact, the possibility of a new technology stasis, whether changes in journals can best be characterised as evolution or revolution, the interlocking systems and interdependencies of the various stakeholders, the archiving dilemma, and the role of technology as enhancer

Statement of originality

I hereby certify that this thesis contains no material that has been accepted for the award of any other degree or diploma in any university or other institution. I further certify that to the best of my knowledge, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signed: Date: / /1998

Acknowledgements

I wish to first of all acknowledge the ongoing support and encouragement of my supervisor, Professor Don Schauder. Endlessly enthusiastic and visionary, he pushed me to do things that he knew at the time (and I realised later) would significantly improve this thesis. I hope that it can stand as a worthy successor to his ground-breaking work [Schauder, 1994a].

My thanks are also due to my Head of School, Professor Andrzej Goscinski, for the way in which he spurred me on to completion and provided me with the study leave to achieve so much. I must also acknowledge my long-suffering colleagues (who covered my teaching and administration while I was away), my students (who have had to deal with my over-enthusiasm for my subject), and Deakin University.

The data collection for the Library Case Studies section was funded by the Victorian Association for Library Automation under their travel grant scheme. Their assistance is greatly appreciated.

No man is an island, and I am no exception. In a very real sense, this is not just my thesis but the work of my extended self, my family. My deep thanks are due to my sons, Mark and Iain, who have put up with extended absences during school holidays, missed birthdays, and a somewhat distracted father when I was working particularly intensely. Hopefully, the return to some semblance of normal family life after four years will not come as too much of a shock.

Lastly, but of course most importantly, my thanks to the other half of me, my wife Dawn. Words alone cannot express the way in which she has enabled me to accomplish this work - in a very real sense her name should also be on the title page.

I Overview & Theory

“I expect that most print-on-paper scientific journals will have disappeared by the end of the decade” Peter Swinnerton-Dyer, 1991

1 Overview

1.1 Introduction

This chapter provides an introduction to the research topic by first introducing the aim of the study in general terms. The structure of the thesis and research design is outlined to enable this topic to be seen relative to the techniques used to investigate. The topic is then located within the wider scholarly field by looking at research in related areas. ‘Related’ will be interpreted fairly broadly to ensure that relevant critical perspectives are brought to bear. Lastly, the contribution of the study to scholarship is discussed.

The intention of this chapter is to provide an overview of the entire thesis (with the exception of the interpretation of results and conclusions) in order to explain the underlying thinking about the topic as well as the rationale for the research program.

1.2 Definitions and scope

The overall domain of interest for this thesis is hypermedia on-line publishing in a scholarly environment. Hypermedia is broadly defined as the combination of hypertext linking mechanisms and multimedia content [Bolter, 1991, pp. 25-7].

Hypertext can be summarised as the ability to link two pieces of text together and traverse that link in some way. In a sense, hypertext has always been with us: the well-established system of scholarly citations is a form of hypertext, defining as it does a mapping from one document to another. With conventional citations the traversal process is far from instantaneous and often involves an inter-library loan request! In the context of current computer systems, the most obvious application of hypertext is the familiar blue underlined link within a World-Wide Web document. A fuller treatment of hypertext will be provided later (see 4.4.1: Hypertext on page 67).

Multimedia is the provision of information in more than one medium. Medium is a term that is used somewhat loosely in the literature. One use is the form of the information: text, still image, moving image, or sound. A different use of the term is the mechanism used to deliver that information: paper, screen, audio or video tape, audio or video disc, and CD-ROM. These uses are somewhat interrelated; there are only certain valid combinations of information form and delivery mechanism.

In a scholarly publishing context this hypermedia publishing will appear as any form of scholarly publishing that provides text and static graphics (both of which can be done already with print) together with some or all of the following:

- hypertext links both within the document and between documents
- links to video, sound, and simulations (or as embedded content)
- ongoing peer commentary
- support for ready communication with the author(s)
- links to underlying data sets

None of these additional features can be provided in a paper-based publishing environment.

The on-line delivery mechanism for this publishing is assumed to be via the Internet or an organisational intranet. It is true that the techniques involved are mostly applicable to CD-ROM or floppy disk delivery, but the communication mechanisms require a networked medium. In practice, all the current e-journals assume the use of the Internet for content delivery.

Electronic scholarly publishing normally takes place in the context of a scholarly journal, but may also appear in conference proceedings, an on-line book, as part of course materials or on a Web site as 'vanity' publishing. This thesis will primarily be concerned with scholarly journals.

1.3 Aims of the study

The programme of study for this thesis aimed to examine the transforming potentials of the new electronic publishing technologies and their effect on the traditional scholarly journal. The emphasis was on the ways in which the potentials of these new technologies will (or may) become realised by the components of the scholarly communication industry. While the technology is an important part of this study, it is not the most important. That place belongs to the potential changes that the new technologies allow and their adoption by the various participants and stakeholders in the ecology of scholarly publishing. By using these technologies, it is possible for both all elements in the ecology to change themselves (and in turn, each other). This study attempted to look at what has happened to date and predict what might happen in the future. The design of the study is grounded in a model of an ecology of communicative transactions. (see 2.2.4: An ecology of communicative transactions on page 28). This model was then applied to the specific ecology that is scholarly journal publishing.

1.4 Research questions

The technologies for communication and computing are becoming increasingly more sophisticated and allowing a greater range of possiblities.Within the ecology of scholarly journal publishing, the new technologies enable a range of transformations. In order to focus the aims of this study, the research concentrated on transformations of the communications artefacts (their form and function) and of the stakeholders (their professional practices and roles).

1.4.1 Artefacts

The communications artefacts are chiefly articles and the journals that contain them. The interaction between form and function is quite a complex one. Form and function influence (but rarely determine) each other. As an example particular journal forms (such as richly hypertextual articles) predispose their readers to particular activities (in this case ready following of links to related material) in ways that alternative forms may not. The specific research questions relating to the artefacts in this thesis were:

- What transformations (if any) will occur in the form of the scholarly journal as it moves into an online hypermedia environment?
- Will the move into an online hypermedia environment alter the functions of the scholarly journal?

1.4.2 Stakeholders

The main stakeholders in this ecology are scholars (as both writers and readers), scholarly societies (as both representatives of scholars and publishing intermediaries), publishers (both for-profit and non-profit), subscription vendors and librarians. Roles and practices are also intertwined. For instance, one should not assume that an organisation with a publishing role could not practice an archiving function.

The specific research questions relating to the roles and practices of stakeholders in the current scholarly journal ecology were:

- How might technology transform the roles of these stakeholders?
- How might technology transform the practices of these stakeholders

1.5 Research design

1.5.1 Design questions

The nature of the questions for this thesis and the theoretical framework that has been used to discuss them place this research squarely within the domain of the social sciences. This should not be surprising: scholarly journals are above all communications artefacts, and communication is a social activity. What then are the possible choices of research methodologies available to social science researchers?

The social science research literature (for instance [Bickman and Rog, 1998], [Allen and Skinner, 1991], [Adams and Schvaneveldt, 1985]), reveals a number of possible choices in both research type and data collection strategies. These are partly related - certain types of research fit better with particular data collection strategies. Possible research types include exploratory and descriptive research (including case studies), field studies and field experiments, and experimental-causal investigations. Available data collection strategies include questionnaires, interviews, direct observation, projective and indirect methods, archival documents and evaluation research.

The challenge for the researcher is thus to match the most appropriate research type and data collection strategies with the foci of their research.

In looking at the research questions for this thesis, a number of possible data collection targets emerged. The artefact questions involved issues to do with the preconditions for change in artefacts (access to suitable technology and current use of existing technology) and early signs of change (new journals and journal forms). The stakeholder questions involved issues to do with both individual and organisational responses. What is an appropriate set of research design decisions? Table 1–1 summarizes the research foci and sub-foci and the corresponding data collection decisions. The next section discusses the reasons behind these decisions.

1.5.2 Research type

An early decision was to situate this thesis within the domain of exploratory and descriptive research, rather than that of field studies and field experiments or experimental-causal research. This was because what underlay the research questions was an attempt to gauge trends on a large scale and to examine merging specific initiatives. Such an exercise lent it-

Table 1–1: Research Questions and Data Collection Strategies

Research Foci	Sub-foci	Collection focus	Collection technique
Transformation of journal form and function	Reader's access to necessary technology	Availability of technology; Comparison between groups in same discipline area	Questionnaire
	Reader's existing use of such technology	Usage patterns; Comparison between groups in same discipline area	Questionnaire
	Reader's attitudes to electronic publishing	Comparison between groups in same discipline area	Questionnaire
	Early indicators of change	Specific new electronic-only journals	Literature/Web survey
Transformation of stakeholder roles and practices	Institutional responses:	Case study on specific initiatives	Semi-structured interview
	Reader's responses	Use of electronic publishing technologies	Questionnaire

self better to descriptive research. Much of this descriptive work was based on survey results, but some relied on case study work to do with specific library initiatives. Despite the overall descriptive focus, the work to do with establishing relationships between variables in the survey section (see 7.3: Survey Results on page 141) should be viewed as analytical [Jones, 1991].

Carrying out field studies on a global scale would have been both extremely difficult to arrange and very expensive. It would be very hard to recruit other sites and ensuring consistency in data collection would pose problems. Experimental/causal research would also face significant practical difficulties if carried out at multiple sites. In addition, it would be necessary to identify a possible type of experiment that would yield useful results. The author is unaware of any such studies either completed or planned.

1.5.3 Data collection

Once the decision had been made about the broad research type, it was necessary to select appropriate data collection techniques. It was decided to primarily use questionnaires and interviews. What are the pros and cons of these techniques, and what other techniques might have been used?

Questionnaires

The term questionnaires can cover a fairly broad range of possible research techniques. For the purposes of this research it means structured surveys delivered remotely and completed without the researcher's input or presence. Such surveys have a number of advantages [Mangione, 1998]:

- relatively low cost
- quick coverage of large numbers of dispersed respondents
- convenience and privacy for respondents
- insulation of respondents from interviewer expectations

These were all seen as advantages for this research topic. Questions of cost and coverage were particularly relevant, but it was seen as important for the attitudes of the researcher not to influence the answers of the respondents.

Surveys are a particularly good choice when:

> "(a) you have limited human resources to help you conduct your study, (b) your questions are written in a closed-ended style, (c) your research sample has a moderate to high investment in the topic, and (d) your list of research objectives is modest in length." [Mangione, 1998, pp. 399-400].

All of these conditions obtained for this research. The resources for the study were largely limited to the researcher. The questions (see 7.2: Survey process on page 127) were mostly closed, with a few opportunities for additions or freeform comments. Both research sub-samples were judged to have at least a moderate interest in the topic. The list of research objectics was capable of being fulfilled with a two-page (albeit information-dense) survey instrument.

Questionnaires do have some areas where problems may occur, affecting the quality of the data and therefore the conclusions that can be drawn. The surveys for this thesis sought to minimise such problems by practising "total survey design" [Biemer et al., 1991]. The four broad error types to be avoided are sample selection bias, nonresponse error, item nonresponse error and response error. The ways in which the surveys in this study sought to avoid these errors are discussed in detail in the chapter dealing with the surveys and their results.

Interviews

Interviews were one of the main techniques used in the case-study phase of the research (see 8: Library Case Studies on page 177). These interviews were with senior staff in the library projects selected and followed a semi-structured pattern. Such semi-structured interviews are an appropriate technique for the conduct of descriptive exploratory research in a new area of enquiry [Bickman et al., 1998]. This is because they provide a consistent framework for the questions being asked while still allowing the flexibility to follow particular lines of enquiry if this seems relevant or appropriate. Because of the open-ended nature of the interviews, the respondents need to be considered as informants only, providing verbal reports about their version of events. Where possible, these verbal reports should be checked by triangulation with other informants and by corroboration with other sources of evidence (i.e. documentary evidence) [Yin, 1998]. The specific interview questions are included at the end of this thesis (see 11.3: Library Case-Study Questions on page 243).

Other techniques

Other possible techniques that could have been adopted include direct observation, documents of the past and evaluation research. Direct observation and evaluation research were not used because of the problems of limited human resources and costs. Documentary research was used as part of the case-study phase, primarily to provide background information on the projects and to corroborate information provided by informants.

1.6 Research implementation

This thesis sought to provide an answer to the questions by focusing on the complex interactions within this ecology from two main perspectives: scholars and libraries.

1.6.1 Scholars

The focus of the research dealing with scholars was on their attitudes to electronic scholarly publishing, together with their access to and use of electronic publishing technology. In contrast to some earlier research [Schauder, 1994], this thesis concentrated on scholars in a single discipline – psychology. Psychology is a good test domain because it straddles the border between the social and ‘hard’ sciences, but does not have highly demanding formatting requirements for its publications. This is in contrast to disciplines like physics and mathematics which rely on precise layout of mathematical equations and figures.

Two groups of scholars were selected, both working in the domain of psychology. The first group consists of the subscribers to an email list of readers of an electronic only e-journal (*Psyche - an Interdisciplinary Journal of Consciousness and Cognition*). The second group was chosen to complement the first and consisted of members of psychological societies in the U.S., U. K. and Australia (the main countries represented in the first group). The first group was surveyed by electronic mail. The second group was surveyed via conventional mail. The two groups were chosen from the same discipline area to provide a comparison between those who could be expected to be comfortable with (and supportive of) the new publishing technologies, and the wider population within their discipline for whom no such assumption was possible

The survey instrument asked the same questions in each case. First were questions about industry category and employment position to gain a sense of the demographic distribution. Next were questions about the respondents access to technology to determine what electronic publishing features they would be able to access. Then it asked how often they used particular electronic publishing fora, both as readers and authors. Lastly it requested their ratings of a number of suggested advantages and disadvantages commonly applied to electronically published scholarly articles.

The intention of this phase of the research was to identify what respondents were doing, what they had access to (in terms of technology) and what they felt about the new scholarly publishing possibilities (see 8: Library Case Studies on page 177). The results are therefore relevant primarily to an analysis of the form and function of scholarly communication artefacts (as used by two different sample populations of scholars), and secondarily to the roles and practices of users as stakeholders.

1.6.2 Libraries

The focus of the research dealing with libraries was on their potential new role as publishers. This phase of the research took the form of a series of in-depth case-studies conducted at five libraries in the U.K. and Europe. Each library was jointly or solely responsible for a significant and innovative electronic publishing project. The research methodology consisted of careful selection of five very different libraries, investigation of publicly available materials about their activities (both published articles and Web-sites), and a site visit combined with interviews. The interview questions covered definitions of publishing, questions of commercial sustainability, the role of libraries, the origins of the particular project, the critical factors for its success and a SWOT (Strengths/Weaknesses/Opportunities/Threats) section.

The intention of this phase of the research was to get a sense of how a range of libraries were approaching the challenges of the new technologies and rethinking the role they could play in the system of scholarly communication (see 8: Library Case Studies on page 177). The results are therefore primarily relevant to an analysis of the roles and practices of one key group of stakeholders and secondarily to the form and function of scholarly communication artefacts as published by these stakeholders.

1.7 Structure of the thesis

This thesis is a combination of a number of elements that together tell a story. More accurately, they tell elements of the same story from different perspectives, somewhat like Akira Kurosawa's *Rashomon*. The story revolves around the potentials of the new technologies and the transformations they make possible.

1.7.1 Theoretical setting

The first element is the theoretical setting for the research program. It is important to situate a piece of sociocultural research like this within a theoretical framework. Chapter 2 discusses in detail the three frameworks used (Constructuralism, Punctuated Equilibrium and Genre theory) and addresses briefly some other candidate frameworks (see 2: Theoretical Perspectives on page 22).

1.7.2 Historical perspective

The second element is the importance of understanding what has preceded the current state of the art. In the case of hypermedia electronic publishing, this requires an understanding of three different (but interrelated areas). The first is the print publishing of scholarly journals. This has been driven both by the needs of scholars themselves and by the availability of low-cost publication and distribution mechanisms. Chapter 3 provides a summary of the development of this area, and raises some of the issues associated with print publishing that scholars are beginning to address (see 3: Print Publishing of Scholarly Journals on page 46). Technology developments in computers and communication since the Second World War have made a range of new things possible both in publication and distribution. Chapter 4 takes up the story of the technologies, but without getting too enmeshed in the technical details (see 4: Technology Developments on page 57). The technology provides an essential background to the story and so needs careful explanation. Chapter 5 considers the potentials inherent in the technologies and their possible impact (see 5: Potentials and Pressures for

Transformation on page 91). The responses by the scholarly publishing community to these possibilities have been somewhat muted to date, but have explored much of what is possible. Chapter 6 provides an account of these responses (see 6: Developing Responses on page 107).

1.7.3 Surveys and case studies

The third element comprises the survey and case study work undertaken towards this thesis (as opposed to the interpretation and synthesis of others' work). This research is in two parts. The first is a pair of surveys (one email and one print) exploring the attitudes of scholars to electronic publishing. Chapter 7 discusses the results of these surveys, and contrasts the two different patterns of responses (see 7: Surveys on page 127). Chapter 8 discusses a complementary piece of research: an in-depth series of case-studies looking at the role of libraries as publishers in the new scholarly publishing environment (see 8: Library Case Studies on page 177). Chapter 9 discusses the results of the surveys and case studies and draws on related work by other scholars where this is appropriate (see 9: Interpretation of findings on page 200). Chapter 10 tries to draw the threads together and look into the future (see 10: Conclusions on page 231).

1.8 Related work

A number of other researchers have examined the issue of scholarly serials in on-line electronic form from a range of perspectives.

Many of these articles and conference papers deal with general issues that affect electronic scholarly publishing and are more of the nature of reflections on themes or discussion of future of possibilities. The articles will be discussed and referenced in the chapters dealing with potentials (see 5: Potentials and Pressures for Transformation on page 91) and responses (see 6: Developing Responses on page 107).

Because this thesis is based on surveys and case-studies, it seemed best to focus on related research with an explicit data collection focus. This section will outline this related research here while in general not dealing with any results. The results will be considered in detail in the context of the results from the surveys and library case studies performed for this thesis (see 9: Interpretation of findings on page 200).

The relevant related research can be categorised into the following categories:

- analysis of journals in the context of tenure and reward structures
- citation analyses
- attitudinal studies
- single e-journal readership studies
- surveys of a range of e-journals
- library studies

1.8.1 E-journals and tenure and reward structures

These studies have looked at e-journals from the perspective of how they fit into the existing tenure and reward system for academics. This is an important, and often neglected, aspect of the way in which journal publication is embedded into the discipline of scholarship.

The work of Julene Butler has sought to address the related research questions of whether publication in electronic journals brings the expected rewards of research publication, and whether the electronic journal is a viable channel for formal scholarly communication. They are related because scholars will only publish electronically if they are rewarded in some way for doing so, and the viability of the channel is dependent on scholars choosing to publish.

Butler uses a sociology of science approach and operates from the assumption that

> the reward processes which sociology has observed in traditional knowledge production processes must be operational in any new scholarly communication channel if that channel is to be adopted by members of the scholarly community [Butler, 1994, p. 59].

Butler's research has been reported in a number of different publications. [Butler, 1994] outlines her research methodology without presenting any specific results. [Butler, 1995b] and [Butler, 1995a] are (as indicated in the text of the latter) essentially identical and report the results of her survey distributed to 511 authors and members of editorial boards of ten e-journals in the science and social science disciplines. Seventy percent of the surveys were distributed electronically with the remaining 30 percent via surface mail. Her survey questions dealt with academic involvement and educational level of contributors, tenure status and academic rank, perceived benefits and disadvantages of electronic journals, feedback received and how electronic publication was perceived by their superiors.

A brief paper by Blaise Cronin and K. Overfeldt [Cronin and Overfelt, 1995] discusses a small survey of universities regarding their policies with respect to electronic publishing.

Very little policy was supplied by these universities regarding refereed e-journals. A range of unsolicited comments indicated that there are inconsistencies in interpretation and practice, both within and across institutions.

1.8.2 Citation studies

Stephen Harter and Hak Joon Kim have been funded by the Online Computer Library Center, Inc. (OCLC) to perform a study of the characteristics of e-journals as measured by number of articles published, frequency of publication, administrative policies, access issues, and references. The main stated purpose of their research was to "assess the impact of scholarly and peer-reviewed (or refereed e-journals) by studying citations to, and references in, e-journal articles as of the latter part of 1995" [Harter and Kim, 1996a, p. 442].

A major component of their work published so far has been an analysis of the citations to and in e-journals. This research was carried out in the latter part of 1995. Two published directories of electronic journals were used to select all peer-reviewed or refereed e-journals available at that time for a total of 131 e-journals. This sample was used for an access and demographic study.

The sample was narrowed to e-journals containing articles that reported the results of research or scholarship which resulted in a sample of 77 scholarly e-journals used for a reference study [Harter, 1996] which examined the patterns of citations to e-journals to identify the most cited e-journals. This study also compared citation and publication data for e-journals in the study with print journals in the same fields. Because of a range of factors, citation counts are not directly comparable between e-journals and print journals analysed in this study. The ISI impact factor was used instead. Because there is in general a direct linear relation between journal production and the impact factor [Rousseau and Hooydonk, 1996], journals with low publication rates will have a lower impact factor. This study found that despite the small number of articles published to date by e-journals, *Online Journal of Current Clinical Trials* (OJCCT), *Public-Access Computer Systems Review* (PACS-R), and *Psycholoquy* all ranked highly in impact factor relative to print journals in their field. The overall conclusion was that the overall impact of e-journals would not increase until publication rates also increased. This requires that authors regard publication in such journals as a legitimate and rewarded activity.

A further study focused on the rate at which e-journal articles referenced other journal articles, as well as looking at problems in accessing such referenced articles from the point of

view of users [Harter and Kim, 1996b], [Harter and Kim, 1996a]. To allow time for scholars to have read e-journal articles the sample was restricted to 39 e-journals that had commenced publication in 1993 or earlier. The study found very low citation rates from e-journal articles to other e-journal articles, a bias towards a few influential titles, and significant problems in accessing cited articles.

The latest extension of this work has examined the extent to which scholars and researchers are aware of or building their own work on research that is published in e-journals [Harter, 1998]. This study sampled scholarly peer-reviewed e-journals and conducted a range of citation analyses. The data showed that the impact of e-journals on scholarly communication has been minimal to date.

1.8.3 General attitudinal studies

Some researchers have examined attitudes to e-journals themselves, either as the main focus of their research or as a component.

F. W. Lancaster [Lancaster, 1995a] distributed a survey in November 1993 to 309 administrators (of which 150 responded) of universities whose libraries are members of the Association of Research Libraries (ARL). The recipients were either library directors or administrators with responsibilities for academic research. The questions related to possible advantages of networked electronic scholarly publishing over existing procedures, factors affecting the implementation of such a networked system, and a ranking of the priority to be assigned to developing such a networked approach to publishing research articles relative to a range of other academic priorities.

Ann Bishop examined the features provided by 7 e-journals [Bishop, 1995] Her analysis is based on her reactions and opinions alone and did not involve any data collection.

Finally, a comprehensive study by Don Schauder examined attitudes of an international sample of academics to electronic scholarly publishing [Schauder, 1994]. A total of 2,229 questionnaires were distributed and 582 returned. The 12-page survey instrument asked respondents to answer questions in the following categories:

- demographics
- behaviours when reading and writing professional articles
- the role of the journal and their attitudes to the current journal system
- use of information technology in professional work

- universities attitudes to and behaviour towards academic publishing

1.8.4 Single e-journal readership survey

As indicated in [Berge and Collins, 1996], there has been very little survey work on e-journal readership. This section of the thesis reviews known survey work on single e-journal titles.

A study of the readership of *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century* (IPCT Journal) gathered information on general attitudes to electronic publishing, preferred delivery formats and storage media, the level of acceptance of e-journals for promotion and tenure, and the demographics of the readership [Collins and Berge, 1994], [Berge and Collins, 1996]. The survey was posted to the IPCT Journal subscription list (N=1604 at the time of the survey) and the IPCT Discussion List (N=1118 at the time of the survey). There was considerable overlap in membership of the two lists. The response rate was around 18%.

[Gotsch and Reich, 1997] performed a series of studies related to the *Journal of Biological Chemistry (JBC)*, published by the American Society for Biochemistry and Molecular Biology) ASBMB) and Highwire Press. JBC is a very high-prestige journal and at the leading edge of parallel electronic/print delivery. These surveys (which took place in late 1996) consisted of:

- A user survey instrument distributed in two rounds: first to 250 members of the Editorial Board (60% response rate), and secondly to a sample of 150 individual subscribers to the online version of JBC (50% response rate)
- An author survey sent to a sample of 500 randomly selected authors who published in JBC in 1996 (48% response rate)
- A survey of 300 institutional subscribers to JBC online (20% response rate).

The user surveys asked which version of JBC was used (print or online), reasons for non-use of the online version, advantages of each, responses to discontinuing the print version, and a ranking of various features of the online version. The author survey concentrated more explicitly on the print versus on-line issue, and also explored whether moving to online only would affect their decision to publish in JBC. The survey of institutional subscribers focused on the process of selecting print or online subscriptions, current subscription assessment practices, assessment of the worth of online subscriptions and reservations about moving to online.

[Woolfrey, 1995] discusses the results of a questionnaire sent in March 1994 to 427 subscribers to the *Canadian Journal of Communication*(CJC) which received a 21% response rate. The questions dealt with interest in receiving electronic copy, their willingness to pay for the added service (electronic as well as print), their reasons for wanting electronic copy and for wanting to search electronic texts, their electronic capabilities, and the other types of electronic information they would want CJC to provide.

1.8.5 Survey of e-journals

[Hitchcock, 1996] reports on a survey of all (with a few restrictions) the online STM (Science, Technology and Medicine) journals as of September/October 1995. The survey considered the novel features of both electronic-only and parallel-published journals, and the funding and commercial status of the 115 journals accessed. An updated version of this survey takes another snapshot as of October 1997 but this time focuses on UK journals (defined as having editorial offices and a presence in the UK) only.

1.8.6 Digital library studies

This category deals mostly with studies associated with digital library projects where there was a component that surveyed use of e-journals. The exception is the Mellon report which looked at the issues facing libraries with respect to serials in general. Because the projects need to have been completed before the results can be analysed, the projects are largely first-generation or ‘early adopter’ e-journal projects.

Mellon report

This was a study prepared for the Andrew W. Mellon Foundation and completed in 1994. The study, outlined in [Mel, 1992] and summarised in [Okerson, 1996], used Association of Research Libraries (ARL) statistics to study 24 major US research libraries. The study examined acquisitions expenditure and patterns, increases in serials prices, and possible technological solutions.

Core Project

The Chemistry Online Retrieval Experiment (CORE) project was carried out at Cornell University from 1990 through 1995. It was a collaborative effort of Bellcore (Bell Communications Research), Mann Library (Cornell), the American Chemical Society and the OCLC Online Computer Library Center, Inc. From the perspective of the Mann library, the primary

goal of CORE was to “create a system with a substantial depth and breadth within a single discipline (chemistry) - a critical mass” [Entlich, 1995, p. 114]. CORE ended up processing over 400,000 page images from twenty ACS journals. The product of this processing was both bitmap and SGML versions of the pages together with hypertext citation links, searching and support for graphics and special symbols. In fact, the average article included:

- text from two different sources
- two full-page bitmaps/page (at 100 and 300 dpi)
- each figure in a separate file
- thumbnail reductions of the figures
- equations in a separate file

CORE included a range of investigations carried out as part of the project. Three studies in particular are relevant to this thesis.

The first was intended to examine the interaction between scholars and the journal literature [Olsen, 1994]. The broad questions were:

> What are the scholars’ purposes in using the literature and what are the actual processes carried out by scholars with journal literature which must be accommodated by the electronic version of journals? [Olsen, 1994, p. 7].

This study was based on interviews with forty eight scholars (sixteen each from chemistry, sociology and English) based at Cornell University and the University of Pennsylvania. These were not scholars who had been exposed to the CORE system and had in fact never previously used electronic journals. The questions dealt with:

- reasons and techniques for locating journal literature
- methods of reading this literature
- when and where the reading too place
- which literature was most useful
- the advantages and disadvantages of printed journal literature compared with the electronic journal
- general demographics
- computer use.

The second study focused on users’ interaction with the e-journal interface [Stewart, 1996]. The data was derived from “open-ended, hour-long interviews with thirty-nine (sic.) users of the Chemistry Online Retrieval Experiment (CORE) at Cornell University” [Stewart, 1996,

p. 340]. Unfortunately only eight of the thirty nine subjects were faculty members - the bulk of the remainder were graduate students; this can be assumed to have affected both the age profile and response pattern. The questions used built on the system functions identified by scholars in the earlier research [Olsen, 1994] as essential in an electronic system. Subjects were asked to evaluate the importance of features for selecting optimal research literature and reading that literature. They were also asked about selected characteristics of an ideal system and impact of an electronic system on reading productivity.

The third study [Entlich et al., 1996] examined the system logs and used online questionnaires to supplement the interview data reported in [Stewart, 1996]. Of the 161 accounts distributed for the system only 75 came into active use. Comparison of the logs with the fact to face interviews enabled the researchers to compare the users stated desires and intentions with their actual behaviour. The online questionnaires were introduced during the final six months of the trial and consisted of four randomly selected multiple choice questions presented at the end of a user's session.

Project ELVYN

Project ELVYN (allegedly from ELectronic Versions - whY Not) was an initiative to take a specific new print journal, *Modelling and Simulation in Materials Science and Engineering* (MSMSE) from the Institute of Physics Publishing and deliver it electronically [McKnight et al., 1994], [Rowland et al., 1995]. The project was particularly concerned with the problems of transmitting a journal from publisher to reader and the costs associated with this [Meadows et al., 1995]. No journal useability studies were undertaken.

TULIP

The University Licensing Program (TULIP) started in early 1991 and concluded at the end of 1995 [Elsevier Science, 1996]. Its participants were Elsevier Science and nine US universities: Carnegie Mellon, Cornell, Georgia Institute of Technology, M.I.T., University of California, University of Michigan, University of Tennessee, University of Washington, and Virginia Polytechnic and State University (Virginia Tech.). The goal was to test systems for networked delivery of journal content to the user's desktop as well as the use of these systems. Elsevier delivered the content (scanned page images, dirty OCR, and raw ASCII full text) to the universities who then or adapted their own systems to manage and deliver this content. The focus of the research was on "technical issues, on user behaviour and on organization and economic questions" [Elsevier Science, 1996, p. 7].

The user behaviour research was both quantitative and qualitative. The quantitative work was based on logfile analysis from the sites. The qualitative research was based on four focus groups with around ten graduate students each and one to one interviews with twenty six university faculty. The questions dealt with:

- current information needs patterns
- access to and use of electronic information
- a detailed evaluation of TULIP
- professional journal usage.

ARL SPEC Survey

Early in 1994 a Systems and Procedures Exchange Center (SPEC) survey was sent to 119 members of the Association of Research Libraries (ARL) in the U.S. The study aimed to determine “the policies and practices of ARL libraries in selecting, acquiring and providing access to electronic journals” [Parang and Saunders, 1994a, p. 1]. A total of 77 questionnaires were returned (65% response rate). The survey also sought to identify issues and trends in the ways ARL libraries were planning to deal with electronic journals [Parang and Saunders, 1994b].

1.9 Conclusion

This chapter outlined the overall aims of this thesis and defined its scope. It identified possible relevant research which will be considered in the context of the thesis results. While there has been some research on attitudes to electronic scholarly publishing and on the readership of individual e-journals, there has been no research comparing practices and attitudes between two sub-groups within a scholarly discipline where one of the sub-groups consists of known e-journal readers and the other sub-group is drawn from the base population of that discipline. The next chapter describes the theoretical basis for this research.

1.10 Previous Publications

Material related to this thesis has already been published in conference proceedings and journals. In accordance with Monash University regulations, and as evidence of peer accountability during the course of the research, those publications are listed here.

1.10.1Refereed

Treloar, A. (1998). Evolving ecological niches: technological change and the transformation of the libraries role in publishing. *Proceedings of the Second ICCC/IFIP Electronic Publishing Conference*, Budapest, April. Available online at <http://www.deakin.edu.au/~aet/Research/Publications/IFIP98/>

Treloar, A. (1998). Six case studies of libraries as facilitators for electronic scholarly publishing. *Communications of the ACM*, 41(4):88-89. Available online at <http://www.deakin.edu.au/~aet/Research/Publications/CACM/>

Treloar, A. (1998). Technology as agent for transformation: five case studies of university libraries as facilitators for electronic scholarly publishing. *Proceedings of The Victorian Association for Library Automation Biennial Conference (VALA '98)*, Melbourne, January. Available online at <http://www.deakin.edu.au/~aet/Research/Publications/VALA98/>

Treloar, A. et. al. (1997). Aftershocks of an electronic publishing experiment: areview of the EPICentre project. *Proceedings of Information Online and OnDisc 97*, Sydney, Australia, January:115-132.

Cameron, J., Denison, T., Jansen, B., Mercieca, P., Reekie, P., Schauder, D., Treloar, A. and Wilken, P. (1997). EPICentre - the electronic publishing innovations centre. *Ejournal*, 7(1). Available online at <http://bond.edu.au/Bond/Library/People/jpr/ejournal/epi.html>.

Treloar, A. (1996). Better than print? Hypermedia scholarly publishing and the world wide web. *Proceedings of The Victorian Association for Library Automation Bi-ennial Conference (VALA '96)*, Melbourne. Available online at <http://www.deakin.edu.au/~aet/Research/Publications/VALA96/>.

Treloar, A. (1995). Scholarly publishing and the fluid world wide web. *Proceedings of AUUG '95 & Asia-Pacific World Wide Web '95 Conference*, Charles Sturt University. Available online at <http://www.deakin.edu.au/~aet/Research/Publications/APWWW95/>.

Treloar, A. (1995). Scholarly electronic publishing and the world-wide web. *Proceedings of AusWeb95 (the first Australian WWW conference)*, Southern Cross Uni-

versity. Available online at <http://www.deakin.edu.au/~aet/Research/Publications/AusWeb95/>.

1.10.2Non-refereed

Treloar, A. (1994). Working in a worldwide invisible college: the internet and the global scholar. *Proceedings of CAUSE `94*, Melbourne.

Treloar, A. and Schauder, D. (1997). Psyche: case study of attitudes and access to an ejournal. *Proceedings of CAUSE `97*, Melbourne.

2 Theoretical Perspectives

2.1 Introduction

The previous chapter discussed the overall aims of this thesis and the specific research questions that would be addressed. These questions can be characterised as broadly relating to the transformation of the:

- *form* of the scholarly journal
- *functions* of the scholarly journal
- and the *environment* within which the scholarly journal operates.

These questions cannot be answered in a vacuum. All scholarly endeavour takes place in the context of a range of possible theoretical perspectives. Before considering the research questions, it is important to consider relevant theoretical work and how this might inform the research and its conclusions. Whatever theoretical framework is chosen needs to relate directly to the research questions, but does not need to explicitly reference the subject domain of the thesis.

This chapter identifies the most appropriate theoretical frameworks and their application to the subject of the thesis. Each candidate theory is introduced together with the main definitions, concepts or axioms. Next the chapter focuses on their general areas of application. Lastly their specific relevance to the research questions is drawn out. Having demonstrated the particular relevance of these theoretical approaches, other possible frameworks are briefly discussed.

2.2 Constructuralist ecology of communication

This is based on the work of David Kaufer and Kathleen Carley ([Kaufer and Carley, 1993], [Kaufer and Carley, 1994], [Carley, 1995]). The strength of their work lies in their rigorous treatment of communication in general and written communication in particular. Its weakness is that the book which encapsulates the fullest working out of their ideas [Kaufer and Carley, 1993] was completed just as the Internet and its associated technologies were coming to prominence. This gives this particular work a somewhat dated feel.

[Kaufer and Carley, 1993] deals with the advent of print and its influence on the world of communication we inhabit. It starts from the fairly standard theoretical model of partners in

communicative transactions exchanging communications with them but then moves on to focus on the entire interaction cycle of communication as the unit of analysis. This shifts the focus from particular elements in the communicative transaction to a single communication *ecology*.

[Kaufer and Carley, 1994] formalizes the discussion of communication concepts in terms of a series of concepts and axioms about both proximate communication and communication at a distance.

[Carley, 1995] moves beyond print technologies and starts to suggest some tentative ways in which this analysis might deal with some of the new communication technologies, without actually discussing any of those technologies. This work still does not address the new hypermedia publishing technologies which lie at the centre of this thesis.

Running through each of these texts are elements from three clusters of themes:

- concepts and axioms relating to communication at a distance
- an open systems ecology of communicative transactions
- application of the theory to particular domains

2.2.1 Initial assumptions

[Carley, 1995] provides the most complete list of the assumptions behind a 'constructural' theory of communication. The assumptions start in the most general way and gradually become more specific. Collected together, they can be summarised as shown in Table2–1.

The focus of Carley's work is communication and so the technologies that alter the ways in which agents can process information are communications technologies (broadly defined). The salient and shared features of many such technologies are that they enable either mass-communication or one-to-many communication, and that they allow for unchanging and intact communication through space and time.

By examining the impact of these technologies, Carley then identifies three stylised classes of agents [Carley, 1995]:

- People, who are "capable of only one-to-one communication, can select interaction partners, and can learn" (p. 553)
- Books, who are "capable of one-to-one or one-to-many communication, cannot select interaction partners, and cannot learn" (p. 553). The term book as used by Carley is shorthand for any written material that is both printed and mass-circulated, and is

Table 2–1: Assumptions behind a constructural theory of communication. Source: Based on [Carley, 1995]

1.			Individuals are continuously engaged in acquiring and communicating information
	a.		Individuals, when interacting with other individuals, can communicate information
	b.		Individuals, when interacting with other individuals, can acquire information
	c.		Individuals can learn the newly acquired information, thus augmenting their store of knowledge
2.			What individuals know influences their choices of interaction partners
	a.		Individuals select interaction partners on the basis of relative similarity and availability
	b.		Individuals engage in action concurrently
3.			An individual's behaviour is a function of his or her current knowledge
	a.		Individuals have both an information-processing capability and knowledge, which jointly determine each individual's behaviour
	b.		Individuals have the same information-processing capabilities
		i.	All agents within a particular class have the same information-processing capabilities
	c.		Individuals differ in knowledge as each individual's knowledge depends on his or her particular sociocultural and historical background
		i.	Agents differ in knowledge as each agent's knowledge depends on the agent's particular sociocultural-historical background and information-processing capabilities
	d.		Individuals can be divided into types or classes on the basis of extant knowledge differences
		i.	Agents can be divided into types or classes on the basis of extant differences in either or both their information-processing capabilities and knowledge
4.			Technology can alter the information-processing capabilities of existing agents
5.			Technology can, though need not, create artificial agents with unique information-processing capabilities

presumably generalisable to 'fixed' electronic text as well.

- Orators, who are "capable of one-to-one or one-to-many communication, can select interaction partners, and can learn" (p. 553). People may obviously also function as orators.

2.2.2 Concepts

Moving beyond the underlying assumptions, a set of key concepts underpin all Kaufer and Carley's discussion of written communication, whether proximate or at a distance. These building blocks are *relative similarity*, *mental models*, *communications*, *fixity*, *signature*, *reach*, *concurrency* and *distance*.

Relative similarity is defined as "the ratio comparing (a) the degree to which one individual is similar to another; to (b) the degree to which that same individual is similar to everyone in the social system" [Kaufer and Carley, 1993, p. 121]. An example of interaction based on relative similarity is the tendency for scholars to form what Crane called 'invisible colleges', based on the degree of fit between their interests [Crane, 1972]. Such informal networks often form the basis of research collaboration and preprint exchanges.

Mental models are defined as "the internal systems through which individuals are able to represent language, meanings, and histories to themselves" [Kaufer and Carley, 1994, p. 20]. Kaufer and Carley argue that mental models are a feature of cognitive agents and that the human mind is a collection of relatively invariant mental models [Kaufer and Carley, 1993, p. 119]. They suggest that such an understanding is an argument against strongly technologically determinist views that the change from oral to print communication altered the human cognitive system. Rather, they argue, "communication technologies affect human behaviour … by altering the content and rate of information supplying an agent's mental models, not by changing the architecture of the mental models themselves" [Kaufer and Carley, 1993, p. 119].

Communications are the externalised aspects of mental models carried through some medium [Kaufer and Carley, 1994, p. 21]. Kaufer and Carley distinguish communications from codes (the physical reality of a semantic representation fixed in some medium) and messages (information that is stabilised by an authority outside the speaker).

Fixity is "the degree to which communication technology enables the communication to be retransmitted without change" [Kaufer and Carley, 1993, p. 100]. Print is an example of a medium that is inherently highly fixed. Electronic text is much less highly fixed (unless stored on a read-only device).

A *signature* is an abstract idea that "links an individual's unique mental model with an external artefact, including a communication … that the mental model is responsible for pro-

ducing" [Kaufer and Carley, 1994, p. 22]. Signatures can be explicitly conveyed through a *handle*, typically a proper name. They can also be implicitly conveyed through the communicator's style - manner, gesture, distinctive vocabulary, choice of phrase or other mannerism.

Reach for an individual can be defined as the number of people whose mental model is affected by a signatured communication from that individual [Kaufer and Carley, 1993, p. 125], [Kaufer and Carley, 1994, p. 25]. The reach of a particular communication at a particular point in time can be characterised as lying somewhere on a number of orthogonal axes: impact (immediate to ultimate), potential (actual to expected), and cognitive comprehensiveness [Kaufer and Carley, 1993, p. 128]. Reach is closely allied to the idea of *diffusion*, "the process by which a communication is transmitted and received" [Kaufer and Carley, 1993, p. 125]. The difference is that diffusion is a property of a communication; reach is a property of an individual.

Concurrency simply says that within a community, multiple interactions can occur in parallel at about the same time. Concurrency implies that "the interactions of one set of individuals may have consequences for others who are not directly involved" [Kaufer and Carley, 1993, p. 153].

While wide-ranging, these key concepts do not of themselves sufficiently delineate the "distancing assumptions of written (as well as print and electronic) communication" [Kaufer and Carley, 1994, p. 12]. Both [Kaufer and Carley, 1993] and [Kaufer and Carley, 1994] regard this distinction between distant and proximate communication as being central to any discussion of written communication. *Distance* "indicates the writer's separation from a reader in space, time, culture or some mix thereof" [Kaufer and Carley, 1994, p. 8]. Recall that reach is a property of the individual. Distance is "a measure of the difference between communication partners" [Kaufer and Carley, 1994, p. 32]. In extending the potential distance between originator and receiver, communication technologies can extend an author's reach. Technology extends distance through *asynchronicity*, *durability* and *multiplicity*.

Asynchronicity removes the requirement that partners in a communicative transaction have to be coexistent in time (and by implication, in space).

Durability is "the length of time the content of a communication is available for interaction" [Kaufer and Carley, 1994, p. 34] and is a property of the medium. Durable texts diffuse more widely and for longer and thus increase the author's reach.

Multiplicity is "the number of communication partners that can be communicated with at the same time" [Kaufer and Carley, 1994, p. 35]. Multiplicity implies greater distance and greater speed in spreading information. Network technologies provide for the largest potential asynchronicity, durability and multiplicity of any communications technologies to date.

2.2.3 Axioms and Theorems

Having identified the underlying assumptions and defined some key concepts, [Kaufer and Carley, 1993] and [Kaufer and Carley, 1994] both identify a series of non-exhaustive axioms that arise from the application of their assumptions to communicative transactions. They can be summarised as follows:

- "Individuals are situated within sociocultural systems" [Kaufer and Carley, 1993, p. 150]
- "Individuals have mental models that limit behaviour" [Kaufer and Carley, 1993, p. 150]
- "People create communications out of their mental models" [Kaufer and Carley, 1994, p. 22]
- "Individual's mental models adapt" [Kaufer and Carley, 1993, p. 151]
- "Individuals continuously engage in communicative transactions" [Kaufer and Carley, 1993, p. 151]
- "Communicative transactions occur concurrently, not serially" [Kaufer and Carley, 1993, p. 152]
- "Communicators can act concurrently" [Kaufer and Carley, 1994, p. 30]
- 'The more relatively similar and individual perceives him or herself to another, the more likely the individual is to attempt interaction with that other" [Kaufer and Carley, 1993, p. 155]
- Interaction and knowledge are "reciprocally related to one another" [Kaufer and Carley, 1993, p. 158]
- "Individuals exert communicative authority by changing the mental models of others

through communications bearing their signature" [Kaufer and Carley, 1994, p. 28]

- "The greater the average potential distance across which an individual can interact, the greater the potential reach" [Kaufer and Carley, 1994, p. 32]

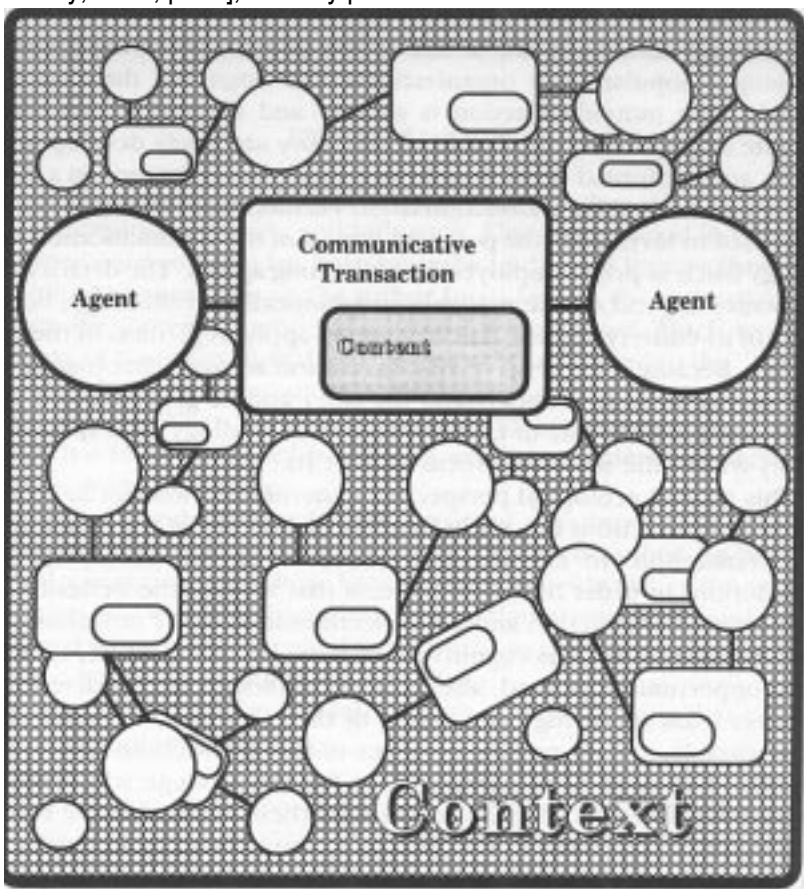
[Kaufer and Carley, 1993] extends these axioms by adding a theorem relating to the relative roles of authors and readers:

- Authors and readers are potentially interchangeable roles [Kaufer and Carley, 1993, p. 152]

2.2.4 An ecology of communicative transactions

These concepts, axioms and theorems play themselves out within an over-arching framework, that of the ecology of communicative transactions as represented in figure 2–1.

Figure 2–1: Ecology of Communicative Transactions. Source: [Kaufer and Carley, 1993, p. 89], used by permission.



Kaufer and Carley argue that communicative transactions, agents, content and context fit into a single ecology of communication. Ecology evokes images of richness, interdependence, and evolving complexity. All of these, Kaufer and Carley argue, are characteristics of the ecology of communicative transactions. Their constructural framework is grounded on

the assumptions already listed and built on its concepts. The two most useful features of their model for this thesis are the idea of an open systems ecology, and the communicative transaction.

Open systems ecology

In a traditional open systems model, an agent is viewed in terms of the inputs into, and outputs from, an agent process. The process is only important because of the way in which it transforms the inputs into outputs. Applying an open system approach to communication "emphasizes the agent, the context (environment), and the content of the transactions (inputs and outputs)" [Kaufer and Carley, 1993, p. 95]. The idea of the ecological constructural framework is to build on this simple model by regarding all these components of communication as "mutually defining, coadaptive, and coevolving components of single ecology" [Kaufer and Carley, 1993, p. 95]. The critical terms here are coadaptive and coevolving. In the context of research into new forms of communication, such a model means that attempting to consider any one component in isolation is fraught with difficulties. Changes in one area will almost certainly prompt unpredictable changes in another. One is reminded of the butterfly beloved of chaos theory that by flapping its wings in the Amazon causes a cyclone in Texas.

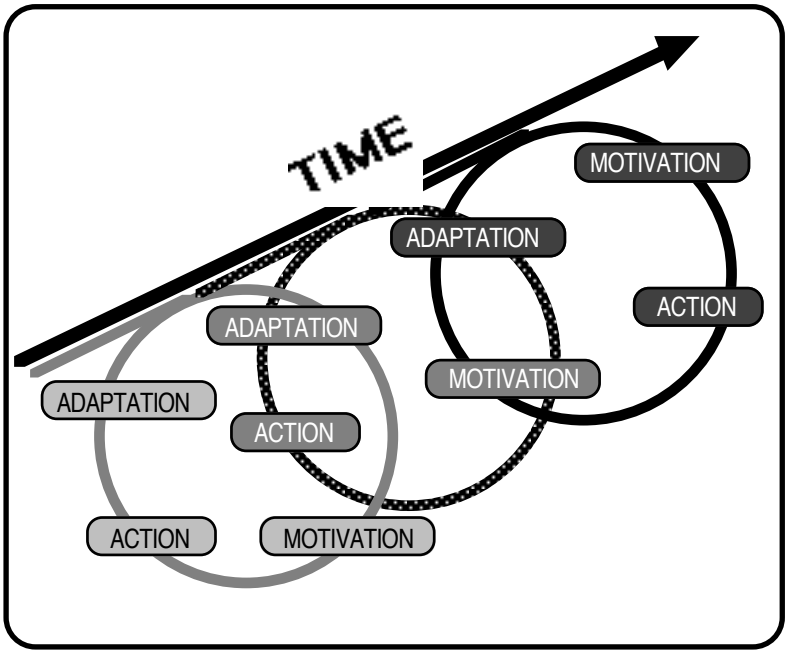
Communicative transactions

Kaufer and Carley describe the communicative transaction as a cyclic process of interaction, communication and adaptation between individuals (shown in figure 2–2). Within each cycle, individuals pass through the phases of motivation (the decision making that decides on an interaction partner), action (the communication of information to that partner), and adaptation (updating one's state of knowledge in the light of the information that has been communicated). These phases do not need to occur sequentially or even in any particular order. Over time, the cycles repeat. In contrast to some theories of communication, any of these phases can drive the communicative transaction.

2.2.5 Areas of application

Having developed their theory in all its richness, Kaufer and Carley apply it through a series of simulation models to a range of areas involving author-reader interaction: print in general, print and the professions, print and academe, and intellectual migration. The first three of these contribute insights for the research questions of this thesis.

Figure 2–2: The Knowledge-Interaction Cycle. Source: [Kaufer and Carley, 1993, p. 144], used by permission.



Print

Their analysis suggests that given the right preconditions, print can "speed diffusion, stability, and consensus" [Kaufer and Carley, 1993, p. 291]. They contrast these results with what they characterise as the Strong Print Hypothesis, based on the idea that print has a range of unconditional effects: extension of author's reach beyond oral communication, origination of an awareness of societal reach, creation of stability and consensus by encoding information. They regard this as far too absolute an idea. Perhaps the strongest statement to come from their own analysis is a fairly bland one about print enabling wider communication through speed of transmission.

Professions

In the case of the modern professions, they argue for a necessary role for print [Kaufer and Carley, 1993, p. 311] in the sense that large diverse professions need to be structured around printed texts. Print is merely a supporting technology, not a deterministic one. The nature of professions depends on the characteristics of a group and not the medium through which they communicate. Like the later technology of electronic mail, print increased the reach of individuals within a profession and thus supported a wider geographical spread of members. Print also bound the members of a profession more closely together through shared experiences of common printed materials in the forms of journals and newsletters.

Using simulation models, they argue that a constructural analysis of the impact print confirms the following hypotheses [Kaufer and Carley, 1993]:

- *Efficiency* hypothesis: print, though multiplicity, can move information faster and bring about more rapid stability and consensus (pp. 318-323).
- *Expanding Member* hypothesis: print makes the expansion in the membership of professions possible (pp. 324-329).
- *Expanding Culture* hypothesis: print allows a profession to have a larger culture yet still reach consensus and stability as effectively as a smaller non-print using profession (pp. 324-329).
- *Weak Integration* hypothesis: print reduces the need for cultural integration of a profession by mitigating the consequences of initial minimal integration and reducing the reliance on co-presence (p. 330).
- *Strong Specialisation* hypothesis: print increases the ability of a profession to be specialised by sharing information more effectively (pp. 330-335).

Academe

In their analysis of academe, they find all these hypotheses also confirmed. They also discuss the scientific journal as a particular print artefact. They argue that in diffusing new ideas journals are simultaneously faster than book publication or face-to-face interaction (due to their frequency of issue and increased reach respectively), and slower than newspapers (due to the gatekeeper function of peer review). The obvious question is whether the current system is too fast or too slow. The consensus according to Kaufer and Carley is that many scientists regard the speed of journals as too slow, particularly in very fast-moving fields. They refer briefly to electronic journals as a possible solution.

2.2.6 Discussion

Kaufer and Carley are looking at the big picture and providing a detailed analysis of print as a communication technology.

The strength of Kaufer and Carley's work with respect to the research questions in this thesis is precisely that it is focused on print and not tied to a particular new communications technology. This means that it does not date (as would be the case had they considered, say Gopher instead of print), and that its theoretical points are cast in the context of communication in general or print in particular. As I am discussing scholarly communication technologies

that move beyond print (but are inevitably viewed in comparison to it) this is entirely appropriate.

One weakness of their work is that it is very difficult to operationalise. With the exception of a citation study discussed in [Kaufer and Carley, 1993], their work relies almost entirely on strongly mathematical computer-assisted simulation and modelling. Another weakness is that it also lacks specific predictive power when considering change within the ecology: either the likely success or failure of particular developments or the process whereby such change takes place.

2.3 Punctuated equilibrium

In order to gain more insight into change within ecologies, a second theoretical approach is necessary. Drawing on Kaufer and Carley's idea that the system of scholarly communication behaves like an ecology (as does the information industry as a whole - see [Odlyzko, 1997]), what happens if ideas from the studies of 'real' ecologies are brought to bear on this new domain?

2.3.1 Natural selection

Natural selection is a theory of evolutionary change in organisms that was defined by Charles Darwin in 1859 [Darwin, 1859]. This theory rests on two initial premises:

1. Organisms of all kinds vary with these variations being inherited to some degree by their progeny

2. In all realistic situations, organisms will produce more progeny than can survive

The corollary of these premises is that progeny with inherited variations that are favoured by their environments will survive and pass these variations on to their progeny in turn. Thus, over time, populations of organisms will fit better into their changing environments.

Stephen Jay Gould [Gould, 1991] argues that two additional constraints are needed to ensure that natural selection is a creative force and not just a removal mechanism. The first is that "variation must be random … if variation comes prepackaged in the right direction, then selection … merely eliminates the unlucky individuals who do not vary in the appropriate way" [Gould, 1991, p. 12]. The second is that variation "must be small relative to the extent of evolutionary change in the foundation of new species … if new species arise all at once, then

selection only has to remove former occupants to make way for an improvement that it did not manufacture" [Gould, 1991, p. 12].

2.3.2 Evolutionary change

This much is generally understood. However, it is important to realise that Darwin also argues that evolution has no purpose (individuals are only concerned to increase the passing on of their genes to their progeny) and no direction (organisms simply become better adapted to their local environment rather than 'improving' or 'developing').

It is also important to realise that evolution, powerful though it is, is not alone a sufficient explanation for evolutionary change. The standard misapplication of evolutionary theory "assumes that biological explanation may be equated with devising accounts … about the adaptive value of any given feature in its original environment" [Gould, 1994, p. 63]. Other factors may well be powerful influences. Darwin focused on organisms, but there are factors that operate both below and above the level of the organism. Research into genetic biology has demonstrated that change in the DNA of individual organisms is essentially random. At the level of species or higher groupings, mass extinctions or environmental change can wipe out entire categories of organisms in similarly random ways, unrelated to their fitness for a particular environment.

The fossil record reveals some of this change. It also reveals that history is quite unpredictable. As Stephen Jay Gould persuasively argues:

> History can be explained, with satisfying rigor if evidence be adequate, after a sequence of results unfolds, but it cannot be predicted with any precision beforehand … History includes too much chaos, or extremely sensitive dependence on minute and unmeasurable differences in initial conditions, leading to massively divergent outcomes based on tiny and unknowable disparities in starting points [Gould, 1994, p. 63].

2.3.3 Punctuated equilibrium and speciation

Evolution in real ecologies occurs on a time scale that is too long for an individual to observe. If one wishes to look at how organisms in an ecology adapt and change, it is necessary to look at the past. The standard model for historical evolutionary change is called phyletic gradualism. This argues that species gradually and continually evolve into new species over time.

An alternative model was first proposed by Niles Eldredge and Stephen Jay Gould in 1972 [Eldredge and Gould, 1972]. Their key breakthrough was to realise that "a standard biological account, Mayr's peripatric theory of speciation in small populations peripherally isolated from a parental stock, would yield stasis and punctuation when properly scaled into the vastness of geological time" [Gould and Eldredge, 1993, p. 223]. This theory, which they named *punctuated equilibrium* suggests that rather than occurring gradually, evolution proceeds in fits and starts. They propose that the development of new species occurs when parts of the breeding population become cut off from the rest of the population in different environments. They adapt to the new environmental challenges by evolving into new species that no longer need to change [Benton, 1993, p. 33]. If the surrounding environment then also changes, a few of the newly developed species are already preadapted and will quickly out-compete their ancestral relatives. In the fossil record, this then shows up as a sudden change from one organism to another. This addition to standard evolutionary theory has now been "accepted … as a valuable addition to evolutionary theory" [Gould and Eldredge, 1993, p. 223].

2.3.4 Other applications of punctuated equilibrium

Punctuated equilibrium has started to be applied outside the domain of biology. Within the science disciplines, punctuated equilibrium is sufficiently well known to be used without explanation in editorial material [Rubenstein, 1995]. Fields of application in the domain of business include organisational change [Fox-Wolfgramm et al., 1998], product innovation [Eisenhardt and Tabrizi, 1995] and technology diffusion [Loch and Huberman, 1998]. The application of punctuated equilibrium to the ecology of scholarly journal publishing has not (to the author's knowledge) been attempted before and may well yield valuable new insights into its likely evolution.

2.4 A genre-based framework for new media

The constructuralist ecology of communicative transactions deals with the overall environment within which print communication takes place. The theory of punctuated equilibrium focuses on the processes of change within that environment. Neither are directly applicable to the design of the specific communications artefacts themselves. The final theoretical perspective looks beyond the technology but is much more obviously applicable to new media and their implications.

A genre-based framework for new media is based on the work of Phil Agre ([Agre, 1995a], [Agre, 1995b], [Agre, 1995c]), but primarily on [Agre, 1995b]. This presents a framework

for media design based on an inquiry into the role of genres in people's activities. [Agre, 1995a] deals with the natures of communities and how they operate, and [Agre, 1995c] focuses more on the forms and use of information. His main argument is that any design of a communications system using new media requires a focus on the "social relationships around a given type of communication" [Agre, 1995b]. New media are never precisely defined, but presumably include mass media (such as radio and television), electronic text on CD-ROM, multimedia offerings of all sorts and the whole cluster of technologies around the Internet.

The central concept is in this social focus is *genre* - "the expectable forms of communication that fit into particular forms of activity involving relationships between communities of people". [Agre, 1995b] The power of genres for Agre can be summarised by saying that "in analytical terms, they are the meeting-point between the process of producing media materials and the process of using them" [Agre, 1995b]. He discusses the insights that applying this theory provides under the headings of communities, activities, relationships, media, and genres. He then provides an extended example to consider how these insights might apply to the design of a set of Web pages for a hypothetical organisation. Last he examines the economic considerations that need to be kept in mind in designing new media.

2.4.1 Definitions

A *community* is defined by Agre as

> the set of people who occupy a given structural location in an institution or society. … Most communities engage in some degree of collective cognition -- the interactions through which they learn from one another's experiences, set common strategies, develop a shared vocabulary, and evolve a distinctive way of thinking. [Agre, 1995b.]

In the context of scholarship, a very broad application of community might be all scholars. A narrow definition might be a particular discipline such as physics or a narrow sub-discipline like string-theory cosmology.

Communities are not isolated constructs. Community implies "shared forms of activity within a particular institutional logic" [Agre, 1995c]. According to Agre

> 'activities' include both the physical actions (sitting, writing, talking, looking, turning pages, pushing buttons, etc.) and the cognitive and emotional processes (identi-

> fying with characters, figuring out what's important, wondering what the professor thinks is important, catching the allusions, etc.). The genre needs to ‘fit’ with the whole complex of ‘external’ and ‘internal’ aspects of the activity. [Agre, 1995b].

Scholarly *activities* include writing and reading articles, conference papers and books, taking part in the processes of journal publishing by refereeing and editing, teaching, researching, collaborating with other scholars, communicating with the wider community, applying for grants, and administration.

Communities express themselves in shared patterns of activity, but the sense of community comes from the quality of shared *relationships* within that community.

> Many of the characteristic activities of a community either directly involve these relationships (asking for a loan, writing a report, casting a vote, holding a meeting, and so on) or are heavily influenced by them (learning skills, gathering ammo, making oneself presentable, thinking about analogous relationships in others' lives, and so on). [Agre, 1995b]

The links that bind scholarly communities together include those that derive from the university system of education (graduates from particular departments, protégès of a particular researcher), those that come from membership in formally constituted scholarly societies, informal networks that often form or are renewed at conferences, and membership of mailing lists. Members of such scholarly communities are in turn linked to publishers through activities like refereeing and editing.

Media are defined as the “specific technical means of communication” [Agre, 1995b] used in activities. Agre notes that the affordances of the medium condition how it will be used.

> For example, it is difficult to carry a VHS playback system, it is physically painful to read a long text on a computer screen, radio is much easier to use while driving than television, overhead transparencies can be projected better onto whiteboards than chalkboards, email requires net access, face-to-face conversation requires travel, and so on. But media should not be confused with genres. [Agre, 1995b]

A genre is a “relatively stable, expectable form of communication” [Agre, 1995b]. In Agre’s analysis, genres are intimately embedded into his framework for thinking about new media: they are designed or have evolved for specific communities and “fit into particular activities in the lives of that community's members” [Agre, 1995b]. They are also usually closely

linked to a particular medium:

> A novel might not change its words in the transition from paper to CD-ROM, but nobody really knows whether anyone has any use for a novel on a CD-ROM, or whether CD-ROM's need new genres that can participate in the activities for which the CD-ROM medium can actually be useful to the members of a particular community. [Agre, 1995b]

Consistent with Kaufer and Carley's theme of an evolving ecology of communicative transactions, Agre argues that "it helps to think of a genre in historical terms as the product of an ongoing process of coevolution between its producers and consumers" [Agre, 1995b]. Genres coevolve with the network of practices in which they participate, they shape the activities of members of communities, and they are shaped by those activities in turn [Agre, 1995c, p. 227].

Other researchers have also found this idea of genres very fruitful. Levy argues that "each genre has a characteristic rhythm of fixity and fluidity" [Levy, 1994]. Nunberg has discussed (in the context of applying genre theory to electronic newspapers) that genres mold their environment [Nunberg, 1993]. Furuta and Marshall have discussed the way in which genres can act as a reflection of the underlying technology in the context of Web homepages [Furuta and Marshall, 1995]. Yates and Sumner argue that "the new burden for providing fixity in communications is being met by increased reliance on genre" [Yates and Sumner, 1997, p. 3].

2.4.2 Application questions

According to Agre, the key concern when designing a new genre is to "Pick a community, explore how existing genres fit into existing activities and relationships, and then consider how a new genre might 'do more' for the people than the ones they already use." [Agre, 1995b]. To aid in understanding this question of 'doing more', Agre poses a series of questions in the context of an extended example - the design process involved in getting an organization on the Web by creating some prototype Web pages (all quotes taken from [Agre, 1995b]:

Users: "Who are these pages for? What defines their relationship to us? What goes on in the life of each community? How is each community changing?"

Purpose: "What are the stages in the life cycle of our relationship with each individual

in a given community, and what role (if any) is each medium and genre supposed to play in each stage of the cycle?"

Activities: "What activities are the people going to be engaged in …? What other media and genres do they employ in the course of these activities?"

Usage: "What existing genres, whether on the Web or in other media, are going to shape their expectations when they encounter the new genre of Web pages we are designing?"

Medium: "What are Web pages going to do for these people that cannot be done better on paper memos or brochures, over the telephone, by electronic mail, in meetings, through posters or newspaper advertisements, and so on?"

Evolution: "How much will our pages change? Will they contain a steady stream of new content? A steady evolution of the existing content? What expectations will the user communities have about these changes, and what expectations would we like to encourage them to have through the design of our genre of Web pages?"

Awareness: "How will the people hear about your Web page and learn your URL? Do your plans effectively require the people to put your URL in their hotlist?"

Access: "How do the practical properties of the Web medium fit with the activities that these people are going to be engaged in? … How powerful are their computers? Do they share their computers with others? What kind of bandwidth do they have to the net? Will they be using our pages at high-load times of day?"

2.4.3 Discussion

The questions posed by Agre's application of genre theory throw useful light on the form and functions of scholarly journals as well as on the roles and practices of stakeholders in the scholarly publishing community. The role of the community is critical. [Nunberg, 1993] points out that any shift to electronic publication will not be possible "in the absence of a social organization that enables scientific communities to compensate for features of print discourse that are lost in the transition" (p. 24).

2.5 Alternative theoretical perspectives

The theoretical perspectives discussed above are not the only ones that could have been applied to this research. A range of alternatives were examined and rejected before making this final selection. The six strongest contenders were modelled scholarly communication, postmodern hypermedia theory, open natural systems models of digital library use, a generalised framework for communication in science, paradigm shifts in science, and social construction of technology.

2.5.1 Modelled scholarly communication

The most influential work on the use of journals by scholars was conducted by W. D. Garvey, based at the Center for Research in Scientific Communication at Johns Hopkins University. Based on the behaviour of 2,030 scholars in the physical and social sciences, he formed a model of scholarly communication. This model, refined in collaboration with B. C. Griffiths [Garvey and Griffith, 1971] tracks the stages from initial research work through informal reports, preprints, journal publication, journal citation and finally appearance in specialised texts or treatises. Their model was originally applied to the discipline of psychology but has since been applied to other disciplines. The time scales on the original model cover up to 10 years for the entire cycle (reflecting its basis in print publication).

A group of researchers [Spink et al., 1998] have proposed extending this model to account for “feedback loops linking the reading of journal articles, books or book reviews with other parts of the scholarly communication process and the initiation of new scholarly work” [Spink et al., 1998, p. 365]. Julie Hurd has also proposed a series of extensions to the traditional Garvey &Griffiths model that she calls the modernized model, the no-journal model, the unvetted model and the collaboratory model [Hurd, 1996].

This model was rejected because it is essentially descriptive and has no predictive power. It also does not take into account the ways in which electronic publishing is changing the patterns of scholarly communication [Mulvaney and Steele, 1993].

2.5.2 Postmodern hypermedia

This derives from the convergence in the late 1980’s of hypertext technologies and postmodern critical theory. Its chief exponents are Jay David Bolter [Bolter, 1991], George Landow

([Landow, 1989], [Landow, 1992]), and Richard Lanham [Lanham, 1993]. For a useful, if insufficiently critical overview, consult [Snyder, 1996].

The theory as advanced in these works deals with writing in general and hypertext writing particular. It lacks direct applicability to the domain of scholarly communication and has little specific interpretive power. It was therefore rejected.

2.5.3 Open natural systems in digital libraries

This is based on the work of Lisa Covi and Rob Kling from the University of California, Irvine. Their research focuses on use of digital library (DL) resources by academic staff and student researchers across a range of universities and disciplines.

[Covi and Kling, 1996] divide models of organisational behaviour into one of four types: closed rational, closed natural, open rational, open natural. They then identified clusters of DL research questions relating to connectivity, content and usability from the perspective of service providers and research users. They found the opposition between closed rational systems and open natural systems to be the most productive in the insights it provided. They argue that both perspectives are necessary. Of particular interest for the research in this thesis is their analysis of faculty research work as an open spiral rather than a document centred loop ([Covi and Kling, 1996, p. 682, Fig. 2]), which is clearly an alternative way of stating the same ideas represented by Kaufer and Carley's Knowledge-Interaction Cycle.

[Covi, 1996] draws on the same body of organisational behaviour theory to contrast the closed rational perspective of bounded database searching with the open natural perspective of social worlds. Social worlds enable the researcher to examine the conventions that govern DL use with a focus on the activity of scholarly communication. This research found that DL users moved in multiple social worlds: the workplace (with proximate and temporal influences), their academic discipline, their occupational niche (and hence research subspeciality). Each of these social worlds influenced DL use in differing ways. Natural systems perspectives see organisations as "competitive organisms who inherit organizational characteristics, develop and change their identities and respond to environmental factors" [Covi, 1996], in a manner reminiscent of Kaufer and Carley's communication ecology.

This body of theory provides some valuable insights into journal use (as a particular DL artefact) but does not sufficiently emphasise the electronic journal as communication artefact.

2.5.4 Communication in science

This derives from the work of Leah Lievrouw ([Lievrouw and Carley, 1990], [Lievrouw, 1992]). The underlying conceptual framework consists of a generalised communication cycle and the concept of social representations.

The communication cycle has three progressive stages: conceptualisation (involving largely interpersonal communications processes within a small group), documentation (where the processes are more organized and involve larger groups), and popularization (where the communication processes encourage the acculturation of ideas and involve potentially millions of people). Electronic journals would be part of the documentation stage, while email lists and discussion groups might be more relevant during the conceptualisation stage. The scientific communication cycle concept is based on the idea of a communication process (particular activities or behaviours that facilitate the sharing and construction of meaning) and communication structures (the relationships between individuals with these shared meanings). Social representations derive from social psychological theory and relate to ways of understanding the world and the nature of social reality.

While a useful way of examining the scientific communication in general, Lievrouw's framework applies at too high a level to be directly useful for this thesis.

2.5.5 Paradigm shifts in science

In the 1960s, the American physicist and historian Thomas Kuhn was studying the work of Isaac Newton and became interested in the way in which Newton's views about the motion of bodies had replaced the previous understanding of the scientific community of the time. Newton's insight (largely still used today, although modified by Einstein's later work in the field of curvature of space-time) was that motion was a deterministic approach. Things moved because they were subject to predictable natural forces like gravity and friction. Prior to this, most people's views about motion were largely based on ideas traceable back to Aristotle. Things moved because they 'wanted to'.

Kuhn [Kuhn, 1970] suggested that the replacement of the Aristotelian worldview with the Newtonian worldview should be regarded as a *scientific revolution*. A particular way of viewing things within a discipline is a *paradigm*, and the revolution replaces one dominant paradigm with another. Such a paradigm shift will usually not occur unless there are significant problems with the current paradigm. Usually, the weight of evidence against a para-

digm will build up and the tensions between the dominant paradigm and the body of understanding in the discipline will become apparent. If a new paradigm becomes available which has at least the same explanatory power as the existing paradigm and which resolves the current problems, the majority of scholars in the discipline will move to switch allegiance fairly quickly.

A good example of such a paradigm shift from earlier this century in the field of geology is the rise in acceptance of the theory of continental drift (now known to be caused by the mechanism of plate tectonics) which was largely the result of ideas first proposed by Alfred Wegener in 1910 [Wegener, 1962]. This was enormously controversial at the time and was still not generally accepted in the mid-1960's [Gould, 1991, p. 160] but is now received wisdom within the field.

Bruce Morton [Morton, 1997] has argued that a similar paradigm shift is necessary in the system of scholarly communication. In this case, the shift would be from the existing system of print based journals to one based on electronic dissemination. Edward Valauskas argues that because of the deficiencies of text on a screen versus text on paper, such a paradigm shift will never happen. Instead, "the future will be rich in print, electronic and mixed media for scholars" [Valauskas, 1997]. The notion of paradigmatic shifts has also been applied to conceptualisations of the broader information economy [Braman, 1995].

The notion of paradigm shifts is a useful one, but much of its explanatory power is also contained within the theory of punctuated equilibrium. Punctuated equilibrium also has the advantage of linking more directly to the ecological framework of Kaufer and Carley. For this reason, paradigm shifts were not used directly as a theoretical framework.

2.5.6 Social construction of technology

The Social Construction of Technology (SCOT) approach [Bijker et al., 1987] proposes that for any given situation a number of technological artefacts arise. Initially there is great flexibility of design with many alternative technologies being available for adoption. Over time, a process of selection and winnowing out takes place. Finally, large constituencies within the users of the technology generally agree on the purpose, meaning and physical form of the technology and what Bijker and Pinch call *closure* takes place.

This framework has been applied to bakelite [Bijker, 1987], brewing [Hård, 1994], and the development of the bicycle [Bijker et al., 1987]. Nancy Fjällbrant has argued that one can

successfully apply this analysis to the development of the scholarly journal [Fjällbrant, 1997].

It is quite possible to see how such an analysis could also be applied to the current move towards (perhaps temporary) closure of delivery and presentation technologies in the domain of parallel print and electronic delivery journals [Hitchcock et al., 1997]. This is an initially tempting prospect. However, closer examination reveals a number of similarities between SCOT and the evolutionary models discussed already. The process of developing many alternative technologies is very reminiscent of the adaptive radiations of organisms into ecological niches so well described by Stephen Jay Gould in his description of the Burgess Shale [Gould, 1989]. The process of closure is then analogous to the selection of one best-adapted organism. Because of these parallels and because of the greater predictive power of the ecological model, SCOT was not selected.

2.6 Conclusion

The theoretical perspectives chosen are directly relevant to the question of transformations within the scholarly journal. It is important to realise that the three main bodies of theory operate on three different, but related, levels. The first, a constructuralist ecology of communication, addresses the overall environment in which scholarly communication takes place and which scholarly journals ‘inhabit’. The second, punctuated equilibrium, addresses the ways in which changes might take place within such an ecology. The third, a genre-based framework for new media, addresses the task of designing a new ‘species’ of journal to inhabit such a changed ecology.

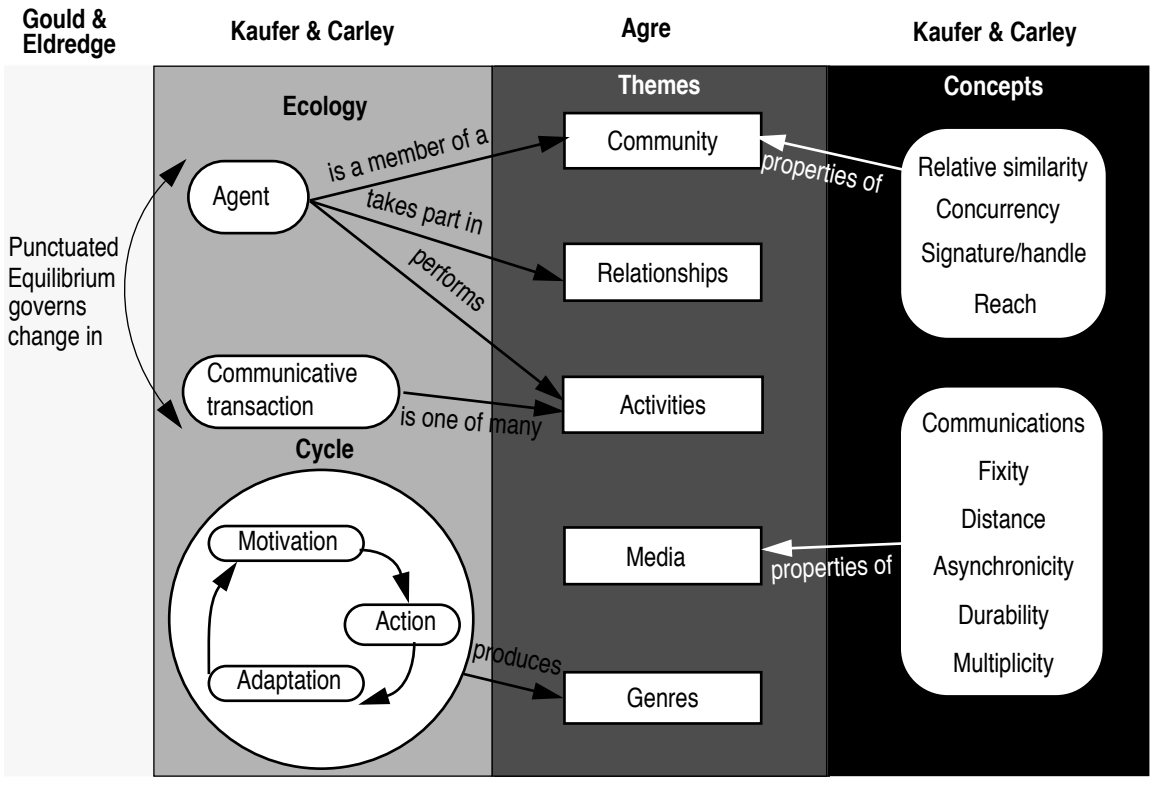
The work of Kaufer and Carley provides a range of precisely defined characteristics of the communication process that can be examined in the context of hypermedia scholarly journals. Their idea of an evolving open systems ecology is an extraordinarily powerful way to think about the interactions with the evolving field of electronic scholarly publishing.They also provide a number of specific suggestions with respect to new communication technologies that will need to be considered in the light of the possibilities inherent in those technologies.

Gould and Eldredge’s model of punctuated equilibrium change in species builds on the idea of the communications ecology to analyse the processes of change and make predictions about both the rate of change and ability to predict developments from known starting points.

Agre's theoretical perspective provides a very useful way to think about some of the key components in the ecology that surrounds hypermedia scholarly journals. The notions of community, activity, relationship, medium and genre are all very powerful ways of conceptualising what happens when we use a journal.

Figure 2–3 shows the linkages between these three theoretical perspectives The concepts of agent and communicative transaction from Kaufer & Carley's ecology of the communicative transaction are both linked to Agre's perspectives. An agent is a member of (at least one) community, takes part in relationships with members of that community, and performs activities (both within the community and outside it). The communicative transaction is one of those activities. Kaufer and Carley's concepts of relative similarity and concurrency are both properties of the community. The members of the community also have particular mental models, signatures, and reach. Communications, fixity, distance, asynchronicity, durability and multiplicity are all properties of the medium of communication. Kaufer and Carley's communicative transaction cycle produces (among other things) during its process of evolution Agre's genres. Finally, Gould & Eldredge's augmentation of the theory of evolution with their mechanism of punctuated equilibrium governs the processes of change within the ecology. Thus the three perspectives complement and inform each other. Each has its own concerns, but links powerfully to the other. Together, they provide a rich base for this thesis.

Figure 2–3: Integrated Theoretical Perspectives



II Publishing & Technology

“Now we live in an electronic field of dreams. We believe that if we build it, they will come” [Hunter, 1994]

3 Print Publishing of Scholarly Journals

3.1 Introduction

The previous chapter considered the most appropriate theoretical perspectives to use when discussing hypermedia on-line journals and rejected a number of possible candidates. Two of the three perspectives chosen make either an explicit or implicit reference to print. Kaufer and Carley are primarily concerned with print publishing. Agre, by dealing with new media, implicitly contrasts them with their print predecessors.

The title for this thesis is 'Hypermedia On-line Publishing: Transformation of the Scholarly Journal'. This implies a transformation *from* something as well as *towards* something. It is necessary, therefore, to look at the origins and development of the system of scholarly communication that is represented by the scholarly print journal, place this system in context, and provide an assessment of its advantages and disadvantages. This ensures a solid base for comparison in later discussion of possible successor technologies.

This chapter first considers the development of print scholarly journals. The role of scholarly societies in formalising communication is particularly important for the development of journals and also needs to be acknowledged. Print scholarly journals are now a mature communications technology, and the characteristics of this technology will be covered next. The various stakeholders in the print journal ecology are outlined and the chapter concludes by looking at pressures for transformation of the print journal system. This chapter does not attempt to deal with the development of print (which is an entire topic in its own right) nor with non-scholarly (or trade) print journals.

3.2 Development of print journals

3.2.1 Early developments in communication

Prior to the recent rise of networked computer and communications technology, there have been four major developments in communication.

The first was the appearance of language, probably some time in the last 200,000 years. Without language, the range of possible things to be communicated is severely limited (although by no means null). Communication of ideas is one of *homo sapiens*' great cultural achievements. At first one could communicate through space only as far as sight and sound

could reach; through time only as far as human memory was reliable. Oral traditions were used to preserve information in ways that are still visible (although now in written form) in works like fairy stories and the Pentateuch, and in living oral traditions like schoolyard rhymes and games that constantly needed to be renewed [Rose, 1993, chapter 4]. The Icelandic society circa 1,000 C.E. even based an entire legal system around the office of the *lögsaguma∂r* (literally 'law saying man') who had memorised the law and could pronounce decisions based on it. However, with only oral communication, space was still a barrier and memory was still fallible.

The second development was the invention of a written form of language circa 5,000 B.C.E. This now permitted communication through both space and time. Assuming durable materials a message could be carried to the ends of the earth, or preserved to the end of time. Knowledge could now be extended to one's contemporaries and descendants. Most importantly, cultural advances did not need to be re-learned by rote each generation, or lost in some catastrophe. Each new generation could start from the level reached by the last. Writing effectively allowed civilisation to take off. The endless inventiveness of humans is seen clearly in the various solutions they have developed for the problem of making concrete records of language, both in terms of symbolic systems (ideographs, syllabaries, alphabets) and recording materials (wood, stone, papyrus, clay, parchment, vellum, paper). It is important to note that the availability of materials often conditioned the form and medium of writing. The Babylonians had no access to papyrus, and little suitable stone. Therefore, they chose to use clay tablets, and adopted a writing system that involved pressing a stylus into the tablets and making triangular marks (cuneiform).

The third development was that of printing, developed around 1,000 - 1,500C.E. (depending on how one defines printing). To use modern terminology, printing provided the ability to broadcast a communications artefact (because of the ability to print lots of identical copies for distribution) in contrast to the narrowcasting inherent in manuscripts (because of the time taken to make each copy). The widespread use of printing brought with it the gradual development of additional navigation features (that had not existed in manuscripts) and greater standardisation [Schaffner, 1994]. Printing also began the knowledge explosion that is still continuing (and accelerating!) [Eisenstein, 1979, [Eisenstein, 1983]. There are interesting parallels between the rise of printing and the spread of the Internet today:

- printing initially a technology for the few
- developments took a while to pick up
- slow uptake in society because people didn't appreciate the impact of technology

- problems in formulating standards
- issues of freedom of access to information and regulation of transborder dataflows

The fourth development was the ability to communicate at a distance, not just in the sense used by Kaufer and Carley [Kaufer and Carley, 1993] when discussing print, but without having to physically move an artefact for the communication to take place.Initially, physical transport of messages on land was dependent on roads. Letters and despatches had been used for centuries, but mostly for government purposes only. The development of postal services for the public at large is comparatively recent. Initial attempts were made by merchants to speed the movement of business information. The 'Penny Post' was introduced in the U.K. by 1688 and by 1702-3 nearly one million letters were carried by this arrangement. The impact of steam locomotion provided increased speed for mail transport, as well as improved direct interpersonal communication through personal mobility. Steam engines also increased the speed of transport of mails overseas. Again, there are significant parallels with developments in the information age: the integration of technologies led to increased growth, and technological change brought about an incremental increase in the speed of social change.

The critical breakthrough in communication at a distance was the separation of the transmission medium from the communication of the message. Prior to the 20th century, there had been some steps along this road: visual signals via beacons or fire (usually on/off messages only), African 'talking drums', flag-signals for the navy, heliographs, and the semaphore system (a 5400 km network of transmission stations built every 16 km) initially constructed to support the armies of the French Revolution. The development of electric/electronic communication in the last two hundred years has brought first the telegraph, then the telephone, and now satellites, fax machines and the Internet.

Despite all these changes in communication technology, the overall trend has been to keep multiple levels of the communications 'revolution' going simultaneously, while the leading edge keeps increasing in speed and capability. We still write letters, for instance (even if some of us spend much more time writing email!).

3.2.2 Rise of the scholarly journal

Prior to the development of printing, there was little formal communication because of the difficulties in creating and distributing copies of scholarly works to other scholars. Scholars did write extensively to one another and often at great length, but this communication was both slow and limited to narrow personal circles [McKie, 1979]. The rise of scholarly soci-

eties and their associated journals changed all this. In fact, a number of strands came together in the late 17th century to cause the appearance of the first scientific periodicals.

Firstly, printing made easy production of multiple copies available and improved communication technologies (in particular postal and courier services) made distribution of these copies possible. These technologies had existed for some time before they were taken up for scholarly communication and did not change significantly over the course of the 17th and 18th centuries [Kronick, 1976, p. 280]

Secondly, periodical publication in other areas had existed for up to a century prior to the first scientific journal: book catalogues, calendars, almanacs and newspapers. The technologies of the newspapers and almanacs suggested a possible format and distribution channel for scientific results [Kronick, 1976, p. 279].

Thirdly, the existing scholarly correspondence patterns lent themselves to transformation into print. The tone of the communications was already somewhat impersonal in tone and contained items of news and recent discoveries. These scholarly letters also passed through the hands of a small number of learned men who acted as gatekeepers and redistributors (a role that could easily be transformed into that of editor) [Kronick, 1976], [Fjällbrant, 1997].

Fourthly, there was a movement across Europe towards the establishment of scholarly societies as a way to communicate and cooperate better. In 1600 in London the Royal Society was founded, followed by the *Académie des Sciences* in Paris in 1666 and the *Kungliga Svenska Vetenskaps Akademie* (Royal Swedish Academy of Science) in 1739 [Fjällbrant, 1997].

In response to these technological potentials and to other pressures, the journal in its earliest form came into being. The earliest scientific journal is generally reckoned to be the *Journal des Scavans*, first published in Paris on Monday, January 5th, 1665. It dealt with news and discoveries in the arts and sciences, but was not formally linked to the Académie des Sciences, being rather intended for interested lay-people. This publication provided a model and impetus for the development of other journals. The first volume of the *Philosophical Transactions of the Royal Society*, London was published in part as a response. It was a medium for the publication of observations and experiments, and can be viewed as a collection of letters now reaching (and addressed to) a wider audience [McKie, 1979]. These ‘new media’ (for their time) allowed small geographically dispersed groups of people to form communities with shared interests [Brown and Duguid, 1995].

Initial growth was fairly slow. By the end of the 17th century, there were perhaps five journals in existence. The first half of the 18th century saw another five added, but then the pace of development accelerated, with a further sixty nine titles added in the second half of the century, over twenty five of these in the last decade [McKie, 1979]. This acceleration continued over the course of the next two centuries, with an analysis of the number of scientific journals available growing from approximately 100 by the end of the 17th century to perhaps 2,000 by the end of the 18th, 16,000 by the end of the 19th and perhaps 90,000 by the end of the 20th [Singleton, 1994].

[Guédon, 1996] has provided an extremely elegant analysis of this process in the context of the advent of science disciplines, the fact that communication technologies are often used in ways unforeseen by their creators, and the organisation of knowledge.

3.2.3 Current status

The current status of print scholarly journals is that they are a mature technology in terms of their production and distribution systems. The trend in print journals has been steadily in the direction of specialisation, with a doubling time in the number of journals on the order of 15 years [Price, 1963]. While it is difficult to be precise there are probably between 70,000 and 80,000 scholarly journals (depending on the definition) currently in print [Rowland et al., 1995, p. 1]. The EBSCO Subscription Services database of serials titles includes more than 256,000 listings [Ketcham and Born, 1996].

Journals are an integral part of the scholarly communication and reward system [Okerson, 1996, p. 195], with promotion and tenure at many universities in the developed world largely dependent on a scholar's publishing record (however in the humanities there is a still a greater emphasis on book publishing). In the United States, the number of articles published "has increased from 208,000 in 1960 or 382,000 in 1977 to 601,000 in 1990" [King and Griffiths, 1995, p. 715], although this largely reflects increases in the population of authors. Print journals are also big business for some publishers, with some journal subscriptions costing more than $10,000/year.

David Kronick has argued that the changes in the scientific periodical have not been "commensurate with the changes in the complexity and the organization of science that have occurred in science that have occurred in the ensuing centuries" [Kronick, 1976, p. 286]. He argues that the periodical has always been forced to play a double role, "that of a repository

for information and as a vehicle for the dissemination of knowledge" [Kronick, 1976, p. 286] and that this is an impossible task which is now threatening the journal with breakdown.

3.2.4 Communication and scholarship

Throughout these developments, scholars have always communicated; indeed one could argue that communication is inherent in scholarship. Can someone really be called a scholar if she works alone and never informs anyone else of her thoughts and findings?

This communication has always been across two dimensions. The first is space - communication with one's contemporaries dispersed across geography. The second dimension is time - communication with one's predecessors and communication with one's successors. Prior to the development of printing these communicative transactions had to be mediated either through speech or writing.

The communication has also been in two modes: informal and formal. The main forms of informal communication have been verbal communication through personal contacts with other scholars. Over time, groups of scholars build up a network of such informal communications channel, often called the 'invisible college' [Crane, 1972]. Now, a large part of such informal communication takes place through the medium of email lists and on-line conferences [Fjällbrant, 1997]. Formal communication has until now required some form of printed or written communication. In fact, one could argue that formal communication was not really possible prior to the advent of the printed journal. This is because of the limited distribution possibilities for any communication artefact inherent in the need to manually copy it or commit it to memory. Formal communication can now also take place online.

3.3 Stakeholders in the scholarly journal ecology

The scholarly journal ecology as it is currently constituted reflects both the history of the print scholarly journal and developments in scholarship.Who are the stakeholders in this ecology and what roles do they currently play?

3.3.1 Scholars

Scholars are broadly defined as those who work in a scholarly environment, most typically a university. According to [Getz, 1997], scholars play three roles in our society:

- they educate the next generation of professionals (teaching)
- they make formal knowledge available to the society they live in (service)

- they generate new knowledge (research).

Within the scholarly journal ecology, scholars can play a variety of roles. They can act as producers of messages (in the form of journal articles) by authoring these articles. They can also act as consumers of articles by reading them. Scholars also perform a critical role as editors for journals in conjunction with scholarly societies or journal publishers. Finally, they act as referees for journal articles, thereby performing a filtering or gatekeeper function.

3.3.2 Scholarly societies

As we have already seen, scholarly societies were closely associated with the first scholarly journals and this association has continued. Most scholarly societies provide a journal as one of the benefits of membership. This journal may be published by the society or may be managed in conjunction with an established publisher (thus benefiting from economies of scale).

Scholarly societies also serve as guardians of shared understandings about a particular discipline, exercising a normative influence.

It is important to realise that scholarly societies act in the domain of publishing both as representatives of scholars (to other publishing companies) and as publishing intermediaries themselves. This is sometimes a source of tension.

3.3.3 Publishers

Publishers in the print world require a range of things in order to function: presses, skilled staff, consumables (paper, in and so on), increasingly complex technology to lay out the print work, salaries for the staff, rent for accommodation and so on. This requires a revenue stream, and the market for many scholarly publications is quite small. A subscription base of less than 1,000 copies per issue is not uncommon.

Scholarly publishing contains two types of publishers. For-profit publishers (typically in the STM area) have found that “scientific and professional publishing has become one of the most profitable – and competitive” areas to work in [Graham, 1992, p. 21]. On the other hand many scholarly societies continue to publish on a not-for-profit basis. The price differential between private and for-profit journals is very large, and may be up to three to one in some disciplines [Metz and Gherman, 1991].

Publishers may actively encourage the founding of new journals. Dale Spender [Spender, 1992] has argued that large for-profit publishers like the late Robert Maxwell's Pergamon Press actively map out profitable areas for new publications. Prestigious academics are approached to act as editors (and gatekeepers) for a new discipline or sub-discipline. Authors and referees then lobby their libraries to subscribe. As Spender puts it:

> While Pergamon Press customarily bore the costs of printing and distribution, it was academics who ran the journal, provided the copy, promoted the journal and pushed for sales. And all this for no payment! [Spender, 1992, p. 19].

An alternative view of the contribution of publishers is that they add significant value to author contributions through a variety of means: management of the gatekeeping or peer-review function, editing, and typesetting and layout [Marks, 1995]. The journal also provides distribution of the published articles and publication as part of a recognised 'brand'. Lyman makes the excellent point that the relationship between authors and publishers should be seen as a "delicate symbiosis" [Lyman, 1995].

3.3.4 Subscription agents

Libraries need to subscribe to journals. Publishers need to sell subscriptions to libraries. At their simplest, subscription agents act as purchasing agents on behalf of libraries to negotiate subscriptions for periodicals with publishers. In recent decades, the subscription agent business has concentrated into a few large organisations like Swets and Zeitlinger or EBSCO, who have developed a range of services to offer to both groups [Singleton, 1994, p. 18]. To the publishers, these agents offer consolidation of accounts and economies of scale. They also offer alternative markets for journals or even single articles, thus increasing the potential revenue stream to the publishers. To the libraries, they offer the potential for a single contact for subscription enquiries and software support for activities like serials claiming. They can also provide alternative document delivery services if required.

3.3.5 Libraries

The final players in the ecology are the libraries. They also perform a variety of roles. One is to service their users, the scholars at their institutions, by having resources to meet their needs. In the past, this has led to a lot of 'Just in Case' resource acquisition to ensure that future needs as well as current ones might be serviced. Librarians have also performed an organising function, cataloguing and structuring information resources for easier access and

use. Finally, they have performed an archiving function, one that the publishers have not been willing to do. A significant proportion of a scholarly library's expenses is in binding and storing (and maintaining the building that houses) back issues of scholarly serials.

3.4 Characteristics of print documents

Existing print journals are so much part of the processes of scholarship that they can easily be taken for granted. It is useful, therefore, to look at the characteristics of print documents (and by extension of print journals) in a little detail before considering reasons for considering alternatives. A number of these characteristics are only notable by comparison with the alternatives (as will become apparent). The literature of Human-Computer Interaction (HCI) describes the affordances of particular artefacts in terms of what they offer to the user [Gaver, 1991]. These affordances are governed by the form of the artefact: print journals will have different (although probably overlapping) affordances to e-journals. What then are the characteristics of the familiar print journal technology and what do they afford?

3.4.1 Strengths

Print journals (even when bound) are eminently portable, and single articles (photocopies or reprints) are even more so. They can easily be transported from place to place and read wherever the user wishes. A common argument against electronic articles (or books) is that they are restricted by where they can be accessed (the standard form of this argument is expressed in [Valauskas, 1994, p. 45] "We will never[1] hear of someone curling up with a good monitor to read in bed").

The resolution of print is excellent, even for what we regard as low print quality. An entry level printer will print at 300 dots per inch (dpi). A normal office quality printer will use 600 dpi. Most journals are printed at 1250 dpi and specialist publications may use 2500 dpi. These resolutions provide crisp type, excellent quality graphics (with no noticeable jaggles on lines) and clear photographs. The contrast between white paper and black type is also excellent.

Paper is almost infinitely annotatable (as anyone who has borrowed a frequently used book from a university library can attest to). Paper journal articles can be highlighted in colours,

1. The author may well be unusual, but he *frequently* reads articles in bed on his notebook computer (at least until his wife tells him to shut it down and go to sleep).

underlined, have comments and notes scribbled in the margin and have pages dog-eared as a form of bookmark. Paper allows the reader to have a vicarious dialogue with the author through the medium of these annotations. Such annotations have even been proposed as the basis for an ecology of hypertext annotation [Marshall, 1998].

Because of our training in school and at university, print journals have the significant benefit of familiarity. No-one needs to go to full-day training courses held in stuffy rooms with titles like ‘Surfing the Scholarly Page’! This familiarity is obvious when one watches someone interact with print. The scholars interviewed by Olsen [Olsen, 1994, pp. 36-38] emphasised the importance of being able to flick through the pages of a journal, quickly scanning tables of contents pages and the text of articles. She argues that scholars are building up a mental model of the context of the material they are reading. She also cites research which suggests that scrolling a long document on a screen interferes with the creation of that context by weakening the visual memory of the location of items on a page. This is sometimes called the ‘everything looks the same phenomenon’ in the literature dealing with e-journals.

Print has evolved over time a range of navigation supports which scholars use both consciously and unconsciously: tables of contents, indexes, sidebars, layout of pages, even memories of issue covers and typography used. Particular journals have a certain ‘look and feel’ that their readers become familiar with.

3.4.2 Deficiencies

Despite all these strengths, print also has some deficiencies as a communication technology.

While portable in small amounts, it can be both bulky and heavy in larger aggregations. Some simple comparisons are instructive. A high-end notebook computer weighs around 2.5 kilos and probably has an internal hard disk of 4 Gigabytes. Assume that 2 Gigabytes are required for system software and applications. This leaves 2 Gigabytes free for user data. The total notebook weight is very close to the weight of 500 sheets of A4 80gsm paper. When printed on using a font like Times 12, each sheet of paper will hold at most 5 Kilobytes of information. This means that 200 pages can store 1 Megabyte of information.The 2.5 kilo notebook therefore can store the equivalent of 400,000 pages of printed information To put it another way, once one moves past the equivalent of 500 sheets of paper, a notebook is a more portable way to carry information.

The physical nature of print also means that it requires physical movement to distribute; the pages have to be sent to someone for them to read them.

Colour is (obviously) printable; most mass-market magazines attest to this. It is much more expensive to print colour rather than monochrome pages and most small-circulation scholarly journals cannot afford the expense.

Print also costs more the more one prints. There is certainly a distinction between first copy and incremental costs, but the incremental costs are significant (particularly if each incremental copy also needs to be sent to someone).

Print is limited to the sort of information that can be represented on paper. A lot of human ingenuity has gone into notation systems for representing sounds, music and even dance on paper, but no one would argue that these notation systems substitute for the real thing.

Paper is also not directly searchable (at least not quickly or realistically for long documents). Print aids (indexes, abstracts, tables of contents) are only partial substitutes.

3.5 Conclusion

The print journal arose out of the need for scholars to communicate effectively and through the possibilities of the printing and communication (rail and sea transport) technologies of the time. It is now a mature artefact embedded in an ecology of scholarly communication with a number of distinct stakeholders. As a mature communications artefact, it has a range of familiar and desirable attributes that need to be kept in mind when considering any replacement. The next chapter will examine some of the technology developments that enable such a replacement.

4 Technology Developments

4.1 Introduction

The previous chapter concluded by reviewing some deficiencies with print as a medium of communication. A number of these deficiencies have only become apparent in the latter quarter of this century in comparison with some alternative technologies. Preeminent among these is the computer with its ability to handle a wide range of types of data, and communication technologies for moving data around.

These technologies can of course also be applied to scholarly communication. This chapter reviews developments in these fields to provide an overview of what is currently available. The emphasis is on what the technologies make possible, rather than the details of how they work. A certain amount of technical detail is provided to ensure that the concepts and capabilities are well understood.

Any starting point in such an interconnected area is in a sense arbitrary. Complicating matters, many of technologies are either invisible or opaque to the end-user. This chapter starts with the most obvious physical manifestation – the hardware that is now commonly found on the desktops of scholars in the developed world. From the nodes on the network, the discussion moves to the networks themselves that link these desktop computers to each other, and in turn to larger minicomputer and mainframe servers. Next comes the emergence of the notions of hypertext and hypermedia, which can be considered a technology for organising and working with information. Finally, the chapter looks at the various software technologies that support the creation of, and access to, hypermedia information.

Before discussing the available technologies, a significant caveat should be made. The electronic publishing revolution is not global, although there are some reasons for optimism [Jacobson, 1994]. There are dramatic disparities in access to information technology and networking between countries (and within countries). The following sections will assume a typical working environment in a knowledge-based organisation in the developed world.

4.2 Desktop Hardware

The last 20 years have altered the image of a typical computer in people’s minds from a mainframe in an air-conditioned machine room to a personal computer on a desk. The rise of

the personal computer (PC) has been one of the defining technological features of the late 20th century. One of the characteristics of this rise has been the extremely rapid rate of change. This change has been driven by the manufacturers of the components that make up personal computers, most particularly by the manufacturers of central processor units (CPUs) and to a lesser extent memory chips.

Because of this rate of change in the hardware, and because software manufacturers are continually providing new features to make use of the increased processing power, desktop computers become rapidly outdated. In the developed world many owners of personal computers have gone through two or more generations of computer technology. The usually accepted lifecycle for a personal computer is just three years. This means that those who seek to use this technological infrastructure for electronic publishing are operating in an environment where change is the only constant and where technologies (both hardware and software) rapidly appear and become obsolete.

From the point of view of electronic publishing, the important hardware advances are in the area of processor capacity, the provision of multimedia support, the improvements in print output and the sheer ubiquity of desktop computing.

4.2.1 Processors

Since the invention of the silicon chip, developments in processor and memory capacity have largely followed Moore's Law, named after one of the founders of the Intel company. Moore's Law states that the power (or capacity) of the process which manufactures silicon chips doubles roughly every 18 –24 months [Moore, 1997]. For a constant price, this means performance of a personal computer doubles on the same time scale. For a constant performance, this means that the price of a personal computer halves every two years.

At the time of writing (mid-1998), a typical entry-level desktop computer has a clock-speed of 233 MegaHertz (MHz). This means that every 244,318,208th of a second, a new instruction is executed by the CPU. Chips with speeds of 333MHz are already available, and speeds of 400 MHz are predicted by the end of 1998.

This phenomenal processing speed enables software manufacturers to introduce new features in the area of multimedia support and improved user interfaces (discussed below).

4.2.2 Multimedia facilities

Early computers (including early personal computers) were largely limited to displaying textual information. Text is only a fraction (albeit a large one) of the possible universe of information that one might wish to present. Current entry-level computers now routinely come equipped with the following features to support multimedia information forms and delivery mechanisms:

- sufficient CPU power to decompress and process complex compressed information streams (including streaming audio and video)
- colour display screens with at least 800 by 600 pixel resolution and 16-bit (= 65,535 discrete colours) pixel depth
- built-in CD-ROM (Compact Disc-Read Only Memory) drives, usually operating at least at 6 times the rotation speed of audio CD's (and hence providing faster access to computer data encoded on CDROMs)
- built-in sound card (or equivalent) for processing sound output and input
- built-in or external speakers for playing sound data.

These features can now start to be assumed by those who wish to deliver content to display on these hardware platforms. It should be pointed out, though, that the resolution of current display screens (at about 90 dots per inch or dpi) is still significantly inferior even to low-end printing devices (300 dpi).

4.2.3 Print output

While it might seem a little incongruous to talk about print output in the context of electronic journals, many people still want to print out articles for a variety of purposes. Printers used to be expensive, slow, low resolution and monochrome. Technology advances in the field of both ink-jet and laser printers have addressed all of the above concerns.

Ink-jet printers are now available at costs that are small percentages (of the order of 10-20%) of the cost of an entry-level computer. Low-end ink-jet printers typically have running costs of less than A$0.05/page for monochrome output or less than A$0.20/page for colour output, print faster than a page/minute, have resolutions of at least 300 dots per inch (dpi) and provide reasonable colour quality. Higher-quality ink-jets provide faster printing speeds, resolutions of up to 1400 dpi and excellent colour output. Ink-jets are best for proofing colour or items where colour output is important.

Laser printers are approximately twice as expensive as inkjet printers and are restricted to monochrome output (unless one wants to pay a *lot* of money) but offer much faster printing (4-6 pages per minute), support for Postscript (see 4.5.3: Page oriented solutions on page 72), and better resolution (usually 600 dpi). Laser printers are best suited for text-intensive work.

The affordability of cheaper printers (and the spread of networking as discussed below) means that most professionals should be able to print out monochrome articles as desired, with an increasing proportion being able to print out colour as well.

4.2.4 Increasing ubiquity

In the developing world, it is now the reasonable expectation that a professional in any field will have ready (and probably sole) access to a personal computer in their workplace. The results from the survey phase of this research (see 7.3.2: Access to technology on page 143) certainly confirm this. Many of these professionals will also have access to a personal computer at home. These computers are usually upgraded on a three-year cycle so improvements to technology will diffuse fairly quickly. For the survey population for this research, over 60% had access to the sort of multimedia facilities discussed above.

This level of ubiquity means that publishers of scholarly journals can start to assume that their readers will have access to desktop computer technology.

4.3 Networks

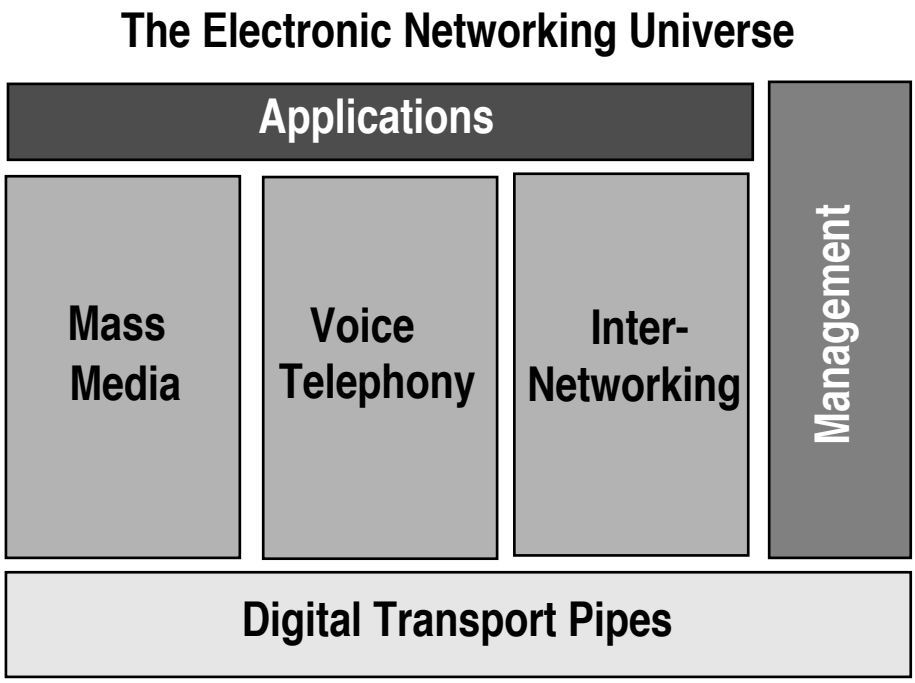
4.3.1 Network infrastructure

Figure 4–2 depicts the components that make up the electronic networking universe.

The digital transport pipes are what allow us to move information around, regardless of its original form. These pipes are increasingly built upon optical fibre technology. As the information we all work becomes created exclusively digitally, the arguments for managing it digitally from producer to consumer become overwhelming.

The management function applies across all activities built upon these pipes. For much of the management software it does not matter what form the information takes. If it is in the form of binary data it can be managed the same way. There are some important exceptions when it comes to distinguishing time-critical media such as sound and video from email and bulk data transfer. Networking technologies like Asynchronous Transfer Mode (ATM) are de-

Figure 4–1: The Electronic Networking Universe. Source: The Internet Society



signed to handle such data differently. The Internet is also implementing mechanisms to support different levels of quality of service (QoS) for different data types in the next generation of the Internet protocol suite (the so-called IPng or IPv6).

The three basic building blocks for all information activities are the mass media (television, radio, and print – increasingly being produced and delivered in digital form), voice telephony (also increasingly digital) and the inter–networking of computers.

Layered on top of these building blocks are the applications that people use to access information and communicate with others. This include traditional computer applications such as electronic mail and the World Wide Web as well as things like voice-response telephony systems and cable television.

4.3.2 The Internet

Internet overview

In the last few years of this millennium, it has become clear that the world is seeing the emergence of a new way of working with information, based on computer networks and the services they provide. This new information domain is called variously cyberspace, the Matrix or simply the Net. It consists of networks interlinked on a global scale. These networks range from small–scale Local Area Networks (LANs), through cross–organisation Wide Area Networks (WANs) to Global Area Networks (GANs). Some of the GANs, like Digital Equipment's DECNet, are proprietary to an organisation. Others like FIDONet are made up of

numbers of small Bulletin Board Systems (BBSs). Others again like BITNet are special purpose research networks. Lastly there are the commercial GANs like Compuserve.

The Internet is a core part of this global network architecture – in a very real sense the ultimate GAN. Originally developed as a way of ensuring the U.S. communications had no single point of failure during the height of the Cold War, it has grown far beyond its research and defence origins. It can now be considered as the 'network of networks', connecting over tens of thousands of separate networks around the world and millions of host machines, many of these being multi–user systems. In Australia, the principle provider of Internet connectivity to Universities and research organisations like the CSIRO is the Australian Academic and Research Network (AARNET). An increasing number of private Internet providers are selling Internet access to the general public.

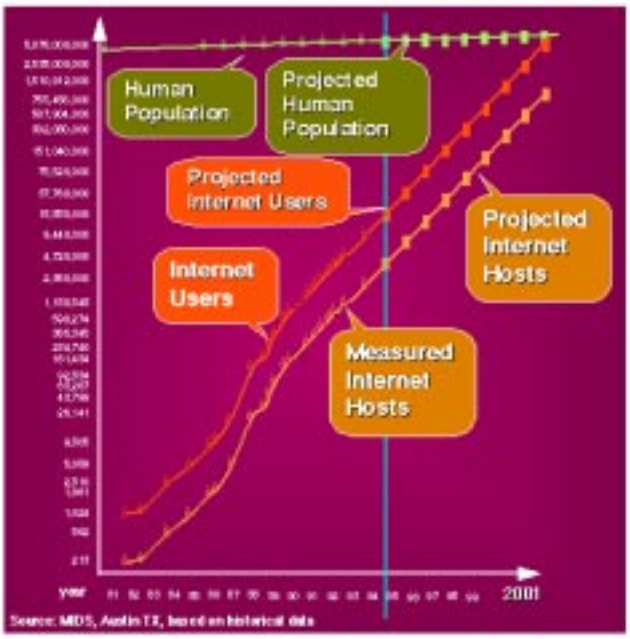
There are other major WANs that are not based on the Internet protocols. However it is possible to communicate between them and the Internet via electronic mail because of mail gateways that act as translators between the different network protocols involved. As well, large computer companies like Apple and Microsoft are providing built–in Internet access with MacOS8 and Windows 95/98. Many of the more innovative and exciting information services that are being developed at the moment, such as the World–Wide Web, are only accessible on the Internet proper.

Any numerical measures of the size of the Internet become out of date very rapidly. At the time of writing, the Internet was growing at 100% per year, versus 10% per year for the voice network. Assuming current trends continue, data traffic will overtake voice traffic around the year 2002 [Coffman and Odlyzko, 1998]. At the same rates of growth, the number of users is rapidly converging with the total human population, as shown in figure 4–3. Of course, this convergence will not take place. The growth in Internet users will gradually taper off at some point, but the figure shows just how rapidly the growth is occurring. The hosts shown in figure 4-2 are computers with an IP address that are connected to the Internet. This includes both single-user personal computers and larger multi-user machines.

What most people think of as the Internet can best be characterised by:

- the use of the TCP/IP protocol suite to regulate communication between machines
- the use of the Domain Name System and the corresponding IP numbers to identify machines.

Figure 4–2: Internet Users and Global Population. Source: Internet Society



Transmission Control Protocol/Internet Protocol (TCP/IP)

TCP/IP is the common name for a family of over 100 data–communications protocols used to organise computers and data–communications equipment into computer networks. As well as its use on the Internet, whose members include universities, other research institutions, government facilities, and many corporations, TCP/IP is also sometimes used for other networks, particularly local area networks that tie together numerous different kinds of computers or tie together engineering workstations.

The Transport Control Protocol (TCP) is responsible for end–to–end links between machines on the Internet. The Internet Protocol (IP) is responsible for connections from one machine to the next on the Internet. Other than TCP and IP, the three main protocols in the TCP/IP suite are the Simple Mail Transfer Protocol (SMTP), the File Transfer Protocol (FTP), and the Telnet Protocol. These three are application level protocols. Together, the TCP/IP family of protocols provides a reliable way of getting information from one computer to another.

IP addresses

In order to identify computers on the Internet, each computer, large or small, which can connect to any of the national networks has its own address or IP number. Each site with a national network connection is given a specific range of numbers that it can use for its internal machine addresses. It is then responsible for allocating this set of numbers within its own organisation. The numbers are in the form of four triplets: for example 128.102.128.50. These are sometimes referred to as a dotted quad. This number identifies a unique Internet machine.

This addressing mechanism enables Internet packets to be routed to the correct machine. Note that just because a machine has an IP number does not mean it is connected to the Internet. A significant proportion of TCP/IP networks have no Internet connection at present, although many are implementing such connections.

Domain Names

Because people are not very good at remembering long strings of numbers, machines can also be referred to by their domain names. The Domain Name System (DNS) is a hierarchical, distributed method of organising the name space of the Internet. The DNS administratively groups hosts into a hierarchy of authority that allows addressing and other information to be widely distributed and maintained. A big advantage of the DNS is that using it eliminates dependence on a centrally–maintained file that maps host names to addresses. A Fully Qualified Domain Name (FQDN) is a domain name that includes all the higher level domains relevant to the entity named. For example, for a host, a FQDN would include the string that identifies the particular host, plus all domains of which the host is a part up to and including the top–level domain (the root domain is always null). For example, atlas.arc.nasa.gov is a Fully Qualified Domain Name for the host at 128.102.128.50 (in other words, with 128.102.128.50 as its unique IP number). In addition, arc.nasa.gov is the FQDN for the Ames Research Center (ARC) domain under nasa.gov.

Domain names are typically in the form of machine.site.type.country, although additional sections may be added and some deleted. For instance sol.ccs.deakin.edu.au is a particular machine (sol) in the Computing and Communications Services (ccs) domain of Deakin University (deakin) in the education (edu) domain of Australia (au). Other common domain names are ac (academic – used in the UK), org (organisations that don't fit anywhere else), mil (military), gov (government) and com (commercial). In the NASA example above, the country code is omitted, and hence assumed to be the United States. Most of the country codes are fairly obvious, although Switzerland is ch (from Confederatio Helvetica, the old Latin name) and South Africa is za (from Zuid Afrika).

Most domain names map onto a unique IP number, allowing software to automatically convert from one to the other. Nameserver software is used to manage the databases that match the IP numbers to computer names and locations and make these translations automatically in a way that is transparent to the user. Having machines referred to by domain name also means that the machine that corresponds to a particular name can be changed (for instance to provide more processing power) without users being aware of this. The domain name is

redirected to point to a different IP address. Not all machines with domain names are on the Internet. The DNS allows machines on other networks to have domain names to facilitate transmission of electronic mail between networks.

Client–Server

Much of the Internet is organised around the client–server paradigm. Users interact with a client process (usually embodied in a particular piece of software running on a machine they have access to) which communicates over the network with a server process (usually embodied in a separate piece of software running on a remote machine). Crudely put, the client talks to the user, and the server does the work. The interaction between client and server is regulated by standard protocols. This means that many different clients can talk to many different servers, provided they are all talking the same protocols. For instance, an electronic mail client running on a Macintosh (or under Windows or Unix or VMS) can talk to a electronic mail server running under Unix (or under Windows or VMS or Macintosh). Users can use the clients they prefer, and providers can use servers that are the most appropriate in their situation. Ideally in such an environment it should be possible to mix and match clients and servers to best meet the organisation's needs and its users' preferences.

Current developments

At present, the Internet is going through a period of enormously rapid and exciting development. There is explosive growth in the number of users connected and hosts providing information services; and in the range of information services available, and the tools to access these. In parallel with this growth, the user population is becoming more diverse as private service providers expand access to the Internet beyond the traditional academic and research base. Many of these new users expect the same richness and responsiveness in their global networked environments that they see in their other interactions with the computer on their desktop.

In terms of the technology, the Internet is increasingly driven by the emergence of client–server systems, running on heterogeneous hardware and software platforms and using standardised protocols to communicate. The client–server paradigm has become increasingly important to all forms of computing as micro–computers (the preferred client machines) have become more powerful. In the context of the Internet, client–server allows the system designer to place the part of the information system that interacts with the user - the client - on a microcomputer and use a more powerful (and perhaps more user–unfriendly) remote ma-

chine - the server - to perform the information processing and retrieval. The proliferation and adoption of desktop machines with a graphical user interface (GUI) allows the client application to use different fonts, colours, and graphics to enhance the interface, as well as to support a wide range of media types.

The client–server paradigm in turn demands well–defined standard protocols to govern communication and interoperability. Fortunately the creation of standards is being facilitated by the push towards open systems across the whole computer industry, with the Internet to some extent being able to piggyback on developments elsewhere. A number of standard protocols have been, or are being, developed for communication between Internet access clients and servers. These include FTP, Telnet, Gopher, Z39.50, HTTP and Prospero. Many of these are *de facto* standards rather than *de jure*. The Internet community has observed the extremely protracted gestation of some of the ISO OSI standards, and has drawn the obvious conclusion: technology develops too quickly for such a process to be acceptable. As a result various small groups often develop and propose their own standards for use. Some of these are adopted by others and flourish (Gopher and WWW being good examples); others wither and die. Evolution replaces the committee. This rapid adoption of standard protocols allows the Internet to flourish in a very heterogeneous hardware and software environment. As long as the client or server software adequately supports the necessary protocols, it does not matter much which hardware or operating system is employed. As a result, client and server software for many of the standard protocols can be found for Macintosh, Windows, DOS, VMS, Unix, and IBM mainframe machines.

The information that is being accessed on the Internet is becoming more varied, expanding to encompass digitised photographs, video and stereo sound. Driven by increases in desktop computing power, more capable display devices, and the widespread use of CD–ROM as a capacious (if slow) publishing medium, multimedia has very definitely arrived at the desktop. Even after discounting all the hype, it appears clear that multimedia in some form will constitute a significant part of future access to networked information.

4.3.3 Trends in access

From its origins largely within the world of academic and research institutions in the United States, the Internet is spreading rapidly.

One dimension of this spread is across countries, so that there are very few countries that do not have at least one Internet point of presence. Indeed, access to the Internet is increasingly being viewed as a very desirable part of any development activities.

The other dimension of this spread is within countries and is seen to the greatest extent within the United States (not surprisingly) with Australia coming a close second on many measures. Characteristic of this spread are the following indicators:

- increase in businesses connecting their internal networks to the Internet
- increase in businesses providing a corporate Internet presence
- increasing appearance of Uniform Resource Locators (URLs= Web addresses) on advertising material, sides of buses, T-shirts, etc.
- increase in home access to the Internet through commercial Internet Service Providers (ISPs)
- increase in governments making information available on-line
- increase in ecommerce activities of all kinds

There are no signs that these trends are abating, and it seems reasonable to assume that they will at least continue at the same level and probably intensify. Within the domain of professionals whose work has a strong information component, it is reasonable to assume that the majority have access to networked information through their work, with many acquiring additional access at home.

4.4 Hypertext and Hypermedia

4.4.1 Hypertext

Hypertext at its simplest level is the ability to create linkages between pieces of text. The start of a link may be delineated in some ways (such as underlining) to indicate the availability of a link. The destination (or target) of a link is not normally so delineated. Activating the link (usually by clicking on it with a mouse, perhaps at the same time holding a special key on the keyboard) ‘jumps’ the reader to the destination and displays the text found there. The term hypertext was first coined by Theodore Holm (Ted) Nelson in 1965 [Nelson, 1965]. The ideas underlying hypertext can be traced back perhaps as far as the earliest Hebrew citation indexes [Weinberg, 1997] with manual hypertext/hypermedia implementations being available as early as 1907 [Rayward, 1994], [Rayward, 1997].

Hypertexts are usually made up of more than one node. These are semi-autonomous pieces of text that are connected together and navigated through by means of hypertext links. Links can also connect pieces of text within nodes. For instance, a journal article structured as a single hypertext node might have a table of contents at the start where each entry was a link to that section of the article. Such internal links are usually less common than external links to other nodes.

This simple model of hypertext links can be enriched by adding link directionality, multiple destinations and link typing. A wide range of link types are possible [Grabinger et al., 1992a], [Grabinger et al., 1992b], [Grabinger et al., 1993a], [Grabinger et al., 1993b]. With link directionality, hypertext links can be either uni- or bidirectional. Some hypertext systems, such as the World Wide Web support unidirectional links only. The usual corollary of unidirectional links is that they cannot be traversed 'backwards' from destination to source. This means that it is impossible to see which links point to a particular node. Such information is very useful if one wishes to know how many times a hypertext article is 'cited' or if one wishes to move a node and notify all the links pointing to it of the new destination. Multiple destinations allows a single link to point to multiple destinations and for the user to select which link they traverse. Link typing enables the designer of hypertexts to characterise links according to a predefined set of types. These might be rhetorical classifications: 'comment', 'amplification', 'refutation', and so on [Landow, 1989]. They might also be type of linked data: movie, sound file, Word document, and the like. None of these enrichments are currently supported by the World-Wide Web.

4.4.2 Hypermedia

For the purposes of this thesis, hypermedia will be broadly defined as the combination of hypertext linking mechanisms and multimedia content [Bolter, 1991, pp. 25-7]. The idea of hypermedia (although it was not called this at the time) dates back at least to a visionary paper by Vannevar Bush [Bush, 1945], which described a theoretical machine to augment the human intellect, based on trails of information. The underlying hardware for this machine, called the Memex, was envisaged as being based on microfilm as computers were still in their infancy. Nevertheless, all the key ideas that have been realised in systems like the World-Wide Web are displayed in this article.

Multimedia is the provision of information in more than one medium. Medium is a term that is used somewhat loosely in the literature. One use is the form of the information: text, still image, moving image, or sound. A different use of the term is the mechanism used to deliver

that information: paper, screen, audio or video tape, audio or video disc, and CD-ROM. Of course, these uses are somewhat interrelated; there are only certain valid combinations of information form and delivery mechanism. The technologies to provide such multimedia information are treated in greater detail below.

4.5 Software

Software is necessary to make the hardware and networks useful, and to provide support for hypertext and hypermedia information. A rich universe of software is now available to support a wide range of information-processing activities. This section will focus only on those software technologies that are directly relevant to the topic of this thesis, and will start at the lowest level and work up.

First, this section will consider the graphical user interfaces (GUIs) that provide the user with the ability to easily work with and directly manipulate information. These GUIs mediate between the user and the hardware and provide the platform for the rest of the software tools. Next come the pieces of software that enable the display and manipulation of multimedia information types. Textual information is still largely delivered in page form and so-called page description languages are considered next. Moving one level above pages involves a consideration of document description languages. The most widely used environment to integrate all of these pieces is the World-Wide Web which is reviewed next. The final category discussed will be the tools for communicating with other users over networks. These are important because they enable a degree of interaction with other scholars. There is no attempt to provide an exhaustive coverage of products or technologies in each software category (that would make for a *very* long document). Rather, the most significant examples will be selected and described. Because the technologies discussed are changing so rapidly, what is depicted is the state of the art as of mid-1998.

4.5.1 Graphical User Interfaces

This chapter has already demonstrated the importance of the personal computer as an information access device. One of the most significant breakthroughs in the use of the personal computer has been the development of *graphical user interfaces* (sometimes called *direct manipulation interfaces*). Elements of these interfaces can be traced to the work of Douglas Engelbart in the mid-60's [Engelbart and English, 1968]. The first fully worked out representation of the principles that are now familiar to personal computer users was the Star workstation developed by Xerox [Johnson and et. al., 1989]. Many of these principles (with

a significant addition of new ideas) were then taken by Apple Computer and first realised in the Lisa workstation and then the Macintosh.

The Apple Macintosh was the first mass-market personal computer to provide most of the features now regarded as standard. Microsoft took many of these ideas and produced a succession of inferior GUIs running from Windows 1.0 through to Windows 3.11. Windows95 was the first version of Windows that bore comparison with the MacOS. Due to a variety of factors, Windows95 has become extremely successful and is now installed on virtually all new Intel-processor personal computers. Both the newly released Windows98 and MacOS8.1 provide a range of sophisticated features to support network software, co-operative work and access to distributed information. These operating systems also provide inbuilt support for display of/access to multimedia information.

Development on user-interfaces is continuing. A good review of the state of the art and future directions is contained in [Halfhill, 1997].

4.5.2 Multimedia

The predominant information forms for scholarly communication are text and still images. This seems unlikely to change in the foreseeable future. Yet, when people think about multimedia information, text and image tend to be ignored in favour of sound and video (or moving images). Print and image will be considered shortly (see 4.5.3: Page oriented solutions on page 72).What software technologies support sound and video?

In considering both of these information forms, an important initial distinction needs to made between streaming and non-streaming media. Non-streaming media require the entire file to be available before they can be played. The combination of a networked environment and a large file size can lead to long delays before the user can view/hear any of the file. Streaming media allow the user to begin to interact with them as soon as ‘enough’ of the file is available. What defines enough depends on the speed of the network and the file type, but can often be less than 5% of the file.

Audio

A range of non-streaming audio formats are available. In the Windows world, the WAV format predominates. On the Macintosh, any of AU, AIFF or System 7 Sound files can be encountered. These sound formats vary in whether they encode mono or stereo sound, and the audio sampling rate. Analog sound needs to be sampled to provide digital data. The higher

the sampling rate, the more data is required for a given length of sound. Typical rates are 11.1 KHz, 22 KHz or 44.1 KHZ (this latter one is the sampling rate for audio CDs).

The dominant streaming audio format at present is provided by RealAudio, a product of RealNetworks. RealAudio is designed for deployment over networks and operates at speeds as low as 14.4 Kbps. The user runs a free piece of software called the RealPlayer client (available for MacOS and Windows) which communicates with a server running the RealServer software. Together, the two pieces of software communicate to stream audio data from the server to the client. The client adapts to the network speed available and allows the user to start at an arbitrary location within an audio stream, as well as to rewind, fast forward, and pause.

This software combination is routinely used to distribute audio files of all types over the Internet. One scholarly use is to archive the audio of conference sessions. An example of this is the *Computers, Freedom and Privacy '96* conference held at MIT. Every session is on-line at `<http://www.swiss.ai.mit.edu/~switz/cfp96/#program>` for later access and replay. A particularly novel use of RealAudio has been the broadcasting of dissident Serbian radio over the Internet [Pantic, 1998].

Video

The main non-streaming video formats are the Motion Picture Experts Group (MPEG) family of encodings, Microsoft's AVI (playable on both Windows and MacOS computers with special software) and Apple's Quicktime Movie (playable on both Windows and MacOS computers via free viewers). Of these, Quicktime is the superior format (and has just been endorsed as the basis for the new MPEG-4 encoding). The majority of video content on the Internet is encoded as Quicktime. Unless the user has a very powerful computer with graphics support in hardware, they will probably not be able to view full-screen video at 25 or 30 frames per second (fps), the standard for television or cinema. Typical workarounds are smaller than screen size windows, lower frame rates or both.

The main streaming video format is Realvideo also from RealNetworks and accessed using the same RealPlayer client. Available network bandwidth on the Internet makes accessing video on-line an exercise in frustration for most users at present. Typically the software has to drop the video frame rate to 5-10 fps just to allow the audio track to continue to play without (too much) interruption. This should improve as bandwidth improves (although modem speeds are at or near their theoretical limits and alternative technologies will be needed) and

compression techniques get better (although there is just so much information that can be packed into a limited number of bits).

The Quicktime video format has just added Quicktime Quickstart which allows the user to start to play a video before it has finished its download. This is not true streaming, but it produces a similar effect.

4.5.3 Page oriented solutions

These are technologies or languages which describe the appearance and content of a page (or series of pages) in such a way that the page can be reproduced by a piece of hardware (usually a printer) or displayed on a screen. A pseudo-Darwinian process of selection has reduced a number of early contenders [Gruman, 1995] to a small field. The two best known technologies are both products of the Adobe software company: Postscript and the Portable Document Format (PDF).

Postscript

Postscript was invented by John Warnock and Charles Geschke, co-founders of Adobe, in the early 1980's. It is a formal programming language that is specialised for producing print output. Postscript is not normally something that end-users interact with directly. The software that they use generates a Postscript file when they print their document. This Postscript document is then sent over a network connection (or rarely a direct connection) to a printer (usually a laser printer, although some high-end inkjet printers also have Postscript as an option). The printer contains an implementation of Postscript in Read Only Memory (ROM) and a Central Processing Unit (CPU) of its own. The printer takes the Postscript code and runs it as a program. The output of this program is dots placed precisely on pages that are then output.

Postscript, in its first implementation inside Apple's Laserwriter printers, was largely responsible for starting the desktop publishing revolution. For the first time, designers could product sufficiently complex output on (relatively) affordable printing hardware.

Postscript is still used as a medium for distributing documents on the Internet, particularly within the physics, engineering and computer science communities. Naturally, the user needs a Postscript printer or a Postscript interpreter written in software (such as Ghostscript) to print or view these files.

Adobe's Portable Document Format (PDF)

Postscript suffers from a number of problems as a universal electronic document format:

- the files tend to be fairly large (perhaps 50 K per page)
- fonts used are restricted to the 13 fonts that are known to be built into all Postscript laser printers
- files are not easily viewable on screen
- files are not easily searchable
- the readers require a Postscript printer (or software interpreter) for output
- the electronic documents are basically print-only; they cannot have any additional navigation/viewing features built into them.

The Portable Document Format (PDF) was invented by Adobe as a solution to these difficulties. PDF provides:

- resolution-independent output
- fixed page-size documents
- a range of navigation mechanisms
- searchable text
- the exact electronic equivalent of the equivalent print document
- links to Web sites
- security restrictions on use (no copy/no print) built into the PDF file
- a range of compression techniques
- ability to include fonts or font subsets into files
- much smaller file sizes than Postscript

PDF (sometimes misleadingly called Acrobat) can be authored in three ways:

- Use a PDFWriter printer driver that intercepts the print stream and produces PDF output
- Use the Acrobat Distiller program that converts Postscript output from any program into PDF
- Use an Adobe program (Pagemaker, Framemaker, Photoshop, etc.) that can produce PDF directly

Once the PDF has been created the Acrobat Exchange program allows direct editing of PDF, importing of images, and the creation of navigational overlays. It is also possible to annotate Adobe PDF files but only with the full Acrobat Exchange product. This means that very few readers of PDF can make use of this facility. It is important to note that PDF is a proprietary

product (although its specification is openly available) and the authoring programs are not free.

In order to access PDF files, the user needs the Acrobat Reader software which is freely available for MacOS, Windows, some Unix flavours. This reads PDF files and renders them on screen or sends them to any printer for output. If the creator of a PDF file used a font not on the reader's computer and did not embed that font in the PDF, the Reader can use Adobe's Multiple Master technology to create virtual fonts that preserve the original line breaks and spacing.

The easiest way to think about PDF is as *electronic print*. It allows content creators to generate documents which match as far as possible the print originals. In fact, when printed on a laser printer they are usually indistinguishable from photocopies of the print originals. Because PDF is based on Postscript (it is essentially a specialised dialect of Postscript) it fits easily into the print production processes of many publishers.

John Warnock likes to talk about PDF as representing the shift from 'Print then Distribute' to 'Distribute then Print'. In other words, creators of content no longer have to have warehouses full of printed materials that they have to physically move to their users (and which become out of date, cost money to store, can't be easily changed, etc.). Creators simply have to make available a faithful electronic version of their document and let the users print it themselves (if they choose to). This, of course, presupposes two things: that people have ready access to a printer, and that they are prepared to accept the lower resolution associated with desktop printers (300 or 600 dpi as opposed to 1250 or 2500 dpi for professional publishers). The first of these is a fairly safe assumption in most corporate environments (and increasingly in home environments as well) in the developed world, and the second depends on the type of material being produced. Colour printers are still fairly rare, but much publishing only requires monochrome output and 300 dpi is usually adequate if not optimal. (see 4.2.3: Print output on page 59).

PDF is an excellent choice for documents with complex formatting requirements and in situations where exact page fidelity is required. Because of the font embedding and freely available reader it largely avoids many problems with cross-platform documents. For all these reasons, PDF is rapidly becoming a standard option for publishers who are producing parallel print and electronic versions of their publications. A typical scenario is for such publishers to make Web versions of their articles available for viewing on screen with a PDF version to download and print out. The printed PDF will usually display the identical formatting (down

to page numbers) of the print original and will probably be of a higher quality than an inter-library loan photocopy.

4.5.4 Document oriented solutions

These consist of markup languages which are focused on describing entire documents rather than individual pages. They either make no explicit reference to formatting or do not specify a particular page size. The three most significant languages are all interrelated: SGML, HTML and XML.

Standard Generalised Markup Language (SGML)

Standard Generalised Markup Language (SGML) is best thought of as a document markup definition language. Defined as ISO Standard 8879:1986, it provides a formal notation for definition of generalized markup languages [Goldfarb, 1991]. Such languages are defined as a Document Type Definition (DTD). The DTD defines the allowable tags for a particular document type, and the permissible sequence of these tags. In effect SGML is "a *metalanguage*, in which tag sets, as well as usage rules for these tags, can be defined" [Marcoux and Sévigny, 1997, p. 586]. A wide range of existing standard DTDs have already been created, and organisations using SGML will often have their own in-house DTD.

The important thing to note about SGML is that it encodes the *structure* of the document and does not define its appearance. This is handled by the Document Style Semantics and Specification Language (DSSSL). Because of this explicit focus on document semantics rather than expression, SGML lends itself very well to repurposing of content. An SGML-encoded document can easily be rendered for print output, delivery on the Web, or distribution on CD-ROM.

Although SGML predates the Web (and indeed was hinted at as early as 1970 [Goldfarb, 1997]) it has been taken up fairly slowly. The *Online Journal of Current Clinical Trials* was an early e-journal user of SGML technology [Keyhani, 1995]. This is because it is a fairly complex system and the benefits are most accessible to large organisations with complex sets of technical documents. It is currently being adopted more enthusiastically by organisations wishing to reuse their content in a variety of media and forms [Marcoux and Sévigny, 1997]. Work is also proceeding apace on further development of SGML and its related standards: DSSSL, HyTime (the Hypermedia Time-Based encoding language) and SPDL (Structured Page Description Language) [Mason, 1997].

HyperText Markup Language (HTML)

HyperText Markup Language (HTML) is formally defined in terms of the ISO Standardised Generalised Markup Language (SGML) as a specialised DTD. It provides a standardised way to create structured textual documents for delivery on the Web (and increasingly elsewhere). In the context of the WWW initiative, HTML is used to encode Web documents and embed the links that together to form the web. Non–HTML documents that are pointed to lie at the periphery of the web – they cannot themselves point to anything else.

HTML defines a small but growing number of constructs which can be used to build up documents of considerable flexibility and power. All these constructs are included in the body of the document and delimited with the < and > characters. Such a delimiter is called a tag. Many of the tags are paired: <X> starts a construct and </X> ends it. This system of tags plus text is similar to other superseded and current systems for marking text to control output like Runoff, troff, and Tex. Another way of thinking about it is that HTML documents are programs, and the client programs 'run' these programs to generate the final document. The range of possible tags covers both structuring elements, and a range of formatting commands. The current version of the HTML tag set, HTML 4.0 has just been released by the World Wide Web Consortium.

Structuring elements govern the logical (as opposed to physical) structure of the document. The two main constructs here are inline images (referred to already) and anchors. Anchors are pieces of text which mark the beginning and/or the end of a hypertext link. They allow links inside a document or to another document. Links within documents are commonly used to provide a table of contents at the start of a long HTML document. The user can jump to a particular section by clicking on an internal hyperlink. Anchors may also be referenced in URLs, allowing links into the middle of documents. Links to another document invoke the full power of the URL mechanism. This means that a single HTML document can refer to other HTML documents on other servers, to Gopher servers, to Usenet newsgroups, FTP sites, and the like. The structuring features of HTML are much more primitive than full SGML allows and are not binding on the author of the document.

Formatting commands allow the designer of a HTML document to control the layout and appearance of the text. The interpretation of HTML documents normally ignores line feeds, form feeds and carriage returns. This requires explicit marking of document formatting. This formatting includes up to six levels of headings, paragraph breaks, various types of lists in-

cluding numbered and bullet points, and character highlighting – bold, italic, monospace text and the like.

The original HTML was not very SGML compliant, although it was SGML-like. With each iteration of the HTML standards it is moving towards closer SGML compliance. The long-term goal is to move HTML into something that is entirely SGML-compliant. This is XML (eXtensible Markup Language).

XML

XML is best thought of as generic SGML delivered over the Web (or 'SGML-Lite'). Its design goals were to provide 80% of the benefits of SGML for 20% of its complexity. The problem is that the fullSGML specification is both hard to implement and more than most Web users need. XML will enable an ISO-compliant subset of SGML to be served, received and processed on the Web. Of course, this will require upgraded servers and browsers to be able to manage documents, their associated DTD's and one or more stylesheets for display. The components of XML are DTD's, XSL, and XLL [Bray et al., 1998].

As with full SGML, the Document Type Definition (DTD) specifies the logical structure (or grammar) of the document. In particular it defines a page's elements and attributes, and the relationships among those elements and attributes. Developers can use existing DTDs or provide no DTDs. In this case the XML parser will only check the document for 'well-formedness'.

The eXtensible Style Language (XSL) specifies style sheets for XML documents. The browser can change the appearance of the document by switching the style sheet. XSL is less complex than SGML's (DSSSL) and provides a subset of its functionality. A mechanical mapping from DSSSL to XSL will be possible.

The eXtensible Link Language (XLL) is a significant enhancement to the linking capabilities provided by HTML, which supports a tiny fraction of all possible hypertextual links. XLL is basically a subset of HyTime (the Hypermedia/Time-based Structuring Language) and will support:

- location independent naming
- bidirectional links
- links specified and managed outside documents
- transclusion (link target appears within link source)

- attributes on links (link types)
- and more!

XML is best suited for applications that:

- require a Web client to mediate between heterogeneous databases
- attempt to redistribute processing load from the server to the client
- require the client to present different views of the same data to different users
- have intelligent Web agents that tailor information delivery to the needs of individual users

The world of the World-Wide Web will gradually make the transition from HTML encoded documents to XML-encoded documents. Support for XML is starting to appear in Web authoring tools and should appear in the next versions of Web browsers.

4.5.5 The Web

> The WWW project merges the techniques of networked information and hypertext to make an easy but powerful global information system. W3 uses the concept of a seamless information space (the “web”), in which all objects including those accessed by earlier protocols (WAIS, Gopher, FTP, etc.) exist. The project allows information sharing within internationally dispersed teams, and the dissemination of information by support groups. Originally aimed at the High Energy Physics community, it has spread to other areas and attracted much interest in user support, resource discovery and collaborative work areas. It is currently the most advanced information system deployed on the Internet. [Foster, 1994]

The World Wide Web (also WWW, W3, W^3, or just ‘the Web’) is the area of fastest growth and most rapid change on the Internet at the time of writing and into the foreseeable future. The most innovative initiatives in information delivery, electronic publishing and electronic commerce are based around the Web and what it offers.

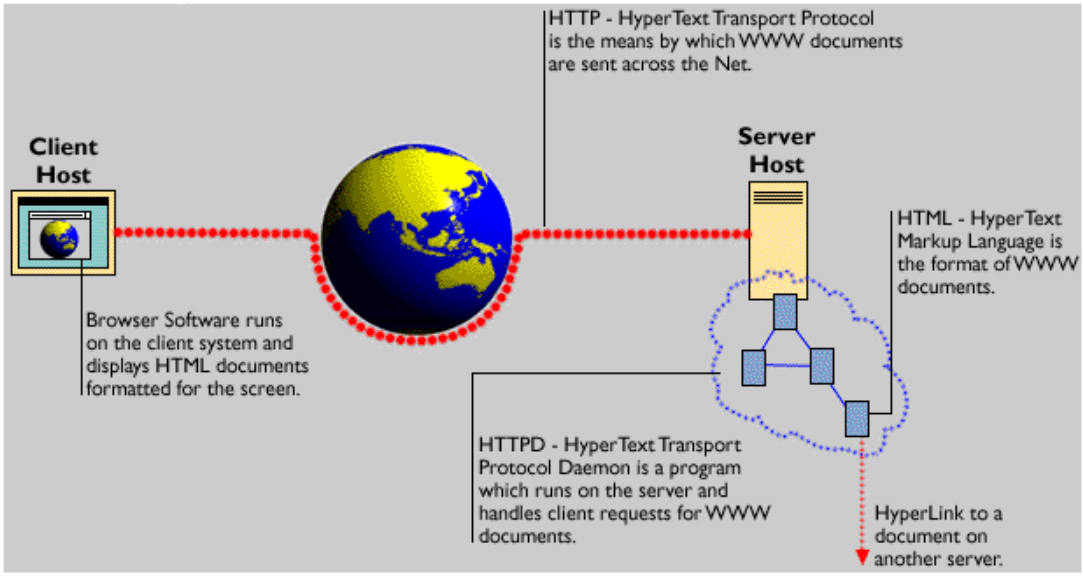
The WWW can be considered as a directed graph of interconnected nodes. Nodes within the graph contain links to other nodes. These nodes are hypertext documents. Hypertext links may point to another location within the same document, or another node (either on the current server, or on a different server altogether). Nodes at the edges of the graph may be stored in many different formats, including plain text, Postscript, graphics (GIF and JPEG), digitised video (Quicktime and MPEG), and sound. Different clients vary in the types of docu-

ments they can display. The Web presents users with a document–centred information space. Links in hypertext documents are represented by bold type, underlining or colour highlighting. To follow a link, a reader clicks it with a mouse (if using a Graphical User Interface client) or moves the cursor to it via the arrow keys and presses Enter (if using a screen–based client). Most clients keep a history of links that have been visited, allowing the user to backtrack their browsing. Some documents in the Web are indexes. The user specifies keywords, and the server executes a search. The result of such a search is another 'virtual' document containing links to the documents found by the search. The end result is that the user sees a seamless web of documents. Navigating through the Web is a matter of starting at the right place or following the right links to get there.

Web clients can be used to access a wider range of information services than just Web content. Most provide Gopher support, UseNet News reading, retrieval from File Transfer Protocol sites, and Telnet to other machines. This enables such clients to be used as all–in–one tools for most Internet work. By providing a News or FTP Universal Resource Locator (URL), the client will connect using that access method, and allow the user to navigate around a news server's message hierarchy or through the hierarchies of an FTP archive.

Web internals,

Figure 4–3: WWW Architecture. Source: Arshad Omari (a.omari@cowan.edu.au)



The Web's architecture is based on the protocol used to communicate between client and server (the Hypertext Transfer Protocol – HTTP), the mechanism used to locate resources (the Universal Resource Locator – URL), and the language used to define the hypertext documents (Hypertext Markup Language – HTML). The interaction between client, server, HTML and URL hyperlinks is shown in figure 4–3.

The HyperText Transfer Protocol is evolving rapidly in response to the demands made by new users of the Web. The HTTP protocol assumes a reliable network connection which is typically TCP/IP and in the context of the ISO OSI Reference Model can be viewed as an application layer protocol. The protocol is stateless, and is made up of atomic transactions. Each transaction consists of:

1. Connection – established by client to server

2. Request – sent by client to server

3. Response – sent by server to client

4. Close – of connection by either party

Requests may retrieve data, store data, create new documents, make or break links to objects, or ask for a search to be performed.

Because there are so many ways to get to networked information resources, The Universal (or Uniform) Resource Locator (URL) mechanism has been adopted. This mechanism provides a standardised way to reference many resources accessible through the Internet. A URL has the general form of `scheme://host.domain:port/path`

Scheme identifies an access protocol or method for the object. Some of the defined schemes are HTTP (the native WWW protocol), anonymous FTP, NNTP (Network News Transport Protocol), WAIS, telnet and Gopher.

Host.domain:port specifies the IP address of the host on which the object resides, and optionally the required port. Most Internet services have well–defined default ports which are used if no port is specified.

Path locates the object in a way that is relevant for the access method. For anonymous FTP, this would include the full directory path and file name under which it may be found. For Gopher, the path would be the menu item hierarchy to be traversed. For NNTP, the path would be the newsgroup and unique ID of the news item. For some schemes, the path may include a search string (or combination of strings) that is/are used to address a ‘virtual’ object formed by searching an index of some kind.

URLs need not be fully qualified. Certain parts of the URL (such as the scheme and machine name) may be omitted, according to well–defined rules. In this case, the URL will refer to

another object, relative to the current location. This facilitates movement of linked sets of objects, provided that their relative positions (for instance within a directory hierarchy) are maintained.

Beyond HTML, Web designers can also provide gateways out to other programs. The Common Gateway Interface (CGI) is intended as a standardised way for programs and information services to interact, although at present only HTTP servers are supported. CGI allows for the transfer of data between an HTTP server and a program written in a conventional programming language. Such a program could, for example, interrogate a relational database and return the results to the server for transmission to a remote client. A number of mechanisms also exist for servers to invoke scripting systems such as Applescript on the Macintosh or perl on Unix systems.

Web browsers

A range of clients exist for most of the popular platforms. A line–mode browser, and the full–screen character–based Lynx client are available for many mainframe and mini–computer configurations. Lynx can only display text. Most Web users opt for one of the desktop client packages. The two browsers that currently dominate the market are Netscape's Navigator/ Communicator and Microsoft's Internet Explorer (both available for MacOS, Windows, and some flavours of Unix).

The Mosaic and Netscape developers are continually leap–frogging each other in the features their software offers.

Web-based interactivity

The Web offers a number of ways in which content creators can add interactive elements to their documents. The two leading candidates at the time of writing are Shockwave and Java.

Shockwave is the generic name for technology developed by Macromedia (creators of the Director software used extensively by CD-ROM designers). Shockwave delivers Director content (and recently content created using the Flash authoring tool) over the Web using a streaming technology. Shockwave applets can provide moving text, digital sound, interactive buttons, branching program options and animation. To access the content, users need a Shockwave plugin installed in their browser.

Java is a platform-independent programming language developed by Sun Microsystems and designed from the ground up to work in a network world. Unlike other programming languages, Java needs an environment to run in. Java *applets* run within a browser 'sandbox', a restricted software area within the browser which restricts what they can access and do. Providing the sandbox is implemented correctly, this should be safe for the user with no possibility of malicious damage or access to confidential data. Java *applications* run with (preferably) OS-level support and have no necessary restrictions; the user can set what the Java environment has access to. Java applets are being used to provide calculation capabilities in Web documents, to enable richer navigation mechanisms, to support embedded simulations and to deliver multimedia content.

4.5.6 Computer-mediated communication

Computer-Mediated Communication has a primary focus on communication between human beings, albeit supported by computer-computer communication along the way. It is likely that for most users, the CMC aspect of the Internet is their primary use, measured in terms of the number of hours they allocate to it. One of the reasons that CMC tools on the Internet are so popular is that people love to communicate. In the scholarly community, the range of Internet tools makes it possible to make links with people one would never have met normally, and may never meet. In particular, it makes it possible to form invisible colleges of colleagues around the world. Such networks may be formal or informal and may use a range of techniques: private email lists, Listservers, newsgroups, MUDs/MOOs and more. However, Lewenstein, in examining the use of bulletin boards during the cold-fusion controversy, argues that "CMC will not replace traditional face-to-face interaction" [Lewenstein, 1995, p. 144].

Electronic mail

Email is probably the most used network application, and for many users is their primary application for the Internet. Email is, of course, not restricted to the Internet. Many organisations run their own in–house email systems, or use one of a range of private email providers. In–house email systems include products like Microsoft Mail, CE's QuickMail and Lotus' cc:mail. Private email providers include MCImail, AT&T Mail and Compuserve. An increasing number of organisations are providing gateways from their in–house systems to the Internet, and most of the private providers provide such gateways. These gateways enable users of in–house proprietary email systems to send and receive Internet email. Email originally provided transmission of text messages only. The development of the Multipurpose In-

ternet Mail Extensions (MIME) standard has enabled a range of other media types (graphics image, digitised movie, program file, etc.) to be attached to a normal email message. A number of gateways provide mail translation across the boundaries between the Internet and internal email systems.

Many mail clients allow the user to attach a file – as opposed to including it. The development of the Multipurpose Internet Mail Extensions (MIME) standard has enabled a range of other media types such as graphics images, digitised movies, documents and program files to be attached to a normal email message. If the email client at the other end knows how to display the attachment, the user will be able to view it onscreen. Alternatively the attachment can be saved to disk for further processing by another program. In this way more complex documents can be transmitted via email. A typical use of this is to attach a formatted word processing file to an email message. The receiver will then be able to open the attachment with their own word processor (assuming the sender and the receiver have used compatible word processing programs).

A typical use of email is for a researcher to maintain an informal list of colleagues around the world who assist with their preparation of articles and conference papers. Early drafts can be circulated for comment, and preprints/reprints sent out once accepted. Of course, scholars have been doing this sort of thing for centuries with print-based information. The difference here is the ease and speed with which it can occur. In fast-moving fields of research, waiting months or years for a paper to appear in a journal may involve an unacceptable delay. The ability to read and cite others work before it has been officially published or presented can be very valuable.

Mailing lists

Listservers provide for automated mailing lists. Typically users need to subscribe to a list via email before they can use it. From then on, any email message they send to the list is automatically copied to everyone else on the list. Such lists can be moderated (someone takes responsibility for vetting contributions and perhaps editing them) or unmoderated (no controls other than peer pressure apply). In excess of 5,000 specialised lists exist to facilitate communication on a bewildering array of topics. Some listserver software allows for automatic retrieval of information via email. The user sends a message requesting a certain file and it is returned by email.

In a more formal way a listserver might be established to enable those interested in a particular topic to correspond. Such listservs either can serve as enablers of existing communities or may act as catalysts to create a new form of ‘virtual invisible college’[Treloar, 1994]. As an example of such an initiative, the National Scholarly Communications Forum was established in late 1993 to bring together parties interested in the effects of the networking revolution on Australian global communication and information access. Its founding members were the Australian Academies (Science, Humanities, Technological Sciences and Social Sciences), Australian Society of Authors, Australian Library and Information Association, Australian Council of Libraries and Information Services, Australian Book Publishers' Association, Council of Australian University Librarians, Committee of Australian University Directors of Information Technology, Copyright Agency Ltd., National Library of Australia, and Council of Australian State Libraries.

Network news

Network newsgroups (also called UseNet News) are another way to enable shared discussions on specific topics. The main differences between news and listservers are that newsgroup traffic is automatically copied to sites that provide news access whether anyone is reading it or not, news disappears as it becomes stale, and that message traffic within newsgroups tends to be grouped into threads (or particular sub-topics of discussion). Newsgroups come into being, evolve and fade away on a time scale of years. Threads arise and disappear on a scale of days or weeks. Special newsreading programs are required to access news. There are upwards of 14,000 globally propagated newsgroups and many more that are specific to particular organisations. Newsgroups can also be used by scholars to correspond with communities of shared interests, although the discussion is not always so focused as with email lists.

Multi-user Dimensions

Multi-User Dimensions (MUDs - originally standing for Multi-User Dungeons) initially evolved for the playing of fantasy-role games set in mythical/fantasy surroundings. They can best be described as text-based virtual reality environments. Users see a textual description of their location and interact with their environment and other users by typing commands. In a MUD, multiple users can simultaneously communicate, interact, and even create new parts of the shared reality. In many ways, a well-populated MUD can be viewed as a virtual community. A good example of such a community is LambdaMOO (MOO stands for MUD - Ob-

ject Oriented). The Uniform Resource Locator (URL) for this MOO is: <telnet://lambda.parc.xerox.com:8888/>.

The facilities provided by MUDs are also starting to be used for serious scholarly discourse, enabling researchers to interact in a natural setting, and work cooperatively to solve problems or produce items of mutual interest. The ability to leave information in a virtual room for others to read when they visit is particularly helpful when cooperating across timezones. A number of examples can be mentioned. Researchers into Post Modern Culture can use the PMCMOO at <telnet:// hero.village.virginia.edu:7777/>, and a MOO for the Institute for Advanced Technology in the Humanities is available at <telnet://hero.village.virginia.edu:8888/>. Biologists are able to hold meetings, discuss research, and critique papers in the BIOMOO. This has a WWW homepage at <http://bioinformatics.weizmann.ac.il/biomoo>. Those without WWW capabilities can use <telnet://bioinformatics.weizmann.ac.il:8888/>. Researchers into Media and related areas can register for admission to the MediaMOO at <telnet://microworld.media.mit.edu:8888/>. This MOO provides a richly textured information space, drawing on the creative talents of a number of the MIT community. Members are encouraged to design their own virtual offices and link them into the shared space.

4.6 Characteristics of electronic documents

An earlier section of this thesis reviewed the characteristics of print documents (see 3.4: Characteristics of print documents on page 54). Electronic documents have a range of defining characteristics that distinguish them from their print counterparts [Schamber, 1996], [Hickey, 1995]. What are their particular strengths and weaknesses?

4.6.1 Strengths

Electronic documents are much easier to move and carry around than print, particularly large amounts of information. For small amounts, print is more portable than the combination of the medium to store it on and the device (even if this a laptop) to access it. Once the amount of print exceeds about 500 pages of A4 paper (about 1.75 MBytes of data), a notebook computer will be lighter. Of course, the electronic documents stored on the hard disk of this computer are not just print-equivalents. They can contain both text and image (as can paper) but they can also contain colour (at no extra cost) and dynamic media such as audio, video and embedded simulations.

Electronic documents can easily be copied and sent from one location to another. If users of the documents go to a server to access them, then the documents can be updated by making the change on the server. The next user to access the document will get the latest version

Electronic documents are (to varying degrees of ease, depending on the format) directly searchable, and can be provided with a range of navigation mechanisms, particularly if hypertext technology is used.

Brown and Duguid also argue that electronic documents (such as e-mail) can be used instead of phone messages to introduce delay and fixity into what would otherwise be ephemeral exchanges [Brown and Duguid, 1995]; one should not only think of electronic documents in terms of mutability and immediacy.

4.6.2 Weaknesses

While electronic document have many advantages, the computer display devices on which these documents are often viewed are significantly inferior to print. Most screens (whether based on cathode-ray tube (CRT) or liquid crystal screen (LCD) technology) offer at best 90 dpi. It is possible to use a technique called anti-aliasing to shade the pixels at the edges of characters and lines but this provides a small improvement only. The contrast ratios between black text on screen and a white background are also deficient compared to paper. In late 1998 Microsoft announced a new display technology called ClearType to work with font rendering on computer screens and improve readability. This is claimed by Dick Brass, Microsoft's vice president of technology development, to “make inexpensive screens look as good as the finest displays, and the finest displays look as good as paper” [Trott, 1998]. Whether this is in fact the case will not be known until it ships (probably some time in 1999), but such an announcement does indicate that software manufacturers recognize that the resolution difference between the screen and paper is significant.

The literature on the relative readability of print versus screen is somewhat confused. One study by Wilkinson and Robinshaw reported significantly higher reading fatigue associated with computer screens relative to paper [Wilkinson and Robinshaw, 1987]. Their results showed a degradation in performance over a fifty minute task. [Valauskas, 1994] also argues for the superiority of print for extended reading. Another study by Gould et. al. [Gould and et. al., 1987] suggests that the quality of the monitors used can remove the fatiguing effects. An article reviewing these studies concludes that

> users do not find reading from monitors intrinsically fatiguing, but that performance levels may be more difficult to sustain over time when reading from average quality screens [Dillon, 1988]

This study also points out that reading is a complex grouping of activities.

> the type of task performed in many of these studies represents … a subset of what is labelled 'reading'. Browsing, light reading and formal studying are probably more frequent interactions with written material [Dillon, 1988, p. 460].

Almost all electronic document technologies have no practical annotation facility whatsoever. The original version of Mosaic (the first WWW browser available for the Macintosh and Windows) had the ability to create either text or recorded voice annotations for Web pages. These annotations were stored locally and displayed in the margin the next time the remote Web page was accessed. The intention was ultimately to provide group annotation features as well. Unfortunately, this feature did not survive the transition of the developers of Mosaic to Netscape and sank without a trace. It is possible to annotate Adobe PDF files, but one needs the full Acrobat Exchange product to do so (rather than the free Acrobat Reader). This means that very few readers of PDF can make use of this facility. Most readers of on-line documents (like the author) therefore have to print out documents in order to annotate them.

Electronic documents need a power source. This is only infrequently an issue for desktop machines but limits portable use. Notebook computer batteries are improving all the time, but 3 hours of usage is still the upper limit for most systems.

Electronic documents also have a number of what might be called system weaknesses. One of the most commonly-cited is the ease with which it can be copied, and therefore the ease with which plagiarism can be carried out. Interestingly, the ACM has been using fast searching and indexing algorithms together with large document stores accessible over the Internet as a mechanism to identify such plagiarism [Denning, 1995]. Another weakness is difficulties in ensuring digital document integrity [Lynch, 1994a]. Thirdly, licensing provisions and copyright mechanisms for multimedia information or for documents that embed multimedia objects are problematic [Lyman, 1995].

4.6.3 Archival issues

E-journals are somewhat different to other digital library archival objects in that they do not need to be digitised. Other objects which began life as paper or film and have been digitised

carry with them a wide range of archiving issues [Conway, 1996]. Objects which begin their life in digital form have as their primary archival focus the challenge of protecting these digits from alteration or loss. This area, called ‘digital preservation’ typically centres on “the choice of interim storage media, the life expectancy of a digital imaging system, and the concern for migrating the digital files to future systems as a way of ensuring future access” [Conway, 1996]. These concerns are dealt with in much greater detail in [Lesk, 1992] and [Task Force on Archiving of Digital Information, 1996].

E-journals as digital objects are likely to consist of text, images and perhaps some attached binary data (compiled computer code, audio, or video). According to [Graham, 1997], digital preservation of such objects consists of three problems: medium preservation; technology preservation and intellectual preservation.

Medium preservation does not (as the name suggests) involve preserving old storage media. Rather, one preserves the information by moving a copy of the digital object from one medium to another (for instance from 1.4 MByte floppy disks to 100 Mbyte Zip disks). [Rothenburg, 1995] points out that traditional media are currently much better candidates for archiving than any existing electronic media. However, provided copying from one medium to the next occurs before the hardware to read the old version becomes obsolete this is not too difficult (the author did not move early enough in one case and now has some binary digits mouldering away and forever inaccessible on an 8” floppy disk).

Technology preservation usually involves taking a digital object in one form and converting it into another (for instance converting a WordPerfect file to a Word file), preferably without losing any information. Again, this needs to be done while the conversion is still possible. Word 6.0 for the Macintosh (or any later version) will not read Word 3.0 for the Macintosh files, as the author discovered when seeking to put his Master’s thesis online.With respect to binary attachments to e-journal articles, the situation is more complicated. As an example, consider a Shockwave file attached to an article in the Journal of Interactive Media in Education (see 6.3.3: Journal of Interactive Media in Education on page 122). In ten years time, it is entirely likely that the Shockwave format (if it still exists) will have changed considerably and that this attachment will need to be run on emulated hardware and software. Old digital objects require old software. Old software often requires old operating systems. Old operating systems often require old hardware. If libraries are not to become museums, then this old hardware and software will need to be emulated.

Intellectual preservation means having the confidence that what is read now is what the author wrote. The ease of copying electronic information carries with it the ease of undetectable change. Mechanisms to combat this (such as embedded digital signatures and checksums) are only now being developed. Adobe's PDF technology also lets the author of a document lock it against changes and/or require a password before opening or modifying it.

From an archiving point of view, the two de-facto standards for e-journal documents, HTML and PDF, both have problems.

HTML is a rapidly evolving standard which is intended to be upwardly compatible. However, documents written to older versions of the standard do not always display correctly with later browsers. Moreover, there are differences between the extensions supported by Netscape's Communicator and Microsoft's Internet Explorer so that documents designed for one may only display partially (or not at all) with the other. This means that in order to correctly view unchanged archived HTML one would need a succession of versions of both browsers. The likely solution to this is a migration of the Web world to the new standard markup language, XML (see 4.5.4: Document oriented solutions on page 75). This appears to be much more robust and extensible and should provide a much better archival base than HTML.

PDF suffers from being a proprietary (although documented) technology. While there are freeware readers, only Adobe products (at this stage) can produce PDF files. If Adobe decides in the future to drop support for PDF this would cause problems for all those journals who have standardised around PDF. However, Apple Computer plans to use PDF as the native image-file format in the next version of its operating system, Mac OS X (10), replacing the current PICT graphics format. This presumably makes PDF less likely to become an orphan technology.

4.7 Conclusion

The technologies discussed in this chapter are now becoming an integral part of the information universe we inhabit. It is important to keep in mind that they were not specifically developed for scholarly journals. Nevertheless it soon became clear to a number of researchers that they could be applied to a number of stages in the process from scholarly author to reader. The next chapter considers the phases in this process and the ways in which the new technologies might be used to support or change the process and outputs.

III Potentials & Responses

“Electronic publishing is the comet and publishers are the dinosaur” [Belanger, 1994], quoting Lorrie LeJeune at the 31st Annual Clinic on Library Applications of Data Processing

5 Potentials and Pressures for Transformation

5.1 Introduction

Section II of this thesis dealt with the two strands of Publishing and Technology. Chapter 3: Print Publishing of Scholarly Journals discussed the evolution of the print scholarly journal to date. Chapter 4: Technology Developments considered the technology developments over the last two decades. These are developments that were largely driven by other forces, but that might potentially be applicable to the communication of scholarly information. Section III now turns to the potentials inherent in the technologies (as well as in the current system of print publication), the pressures on the current system, and some of the developing responses to both the technology potentials and the pressures.

This chapter focuses on the potentials and the pressures for transformation. It aims to describe the potentials for transformation both of publishing functions and stakeholder roles. It also looks at what the pressures are and how they might act on the stakeholders and the functions of the journal.

5.2 Transformation of publishing functions

Turning potentiality into actuality is a large step. Technologies evolve, and users of technology are normally fairly resistant to change. Inertia, particularly with respect to communication technologies, plays a large role. People are not willing to change just for the sake of change – they need a compelling reason. Applying this to scholarly journals means that in order to be successful, any new technology needs to provide either equivalent functionality to print, or if this is not possible, enough alternative functionality to compensate for any deficiencies. For any scholarly publishing medium to be useful to the user, a core set of functions is needed [Treloar, 1996]:

- ability of authors and publishers to produce and format the publication;
- ability of publishers to notify users of new issues of the publication;
- ability for users to access the publication.

Peter Boyce argues for an expansion of this list [Boyce, 1996] [Boyce et al., 1996], focusing on the functions that need to be performed by the publishers and/or librarians. His list of essential components of the publishing process is:

- Author preparation

- Peer review
- Copy editing and typography
- Distribution
- Archiving

Michael O'Donnell proposes five roles [ODonnell, 1995] which he defines in terms of the functions they perform:

- author - writes article
- editor - judges quality and relevance
- publisher - announces acceptance of article and makes it public
- archivist - preserves article
- readers- read accepted text

Boyce does not mention the issue of access. O'Donnell makes no reference to those who perform the refereeing function (at least in this list- they are mentioned later). Because access to a publication includes delivery to the user, presentation of the publication, and navigation through the publication, these need to be considered separately. Conflating these three lists and making explicit some assumptions therefore produces the following set of functions:

- Authoring
- Peer review
- Production and formatting
- Notification of users
- Distribution and access
- Navigation
- Archiving

This section will look at each of these functions in turn, discussing both how they are provided in print journals and how they might be re-implemented in electronic journals.

5.2.1 Authoring

Most authors these days create their articles with the assistance of a word-processor (either directly or through the medium of a secretary). There are undoubtedly a few who still write longhand but they are undoubtedly a minority. Many (although not all) journals will accept submission in electronic form (usually with a fairly restricted) set of choices, although some still ask for multiple copies printed out double-spaced as well (probably to be sent to the referees). A move to e-journals would therefore not involve a large change for most authors.

5.2.2 Peer review

The most rigorous model of peer review (or refereeing) is the so-called 'double-blind' peer review. Submitted articles are sent to two or more reviewers, and neither the reviewers or the author are aware of each other's identity. This is the traditional form of quality control in p-journals and provides a filtering/gatekeeping/censorship function, depending on one's perspective. This process has been criticised for the delays it introduces into the publishing process. Such delay can be broken into five components:

1. the time the paper takes to get from the journal office to the referee
2. the time the paper sits in the referees' in-tray before they begin to review it
3. the time taken until the referees have completed this review process
4. the time the paper takes to get from the referees to the journal office
5. any further time required to resolve discrepancies (if any) between referees' reports

On the face of it, faster communication through email will only be of assistance in those steps involving the transport of the paper (1 and 4). There is some anecdotal evidence that people feel a subtle pressure to respond to email quicker than paper mail, which may assist in step 2 as well.

Email is gradually becoming more and more widespread and in developing countries an increasing proportion of professionals will have access to it. It is quite feasible to attach electronically submitted manuscripts (minus any identifying information if blind refereeing is being used) to an email message to referees and receive responses in the same way. If the referees wish to annotate the received article they have at least two choices. Adobe's Acrobat Exchange product can be used to create annotations to a PDF file. A number of journals are starting to experiment with this. Alternatively, current versions of Microsoft's Word product allow for embedded annotations within a Word document. A detailed proposal for editing a journal through email (although based on ftp and Postscript rather than email and PDF) is contained in [Appel, 1996].

5.2.3 Production

In all cases, the technology chosen by publishers for distribution (see see 5.2.5: Distribution and access on page 96) places constraints on what can be represented and how.

At first glance, print publishing might seem to provide few restrictions; multiple fonts, sidebars and images are all possible. However, hyperlinks within the one publication are clumsy,

and links (footnotes and citations) to other publications rely on the scholar having ready access to the publications linked to. As well, print is limited to information that can be represented statically on paper. Audio and video are impossible. For most publications, colour still images are technically possible, but prohibitively expensive. However, even relatively cheap print journals will have good resolution text.

Due to current display limitations (see 4.2.2: Multimedia facilities on page 59) any electronic format on a screen (as opposed to printed out by the user) will have inferior resolution to print. Electronic formats may also impose other restrictions.

In the electronic world, Listserv (mailing list) archives are usually restricted to documents in 7-bit ASCII. This is because of the need for such documents to pass through email gateways in transit and because no assumptions can be made about the display device at the other end.

Anonymous File Transfer Protocol (AFTP) archives can be used to store any kind of file. In practice, most e-journals using this technology have tended to use 7-bit ASCII text documents. Some journals are storing articles in richer formats like Hypertext Markup Language (HTML), Postscript or PDF.

Gopher servers can also provide a range of document types, but most e-journals mounted on gopher servers also store documents in 7-bit ASCII text. A wider range of Multipurpose Internet Mail Extension (MIME) types is now supported by available gopher clients and servers - the lack of adoption of this facility to distribute documents in other formats is probably being affected by the general rise in popularity of the Web.

World Wide Web documents are written in HTML. As already discussed, this provides for formatted text, inline graphics, hyperlinks within documents, links to other HTML documents, and links to documents in other formats altogether. However, the scholar writing for the Web needs to be aware that a wide range of browsers will be used to access their work. Not all browsers format HTML in the same way, and the available range of markup tags is restricted, particularly compared to SGML. Typography choices are also limited. Thus a lesser degree of control over the final appearance of the document is inevitable, compared to the richness of print.

PDF is a very good solution for complex electronic documents with a high graphical content or lots of formulæ. An example of a stable of e-journals using PDF is the Cajun Project being developed by the Electronic Publishing Research Group [Smith et al., 1993]. Integration be-

tween ‘PDF-space’ and ‘Web-space’ is now very good. The latest version of Acrobat provides for links from PDF files to Web documents and the PDF plugin for popular browsers allows for display of PDF files within the browser window.

HTML provides good navigation mechanisms and works well for onscreen reading, but is poor for printing. PDF versions of the articles are difficult to read on the screen (unless specifically designed for it) but provide high-quality printouts for archiving or annotation by the user. The optimal solution is to use the Web to locate, browse and select the articles desired, and print out the PDF versions if required.

Some publishers are now choosing to create a single database of content from which they can generate multiple representations (SGML, HTML, PDF, etc.) as required for output to print or storage on file/web-servers. Such a ‘neutral’ database can be used to generate complete journal issues, individual articles and current awareness notifications [Campbell et al., 1997].

5.2.4 Notification

In order to access a new scholarly publication, the scholar needs to be notified of its existence. To use current network jargon, such notification can either be ‘push’ or ‘pull’. Push notification is where the scholar receives notification about new publications on the basis of some previous registration of interest. SDI searches are one example of this. Pull notification is where the scholar needs to take the initiative to look for new information, perhaps by scanning contents pages or a publications database.

In the print world, notification is often limited to the physical arrival of a new issue of a journal (often on a semi-regular, predictable schedule). If the journal comes to a library, the scholar has to check the shelves periodically, or rely on some sort of alerting service. Such a service might be provided by the library (e.g. in the form of photocopied contents pages) or a commercial information provider like DIALOG (via the results of an SDI search on a contents database). Alternatively, scholars can directly search on-line databases of abstracts and citations looking for relevant information, but this requires them to take the initiative and can easily get crowded out of a busy schedule.

In the domain of electronic journal publishing, the standard solution to the notification problem is to use one of a number of computer-mediated communication technologies. By far the most popular is electronic mail, with network news a distant second. Two distinct strategies

can be employed. The first is to email the entire text of the latest issue of an journal direct to a scholar's mailbox. In this case, the notification is directly analogous to the arrival of a print journal. An alternative increasingly being adopted is to notify the scholar of the publication of a new journal, include author, title and abstract information, and provide advice on how to access either the entire journal or particular articles of interest. For FTP, Gopher and Web journals, this access information is usually in the form of a Uniform Resource Locator (URL). Most email clients now allow the reader to click on a URL in an email message and automagically load the document in the web browser of choice. This makes for a seamless transition between notification and access.

5.2.5 Distribution and access

Once notified, the scholar needs to be able to gain access to the information. This includes locating the journal, and being able to identify and read articles of interest.

In the print world, the journal has to be physically delivered to its destination. For readers of the journal located outside the country of publication this involves the invidious choice between fast airmail (but expensive delivery charges) or slow surface mail and delays of weeks or months. For some journals (particularly those which are both large and frequent), the cost of shipping is a significant component of the total subscription cost. If the journal is delivered directly to the user, the problem of journal location is limited to finding the journal within the context of the scholar's own personal information management system. If the journal is delivered to the library, it will be filed in some well-defined sequence. To assist with locating articles within journals, the publishing industry has developed a range of standard tools: contents pages at the front of issues, yearly cumulative printed indexes, and the like.

For e-journals, distribution and access are quite different. In the first place, for most e-journals nothing is actually delivered. The exceptions are those e-journals that still use CD-ROM distribution and so require the movement of a physical object. For all other e-journal types, either a sequence of bits is delivered automatically or the user connects to a server to access the content of an article or issue.

Listserv archives enable scholars to access information via email. All that is required is to email a GET command to the listserver address requesting that a specified file be sent by return email. As email is the lowest common denominator for users of the Internet, this provides the widest possible audience. As an example, consider the reference in this thesis to Harnad (1991). This article in the refereed e-journal *Public-Access Computer Systems Re-*

view (PACS-R) can be retrieved by sending the email message `get harnad prv2n1 f=mail` to listserv@uhupvm1. Of course, before issuing a GET command, one needs to know that the file exists. Some journals, including *PACS-R*, handle this by sending the table of contents and abstracts to users subscribed to the PACS-L or PACS-P mailing lists. Alternatively, it is possible to email commands to some listservers instructing them to search a database and return a list of articles that match the search criteria. These articles can then be retrieved as above. It is also possible for the entire issue of a journal to be delivered by email.

Scholars can access articles in anonymous FTP archives either by using a dedicated FTP client, or by providing an FTP URL to a Web browser like Lynx, Mosaic or Netscape. If the URL formalism is not being used, then the FTP location of the article will need to specify host machine, directory path and filename. For example, the information encoded in the URL FTP://cogsci.ecs.soton.ac.uk/pub/harnad/Harnad/harnad95.quo.vadis can also be expanded into (more or less) plain English as 'Make an anonymous FTP connection to cogsci.ecs.soton.ac.uk, move into the directory pub/harnad/Harnad/ and get the file harnad95.quo.vadis'. The URL formalism has the advantage of being more compact as well as parseable by both humans and machines. One example of a journal accessed by AFTP is *Psycholoquy*, edited by Stevan Harnad.

Gopher was initially developed to provide a basis for mounting Campus Wide Information Systems (CWISs). It is based around the idea of hierarchical menus, and allows the server administrators a lot of flexibility in how they structure their information space. One fairly standard way to mount e-journals on a gopher server is to have a menu of possible journals. Each journal points to a menu of issues for that journal. Each issue points to the individual articles. Given unambiguous information about the path to be followed, scholars can navigate through the menus until they locate the files they want. It is also possible to provide Gopher URLs for direct access using a Web browser. An example of a journal available through Gopher is the *Mathematical Physics Electronic Journal*.

The Web, with its non-hierarchical document-based networked hypermedia architecture provides a much richer environment for electronic publishing. Documents can either be reached by following an existing link, or can be accessed directly by entering a valid URL. Documents can in turn refer to other documents and provide direct links to them (something that is not possible with documents accessed using a Gopher client). The Web can also be used to point to documents in PDF format (as discussed already)

5.2.6 Navigation

A range of navigation mechanisms are possible. This section will consider first the question of navigation granularity and then browsing and searching.

Navigation granularity

Navigation granularity can be defined as the smallest item that can be navigated to or cited. A range of increasingly fine granularities is possible, depending on how much information has been provided: Journal, Issue, Article, Section. Beyond the level of a section within an article, there are some difficulties.

The most common navigation mechanism in the print world is page numbers. This presupposes that the page cited will be the same in all versions of the work referred to. This is a safe assumption as long as editions and reprints are carefully flagged in the citation. E-journals in general have no direct equivalent to page numbers. Some first-generation e-journals, such as PACS-R introduced 'page numbers' every 60 lines into their ASCII articles. Unfortunately, these page numbers did not align with those in the parallel print incarnation of the journal. This meant that locating a particular cited page required the citation to include which of the versions of the article was intended. PDF-based e-journals usually inherit the page numbers of the print versions they replicate and so do not have this problem. HTML-based e-journals cannot have page numbers because the concept of page is meaningless in a user-reformattable document space. Changes in font size, browser window dimensions or paper used for printouts will all affect the size of pages. Some journals are experimenting with making the citation granularity more useful and introducing either paragraph numbers (long the practice in legal circles) or defining visible HTML anchors (which can then be linked to) for each section heading.

Chemical Journals Online (a commercial full-text service run by the American Chemical Society using STN network which includes all of the ACS journals as well as those of several other major publishers of chemistry) numbers the paragraphs of the online versions of articles. This allows users of its service to just look at those paragraphs referenced in a search set.

Browsing

Browsing through print is actually a very sophisticated set of activities, many of them unconscious (as noted in [Olsen, 1994]). Users make use of tables of contents, highlighted articles

on the front cover of a journal, known ‘favourite sections’ of a particular journal, flicking quickly through all the pages, skimming abstracts and authors names, consulting the index at the end of the issue (if it exists) and so on. These are skills acquired over years of use of these particular communication artefacts and are, in part, tied to the affordances of that artefact. The serendipitous nature of browsing (as opposed to searching) was an important factor for a number of Olsen’s respondents.

Browsing in the electronic world is constrained by the inability to quickly ‘flick’ through content in the same way, and by the similarity in appearance of many online journals. Browsing mechanisms therefore have to use the affordances of the online media in different ways. Examples include use of images of front covers to select a particular issue, hyperlinked tables of contents, and hierarchical arrangements of content by year, then issue, then article.

Searching

Searching print on a small scale can be done manually. Skimming an article is partly a process of searching for words or phrases (although these may not have been decided on before starting). For larger amounts of text, the user needs to work with document surrogates such as indexing or abstracting services, usually in some computerised form.

The searchability of e-text is of course one area where it provides significant additional functionality over print. Because the entire content of an online journal is already machine-readable, it can be indexed relatively easily (depending on its format) allowing for free-text searching across issues or even sets of journals.

5.2.7 Archiving

Archiving is usually performed by libraries not by publishers. In fact, most publishers regard their responsibility as having ended once they have shipped each issue of a journal. Paper has its difficulties as an archiving medium (acid-decay, flammability, storage requirements) but they are familiar difficulties and techniques for dealing with them are well known.

The new digital media on which electronic publications are stored (both on servers and on user’s own workstations) pose a range of challenging problems to librarians and archivists needing to store them for long periods. The Commission on Preservation and Access has devoted considerable resources to this problem. Some of the issues raised in their best-known report on the subject [Lesk, 1992] are:

- technological obsolescence is a greater danger than deterioration

- devices (and their associated media) may disappear for business reasons
- many of the technologies involve some kind of format, as well as a physical device.
- it is not possible to detect the state of the medium via mere physical inspection
- format, software and hardware are often intermingled: information may be preserved but if the software to print, search, and edit it has gone, it may be quite costly to make any use of it.

On the plus side, digital media can

- allow for very fast retrieval of archived information
- store information very compactly
- allow for easy copying of information for preservation.

These characteristics will require all those involved in the scholarly journal ecology to re-think their role with respect to archiving.

One initiative that may point the way to a class of solutions to e-journal archiving is Electronic Collections Online (ECO) from OCLC. ECO has the following characteristics:

- Publishers have granted OCLC archival rights to journals in the collection.
- OCLC builds and maintains a subscription profile for each participating library.
- A library's rights to the archive remain active regardless of their current subscription status.
- OCLC will ensure that they migrate the service to provide ongoing access.

Mindful of ongoing concerns about future access, OCLC is also working to ensure that libraries will have access to their journals even if Electronic Collections Online is discontinued.

5.3 Transformation of stakeholder roles

As well as the potential transformations in publishing functions, the new technologies also make possible transformations of stakeholder roles within the scholarly journal ecology.

5.3.1 Scholars

Scholars are now able to perform all the publishing functions discussed [Boyce, 1996]. In particular, they are able to publish their ideas without going through the gatekeeping functions of a journal because the barriers to publishing are so much lower (indeed almost invisible) in an online environment. Critics of this possibility point darkly to the early days of

desktop publishing as illustration of the dangers of enthusiastic amateurs. Others mutter about ‘vanity publishing’ being out of control on the Web. The biggest problem for such self-publishers may in fact be the problem of getting noticed by their audience amidst the hundreds of millions of other Web pages. The new economic barrier to being read may in fact be the attention barrier [Goldhaber, 1997], [Lanham, 1994].

5.3.2 Scholarly Societies

Scholarly societies are probably less able to take on new roles due to advances in technology. Thanks to lower barriers to entry they may be able to publish themselves electronically rather than subcontract this to a commercial publisher. They are also able to do better what they already do and provide greater functionality to their members or perhaps to reduce the price of receiving the journal (although by how much is a matter of debate).

5.3.3 Publishers

Publishers are also able to reengineer their processes to provide greater functionality to their customers and perhaps reduce the price they charge them (or provide higher returns to their shareholders). Publishers are also able to start thinking about managing a constantly growing database of digital content, rather than just publishing issues of print journals [Boyce, 1996].

5.3.4 Subscription agents

The new technologies will also help subscription agents to improve and expand their services. Access to databases of content will allow them to provide tailored collections of articles rather than entire journals. New Web-based information architectures will also let them act more effectively as an interface layer between publisher and library.

5.3.5 Libraries

Librarians are now able to broaden their roles beyond access and archiving to become publishers themselves and provide greater access to information for their users. This theme will be taken up in greater detail later (see 8: Library Case Studies on page 177). O’Donnell argues that the role of a publisher may be able to be performed by a library organisation if scholarly journals are electronic [ODonnell, 1995, p. 195]. [Butler, 1986] suggests that libraries may find themselves moving towards one or both of two possible extreme positions: libraries without journals or journals as libraries.

5.4 Pressures for transformation

Despite the obvious advantages of the print journal, over the last twenty years there has been a gradual rise in dissatisfaction with the current system of producing and providing access to these journals.What are some of the most serious pressures for the transformation of the current system?

5.4.1 Journal economics

One of the most significant difficulties with print journals from the point of view of libraries has been the increase in serials subscriptions prices. The Mellon report summarised by Ann Okerson in [Okerson, 1996] found the following picture in the U.S.:

- serials prices have consistently increased at nearly double inflation, with STM material being the worst offenders
- the most expensive serials show the largest relative price increases
- foreign publishers apparently put prices up to compensate for relative exchange rates but did not adjust the prices when exchange rates changed
- three European commercial publishers (Elsevier, Pergamon and Springer) accounted for “43 percent of the increase in serials expenditure at one university between 1986 and 1987” [Okerson, 1996, pp. 188-190].

Some of this increase in cost is due to the very healthy profits earned by a few large publishers in markets characterised by near-monopoly positions and little competition. Estimates for the latest financial year (based on published information) show Return on Equity (ROE) running at 41.7% for Wolters Kluwer and 28.2% for Reed Elsevier [Wyly, 1998].

The impact of these pressures has been a trend towards decreased subscriptions for titles and a corresponding increase in the cost of these titles for the remaining subscribers. Unless unchecked, this will lead to what has been described as the ‘downwards spiral of death’ as fewer and fewer subscribers will have to pick up a greater share of the cost of production.

According to [Agre, 1995a]:

> All types of fixed costs of consumption can raise distributional questions when they are high… This is particularly true when media that have high fixed costs of consumption (e.g., television or networked computers) compete against media that have high fixed costs of production (e.g., newspapers or books). As the latter lose their

needed economies of scale and are forced to distribute their fixed costs among ever-fewer units, they will consolidate among themselves and may ultimately collapse.

In addition to the initial subscription costs, there are the ongoing costs associated with a serial subscription. Robin Peek has summarised the incongruous nature of the current system particularly well:

> In this system libraries purchase and store paper that may never be read. Then they bind the journal, perhaps purchasing it again in microfilm, all of which can cost more than the original journal subscription. Then they spend even more money to house the title and maintain it until, in the not-so-distant future, the paper is about to crumble in their hands, and then they decide whether to spend more money to preserve it [Peek and Newby, 1996, p. 9].

Dyson argues much of the above discussion is in danger of being made obsolete by the new technologies and that in the new economic environment of the net the physical manifestation of content is irrelevant. What matters is intellectual processes and services, not intellectual assets and property [Dyson, 1995].

5.4.2 Problems with refereeing

The most rigorous model of peer review (or refereeing) is the so-called ‘double-blind’ peer review. Submitted articles are sent to two or more reviewers, and neither the reviewers or the author are aware of each other’s identity. This is the traditional form of quality control in p-journals and provides a filtering/gatekeeping/censorship function, depending on one’s perspective. There are a number of criticisms of this model:

- It is essentially a binary process. The articles are either accepted or rejected (acceptance with changes can be considered ultimately in the same way as acceptance). The journal does not attempt to rank the quality of the articles published. There is an implied ranking between journals but not within (except between refereed and un-refereed sections).

- It is anonymous. There is no way for the rest of the scholarly community to observe the process and its affect on the final product. There is also no way to effectively challenge a negative referee’s report that may have been motivated by malice or professional envy. There is also no overt reward to referees for their efforts and often significant contributions [Gaines, 1993].

- It may involve either errors in the judgement of the referees or conscious or unconscious bias towards the article or author [Meadows, 1977].

- It may well involve referees in a conflict of interest. Fjällbrant has argued that because of today's narrow specialisations, referees will almost certainly know those whose work they are refereeing. Given the competitive nature of modern research, "the temptation to delay or downgrade information from a parallel research team is obvious" [Fjällbrant, 1997]

- It stops once the article has been accepted. There is no real opportunity for ongoing discussion of an article within a scholarly community. Letters to the editor often appear months after the original article and the discussion is limited to one round of comment and response.

- It is reliant on individuals who are usually eminent in their fields and hence have very busy schedules. It is therefore frequently prone to delays. An analysis in suggests that under reasonable circumstances authors of articles will end up 'owing' the journal many more reviews of other articles than most authors are prepared to accept [Gaines, 1993, p. 151].

- Articles that challenge the accepted wisdom have real difficulties in getting published because of the conservatism of the referees (or because they require extra space to expound their non-standard assumptions [Gaines, 1993]). It has been suggested in fact that the refereeing system may actually suppress new ideas [Zuckerman and Merton, 1979] rather than only allow the best ideas through.

5.4.3 Delays to publication

Delays of up to one year between initial submission and appearance in print are routine in many fields with evidence of a three year delay for the proceedings of a conference [Hayes, 1989]. *The Journal of the American Society for Information Science* (JASIS) in common with many journals shows at the start of each article when the article was first submitted and when accepted. A common delay between final acceptance and appearance in print is 12 months. Some of this delay is no doubt due to the processes of preparing the article for publication, but the rest is probably caused by the need to defer printing the article until an issue is available with free space. [Peek and Newby, 1996] argues that some of the delay is caused by the uncertainties inherent in the refereeing process. Without any firm idea of when an ar-

ticle will be returned by a referee, and with no control of the submission of articles to the journal in the first place, the journal is forced to maintain a backlog of articles working their way through the production pipeline to ensure a regular printing schedule. Of course, in the worst case, an issue of a journal may be delayed until it is full.

In fields with a significant information technology component, 12 months can encompass an entire generation of technology and is simply unacceptable. Fields like computer science have a strong tradition of refereed conference proceedings for this reason, and the high-energy physics community has always made extensive of rapidly circulated preprints [Ginsparg, 1994].

5.4.4 Limited interaction

The possibilities for interaction with the author of a journal article are extremely restricted. The reader can write a letter directly to the author, but this correspondence will not be visible to the other members of their scholarly community. Alternatively, the reader can write to the editor of the journal and hope to get their letter published. Unfortunately, space constraints and the delays in the publication process already discussed mean that there will be a lag of at least one issue before the letter appears. This may well be three months by which time the urgency of the response to the article has gone. Scholarly debates that take place in this way can last quite a long time and do not have the immediacy of a question session after a conference paper (for example).

5.4.5 Loss of ownership of knowledge

The traditional way for scholars to get published (in the non-trade press, which is the overwhelming majority of scholarly publishing) is to ‘give’ their content to the publishers by assigning copyright to them for no monetary return. The publishers then add some value to the content through editing, distribution and promotion and sell it back to the universities. In effect the universities are paying the publishers for their own content. A number of commentators have been expressing concern about scholars losing control over their own information in this way for some time [Okerson, 1991], [Thompson, 1988], [Guédon, 1995], [Metz and Gherman, 1991].

5.4.6 Need for associated intermediary processes

The size of the journal literature brings it with it significant problems of access. Authors want to ensure that their work is read, and preferably by those who are most interested or affected

by its content; usually their own invisible college. Readers want to find the most relevant articles among all those continually being produced. Unfortunately, it is not possible to search the print literature directly. This has given rise to a wealth of secondary indexing and abstracting services which allow researchers to search document surrogates. These services are delivered on CD-ROM and/or online and are a significant industry in their own right. Having a journal listed in the right service is almost essential for access to readers (and hence acceptance by authors). The new (in the early 1990s) *Online Journal of Current Clinical Trials* (OJCCT) [Keyhani, 1993] suffered a significant credibility gap until it was indexed in *Index Medicus* and *Medline* (the two premier indexing services for the medical community) [Peek and Newby, 1996, pp. 7-8].

5.5 Conclusion

This chapter has shown that the new publishing technologies enable both the transformation of existing publishing functions and stakeholder roles. The pressures towards, and the potentials for, transformation have grown in parallel as the technologies have made more and more possible. What have been the responses to these potentials and pressures? This is the topic of the next chapter.

6 Developing Responses

6.1 Introduction

This thesis has already discussed some of the pressures on print scholarly journals (see 5.4: Pressures for transformation on page 102). Chapter 4: Technology Developments reviewed the range of technologies for working with information (particularly print-intensive information) that have become available during the last two decades and Chapter 5: Potentials and Pressures for Transformation considered the potentials inherent in these technologies.

Neither of these chapters dealt specifically with particular solutions. This chapter tells a story that both runs in parallel with the developments outlined in chapters 4 and 5 and makes more specific some of the possibilities already presented (see 5.2: Transformation of publishing functions on page 91). It deals with the initiatives that evolved in response to the pressures already discussed, and shows how these initiatives built upon (and in some cases anticipated) the technologies discussed in Chapter 4 as they developed and became more mature.

The particular responses discussed are electronic text journals, electronic paper journals, the addition of hypertext features, multimedia enhancements to journal articles, the consequences of network distribution, increased interaction capabilities, the opportunity to reengineer scholarly communication, workflow support for the creation of scholarly journals and new economic models for journals. Where appropriate, illustrative examples are provided for each response. While these responses are discussed in rough chronological order, it should not be assumed that they developed in lock-step fashion. In most cases there is considerable overlap, and some journals provide a combination. It should also be kept in mind that many of these responses are still developing over time as the environment they operate changes around them and as they interact. The chapter concludes by examining three leading-edge e-journals as examples of what is already possible.

6.2 Responses

6.2.1 Early stirrings

Even comparatively early in the development of mainframe computers researchers were starting to think about their possible use to enhance scholarly communication. In 1972, Bamford proposed a system based around editorial processing centres and using OCR and mag-

netic tape technology [Bamford, 1972]. Somewhat later Senders proposed an alternative model (without apparently being aware of Bamford's work) using paper tape or magnetic tape and some unspecified networking technology [Senders, 1976], [Senders, 1977 - now citing Bamford]. Four years later, having worked with an early and very experimental system, he famously wrote "I have seen the future and it doesn't work" [Senders, 1980]. As early as 1982, some commentators suggested (fairly presciently) that it would be at least a decade before electronic journals would substantially supplement print [Turoff and Hiltz, 1982]. The problem with these early systems was that the technology simply wasn't ready - [Senders, 1980] contains a positive litany of complaints. It wasn't until the development of better networking and display technologies in the late '80s that it was possible to start realising some of the earlier ideas. For an overview of the evolution of electronic publishing, see [Lancaster, 1995b]

6.2.2 Electronic text

The earliest e-journals produced can be described as *electronic text* only. These were journals that were restricted in their formatting to 7-bit ASCII (128 characters in all). Because of this, only very restricted formatting was possible, with only character-based graphics. These journals were (and still are) distributed by automated mailing list software, or stored on ftp, gopher or Web servers for retrieval.

Such journals had the advantage that they were directly searchable and had very low bandwidth requirements (and entire issue might be less than 100K in size). A number of these publications, typical of pre-Web e-journals, still survive today (although they may have added other facilities in the meantime). A good example is the *Public-Access Computer Systems Review* (PACS-R) which first appeared in 1990. It is distributed via the PACS-L mailing list (to join, send the following email message to `listserv@listserv.uh.edu`: `SUBSCRIBE PACS-L First Name Last Name`), as well as being available on the Web at `<http://info.lib.uh.edu/pacsrev.html>`.

6.2.3 Electronic paper

Electronic paper describes the use of publishing technologies to provide a similar appearance and functionality to traditional print publications. The two main technologies used for this purpose are page images and PDF.

Early e-journal projects (CORE, TULIP) used scanned page images stored as TIFF files as one (or the main) format for viewing. The problems with this approach are numerous. The page images are large (1 Megabyte/page for uncompressed A4 pages at 1 bit/pixel, more for higher bit-depths in greyscale or colour) and therefore take significant amounts of bandwidth to move around the network. Displaying such images also takes an appreciable time on all but the most high-end workstations. 'Flipping' through such pages is like swimming through treacle. Finally, the text of the images is not directly searchable, as it is represented internally only through a pattern of bits.

The most widely used current technology for providing electronic page images is Adobe's PDF/Acrobat. PDF has the significant advantage over HTML that it can cope easily with complex page layout (such as multiple columns and text wrapping around graphics), and large amounts of mathematical formulæ. It is however often hard to read on screen, particularly if the PDF has been created for portrait orientation pages and is being viewed on a landscape orientation screen. [Kasdorf, 1998] argues that PDF alone is an inadequate solution and the e-journal publishers should use both PDF and SGML.

Acrobat is an appropriate way to provide fairly low-bandwidth delivery of page images that will print at the highest available resolution of the user's output device. It is often possible to generate the PDF files at little extra expense as part of the p-journal production process, particularly if this process already creates Postscript files on the way to the print device. The Postscript files can be automatically converted into PDF using Adobe's Distiller program.

PDF also has advantages in other areas. It can be used to 'print' electronically things that would be prohibitive to produce on paper, or that change frequently. Publications containing large amounts of colour graphics, or long publications with a limited market (such as conference proceedings) are good examples of the former. Catalogues and internal policy manuals are good examples of the latter. PDF in this case can be used to support 'print on demand'.

PDF also allows publishers to produce exactly parallel electronic and print versions of a journal. This may well prove to be quite important while we remain with parallel versions. From the user's perspective receiving a PDF file just means that they now have the *option* to print it out rather than being *required* to get it in paper.

An example of a stable of journals provided only in PDF format is the nine journals being made available by the CAJUN (CD-ROM Acrobat Journals Using Networks) project and

available on-line at <http://cajun.cs.nott.ac.uk/> and described in [Smith et al., 1993].

6.2.4 Hypertext articles

The addition of hypertext technologies to e-journals provides a significant improvement in the features provided and the useability of the journal. The journal can now provide linkages

- within articles (i.e. footnotes)
- between articles on same server (i.e. tables of contents)
- between articles on different servers (i.e citations)
- to other document types (i.e video, audio, etc.)

This was very hard to do before the advent of the Web, but is trivially easy to do now. Because the hypertext functionality is only available on the screen, such articles should be designed to be read on the screen. In particular, they should take into account its landscape orientation, rather than the portrait orientation of paper pages. Hypertext linking allows an e-journal to have a rich set of links within, between and out of articles.

Such links might include navigation links:

- a table of contents for the journal issue
- a table of contents for each article
- links to enlarged versions of figures from thumbnails

They might also include citation links:

- external links to cited articles (backward references)
- external links to articles which cite this article (forward references).

This latter feature needs to be deliberately added because the architecture of the Web does not directly support bidirectional link traversal. The addition of forward references dramatically increases the usefulness of the article and situates it in a web of scholarship.

PDF now supports hyperlinks internal to a document and external hyperlinks to other PDF files. It also supports links to Web documents through the Weblink plugin.

Examples of these linking mechanisms will be provided when discussing the exemplar e-journals (see 6.3: Leading-edge examples on page 116).

6.2.5 Multimedia enhancements

Moving to electronic content rather than printed content also allows for a range of multimedia content that is expensive or impossible to provide in print.

A good example of something that is possible but expensive in print is colour. Once one moves past the cover of most scholarly journals, one rarely encounters any colour (either in graphics or photographs). This has led to a long tradition of multiple barely-distinguishable crosshatch patterns to differentiate variables in graphs. For some applications, colour is nearly essential to an accurate description of the research and the lack of it causes significant difficulties and circumlocutions. In an on-line environment, colour is nearly free: there is some increase in transmission speed due to larger file sizes and those user without colour screens (a shrinking minority) will be unable to view the colour but these are fairly minor problems. It is true that the hundreds or thousands of colours available on most display screens are inadequate for some demanding image requirements but these are fairly rare in scholarly communication.

Examples of things that are impossible in print are links to digital objects such as sound or video files as part of an article. Sound can be used for music clips (for an example of this see [McNeilly, 1995]) or for phonetic samples in a linguistics journal. Video can be used for performances (for an example of this see [Magrini, 1995]) or visualizations of complex phenomena [Mustard, 1994]. Extending the concept of document to include such objects (or even collections of such objects) is an ongoing research challenge [Furuta, 1995].

6.2.6 Embedded simulations

On-line journals can also contain embedded simulations as part of an article's content. For example, an article discussing a piece of software might include a simulation of that software that the reader could interact with. An economics article might include an embedded interactive model to illustrate the relationship between variables. These simulations can be provided by using a back-end server running Common Gateway Interface (CGI) scripts, embedded Java applets or Shockwave. For an example of the latter in action, see [Durbridge and Stratfold, 1996].

6.2.7 Increased interaction

An on-line publishing environment makes it possible to dramatically improve interaction between the reader and both the author and other members of the community interest. At the

simplest level, an article can contain the author's email address enabled so that the reader can send a message by selecting a link. A more sophisticated option is the ability to make available on-line commentary on an article together with the article. In this case, the article becomes the focus for scholarly dialogue much more so than at present. It is also possible to have a link from an article to scheduled on-line live chat session.

For an example of a journal that uses peer commentary see 6.3.3: Journal of Interactive Media in Education on page 122.

6.2.8 New models for peer review

On of the criticisms of the peer review mechanism is its lack of transparency (see 5.4.2: Problems with refereeing on page 103). Two main alternatives have been proposed in an on-line environment: open peer review, and scholarly skywriting.

Open peer review comes in a variety of flavours [Odlyzko, 1994], [Peters, 1995]. One version suggests that everything that is submitted to a journal should be regarded as published, but with open peer review providing an ongoing ranking process [Nadasdy, 1997]. This would provide for a wider assessment of the quality (and hence status) of an article. One can imagine authors withdrawing from publication articles which had attracted too much criticism. Quality control would still exist but would be less of a pass/fail ranking. [Roistacher, 1978] even suggests a system where referees provide a numerical ranking between 1 and 100. His proposed 'virtual journal' would then "publish all papers submitted, but would also allow readers to treat papers as if they were published in a series of journals of differing prestige" (p. 20). A much more elaborate system with some similarities called 'consensus journals' has been proposed in [Stodolsky, 1995]. This aims to eliminate the editor in favour of a mediator and use anonymised reviews based on agreed dimensions.

An alternative is scholarly skywriting, proposed by Stevan Harnad [Harnad, 1990], [Harnad, 1995a]. This proposes a model where after a journal has accepted an article and had it refereed, it is then circulated to commentators around the world. They are invited to submit critical commentaries to which the author will respond. Each article is then co-published with the commentaries and response. Harnad envisages a version of this model but not tied to print production systems radically transforming scholarly communication. For an entertaining and vigorous debate about this model see (in sequence) [Harnad, 1995c], [Fuller, 1995b], [Harnad, 1995d], and then [Fuller, 1995a].

6.2.9 New economic models

There is a very large literature dealing with the economics of scholarly journals, much of which predates the arrival of e-journals.With respect to the possibilities for changing the prevailing economic models, the best known protagonist is Stevan Harnad. His economic arguments against the current system of economics for publishing scholarly research are detailed in [Harnad, 1991], [Harnad, 1995b], [Harnad, 1995a], [Harnad, 1995d] and summarised in [Brent, 1995]. Harnad's views on the economics of on-line versus paper are neatly (if somewhat tersely) summarised in [Okerson and ODonnell, 1995']:

> In brief: Paper means substantial expense. Substantial expense means copyright protection. Copyright protection means fees. Fees mean 'protection' of the scholar's work from nonpaying eyeballs. THAT is precisely what the scholar does NOT want. (p. 33)

In reviewing the recent literature, much of the debate surrounding journal economics revolves around the relative proportions of first copy costs versus the incremental copy + distribution costs for journals in general and e-journals in particular. This has come to be characterised as the '70:30' debate. Harnad's view is that the ratio is 30:70 and that therefore considerable savings are possible by moving to electronic journals only and saving the distribution costs. The consensus view of most publishers is that the ratio is at best 70:30 and that therefore little saving is possible [Marks, 1995]. In a carefully worked out case study, Malcolm Getz shows that the costs for printing and mailing (both fixed costs and variable costs) for the *American Economic Review* are 38% [Getz, 1997]. [Lynch, 1994b] points out that one should also include what he calls "marketing and fulfilment costs" (p. 24) or the costs of doing business.

In fact, these two apparently diametrically opposed positions may not be analysing directly comparable processes. The costs inherent in a system designed from the start for electronic processes and output are very different to those for a system structured for print that is adding electronic output on as an option. Indeed, most existing print journals that are adding parallel electronic delivery are proposing 10% *increases* in subscription charges.

Phil Agre's analysis [Agre, 1995b] argues that the situation with regard to economic models for new media is in fact significantly more complicated than a simple analysis of first copy versus incremental copy costs. He argues that both the costs of distribution and consumption have fixed and marginal components, and that all combinations need to be considered.

The *fixed costs of distribution* are the costs necessary before distributing the first copy of an e-journal. If we assume network distribution (rather than, say, CD-ROM) then much of this is already in place in the developed world. Where it is not in place, it will probably be rolled out for other reasons not associated with electronic publishing. This will allow epublishing to piggyback on other digital services.

The *marginal costs of distribution* are the costs associated with each successive copy of an e-journal. In common with most information products, these costs will be low because of the ease of production. If no physical e-journal is being distributed then copying the bits and sending them down a network is a very low cost operation.

The *fixed costs of consumption* are the costs to acquire the hardware and software and the skills to use both of them in order to access e-journals. Because both machine costs and skill-acquisition costs will be largely paid as part of other activities (need for a networked computer to access things other than e-journals, need to learn a Web browser to access the corporate intranet), these costs will be low for e-journals. For nonprofessionals, these fixed costs of consumption may raise a barrier to access. They are unlikely to be an issue for the main target audience of an e-journal.

The *marginal costs of consumption* are mostly the price of the e-journal (assuming it has one). It may also be necessary to factor in access time and usage time. Access time for an e-journal that is delivered to a scholar's desktop is lower than a journal that requires a visit to the library and much lower than a journal that can only be accessed via document delivery. Of course, a document delivery service that automates the request process for the scholar will only then involve the inconvenience of waiting for the document and re-invoking the scholarly processes that were interrupted by the need to request the document in the first place.Usage time can be quite significant if the technology or user-interface for the e-journal act as a barrier to efficient use. As a minimum starting level an e-journal should be *as least as easy* to use as a print journal.

No matter what the publication technology, management of the publishing process is a necessary activity if one wishes to achieve acceptable and reliable quality, and this management will cost ((someone) money. Fytton Rowland, coming from a background of 25 years experience in working for not-for-profit learned-society publishers, argues that:

> while a journal publishing 15 papers a year could be run on an "amateur" basis, one publishing 1500 papers a year cannot, regardless of the medium it is published in.

> The sheer administrative load of organizing the input, refereeing, copyediting, formatting, and distribution of that many documents (including the ones that get rejected, which generate work too) requires full-time staff. And since these people have to eat, they need a salary. Contrary to what some participants in discussions of electronic journals have alleged, it is this area of "first-copy cost" that is responsible for most of the cover price of a journal, not the paper, printing, binding and postage costs. Yes, a purely electronic journal is inherently somewhat cheaper than a paper one; but not a tiny fraction of the cost. [Rowland, 1994]

Karen Hunter, Senior Vice President of Elsevier Science has similarly argued that "It is expensive for a publisher to make the transition for existing journals from paper to sophisticated, robust electronic publications that smoothly link with those of other publishers" [Hunter, 1998].

There is little available internal cost data from existing publishers; they (rightly) regard this as commercially sensitive information. There has been some research to try and quantify these costs. One of the foci of Project ELVYN was to examine exactly what the costs of would be to libraries and publishers in providing an electronic version of an existing print journal [Rowland et al., 1995] [Meadows et al., 1995]. This project found that the "effort and cost of starting up an electronic version of a journal- where this includes graphics and equations - are not negligible" [Meadows et al., 1995, p. 231].

At present, a range of e-journal financing models are being tested in the market place, including free access, site licences (often to large consortia of sites), page charges levied on authors, pay per view, and conventional subscriptions or even a combination of some of these [Day, 1995]. No-one seems brave enough yet to predict what a long-term sustainable model might look like.

6.2.10Specialisation

According to [Agre, 1995a], all information commodities have to deal with economic pressures that push in opposite directions. On the one hand

> their high fixed costs of production and low marginal costs of production create powerful competitive incentives for distributing them to the largest possible audience. On the other hand, there often exists a pressure for specialization to particular communities … Content producers are developing a range of strategies to deal with these con-

tending forces.

To date, the specialisation strategy has won out for print journals. In the area of e-journals, the dynamics may be quite different. One possible strategy for dealing with both pressures is to develop a large database of material [Hunter, 1994, p. 129], [ODonnell, 1995'] that can be repurposed for delivery to a range of communities, in a range of genres and via a range of media. Another is to produce a ‘virtual’ journal drawing on a wide range of authors containing articles tailored to the stated interests of a particular subscriber. Another again is to have articles written at several increasing layers of detail, allowing a reader to ‘drill-down’ on topics they find particularly interesting.

Hal Varian [Varian, 1996] has suggested that publishers might distinguish between site-licensed and individual subscription e-journals by providing a ‘utility gap’ between the two, thus encouraging the scholar to maintain a personal subscription.

6.2.11Preserving brand identity

Most people probably do not think about scholarly journals in the same way they think about toothpaste, and yet both are examples of branded content.

> A brand is a set of expectations and associations that a given community has about a product, and attaching a brand to one's content stream is a way of explaining what it is and enabling satisfied consumers to get ‘more like that’… Brands increasingly cross media boundaries [Agre, 1995b]

A number of print journals are moving into parallel electronic delivery. In every case, they are taking their brand identity across into cyberspace. Tim O’Reilly (from O’Reilly and Associates) singles out brand identity as a “critical part of publishing success in what might be called commodity information businesses” [OReilly, 1995', p. 30]. ‘Brand identity’ for journals has always been important. As the amount of networked information available increases scholars may come to rely more and more on brand identity for journals or even individual scholars.

6.3 Leading-edge examples

In order to show how some of these responses have been packaged it is instructive to look as some example e-journals. These are not necessarily a representative sample but are all at the leading edge of the field. The initiatives they are pioneering are therefore likely to be emu-

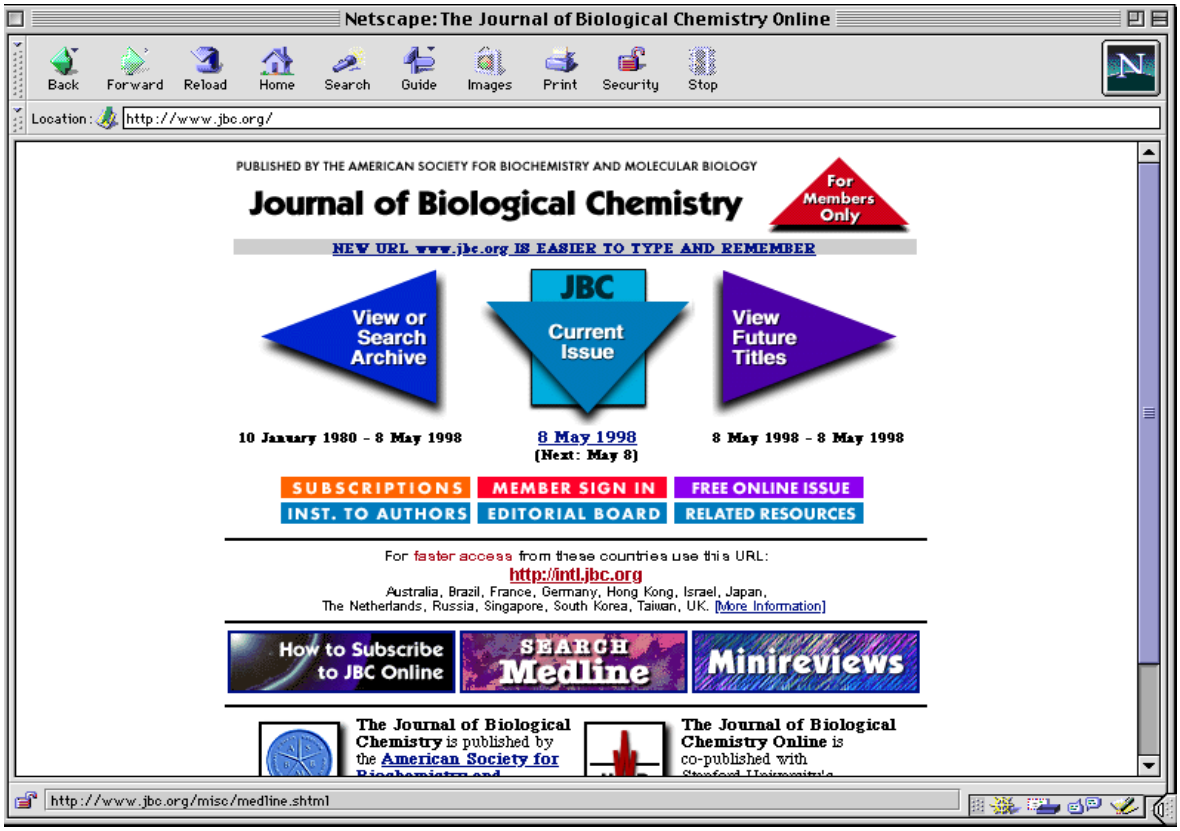
lated in the future. They also exemplify three different models for electronic journal publishing. *The Journal of Biological Chemistry* is a print journal that has moved into parallel delivery. *The Journal of Artificial Intelligence Research* is an electronic journal that is also made available in print form. *The Journal of Interactive Media in Education* is an electronic-only journal.

6.3.1 Journal of Biological Chemistry

Overview

The Journal of Biological Chemistry Online (JBC) is published by the American Society for Biochemistry and Molecular Biology in collaboration with Stanford University's Highwire Press (see 8.3: Highwire Press on page 182). It is available on the Web at <`http://www.jbc.org/`>. For a representative opening screen see figure 6–1. It was the first e-journal

Figure 6–1: Opening Screen for JBC Online. Source: JBC Online Website



published by Highwire and is in some ways still their flagship product. JBC is a parallel delivery journal and in its print incarnation is one of the larger journals. It is produced weekly and each issue can contain between 700 and 1,000 pages. In a sense it can be regarded as seven journals dealing with different aspects of biological chemistry bound into one.

Access

Access to *JBC Online* is through an individual subscription or institutional (which can cover consortia of institutions) license. ASBMB publishes the JBC on a non-profit basis and recognises the pressures on library budgets. For 1999, the institutional subscription price for the print only version of the journal (over 33,000 pages) is $1600. The price for online only (single site licence) is $1100. This presumably reflects the reduction in printing and distribution costs. There is no discount for taking both print and online - the combined price is $2700. On their Web site, they actively encourage institutions to convert multiple print subscriptions to a single online subscription. Individuals who are not members of The American Society for Biochemistry and Molecular Biology (ASBMB) must subscribe at the institutional rate. In common with many e-journals, consortium licensing attracts a discount and access is controlled through the IP number of the computer used. Without a subscription, users only have access to tables of contents, abstracts, and full text searching - enough to see what they are missing!

Navigation

Navigation through the *JBC Online* site is either hierarchical or via searching. The archive of full text goes back to the start of 1995. The archive of abstracts back to the start of 1980. Either can be viewed hierarchically starting by year and then issue. Because the cover of each issue looks essentially the same, there is no point in showing an image of the cover. Highwire also produce *Science Online,* the online version of the prestigious journal of the American Association for the Advancement of Science. Each cover of this journal is very different and the Web site for this e-journal at `<http://www.sciencemag.org/>` does provide thumbnails of the covers as a navigation technique, keying into the user's visual memories of a particular issue.

Searching is supported for volume number, author/title/abstract keywords or by words anywhere in the article, as well as limited by date range. Because each issue is so large, it is also possible to search for keywords once a user has selected an issue. It is also possible to search across multiple online journals in a single step.

Presentation

The contents of an issue are presented as author and title, with links to either abstract or full-text. The onscreen presentation of individual articles is done using HTML. Articles are or-

ganised into sections (typically Abstract, Introduction, Materials and Methods, Results and Discussion, and References) with a mini table of contents available at the start of each section to aid navigation within an article. Footnotes are implemented as internal hyperlinks. Images are shown as thumbnails which link to larger versions for detailed viewing.

For users who wish to print the article out, a PDF version is available. This is identical in appearance to the corresponding pages in the print version as it has been produced from the same original electronic version. The quality of the printed PDF will probably be better than an inter-library loan photocopy of an original.

Additional features

JBC Online provides a range of features that are only possible in the online version. These include:

- automatic creation of hyperlinks to MEDLINE citations provided by the National Library of Medicines PubMed service, for the article itself and for cited articles, and for related articles.
- links from Genbank accession numbers to full Genbank records
- bidirectional links between citing articles and cited references (where available)
- forward citations
- 'toll-free' links between the references from one journal article to the full text of the cited article, where this is also published by Highwire and subject to the subscription policies of the cited journal
- predefined Medline searches by article author
- automatic alerting service via email when new articles cite an article of interest.

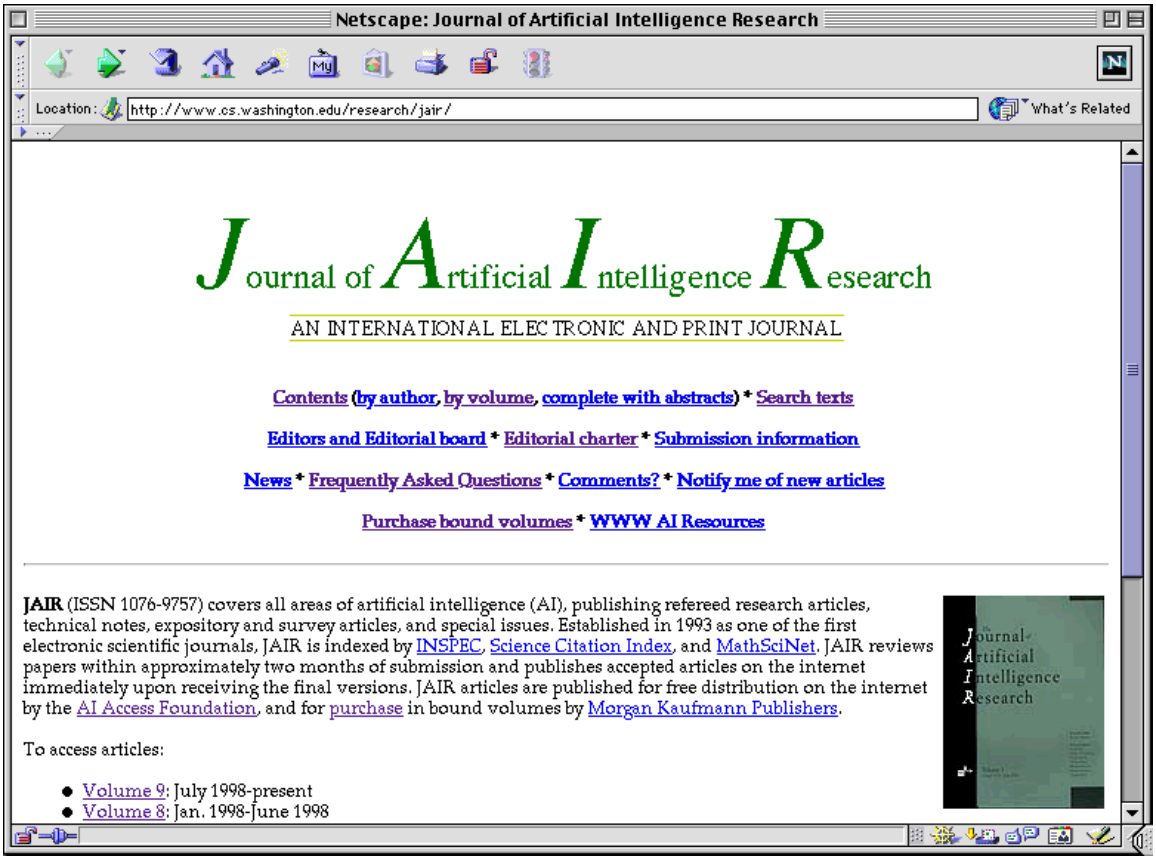
6.3.2 Journal of Artificial Intelligence Research

Overview

The Journal of Artificial Intelligence Research (JAIR) is a peer-reviewed e-journal that began accepting submissions on June 15, 1993 and published its first article in August, 1993. The publisher is the AI Access Foundation. The journal is available online at <http://www.cs.washington.edu/research/jair/>. From its inception, it has been managed as a grass roots exercise with no formal budget. The labour and resources required to run the journal are donated by individuals and scientific organizations on an as needed basis. A small seeding grant was provided by the The American Association for Artificial Intelligence to

pay for JAIR's start-up legal expenses. The opening screen for the JAIR Website is shown in figure 6–2.

Figure 6–2: Opening Screen for JAIR. Source: JAIR Website



JAIR has been described as the "Stealth E-journal of Artificial Intelligence Research" [Kling and Covi, 1995] because of the way in which it exploits the Web to provide easy and broad distribution while providing authors with publication in a medium which looks like a print journal when reproduced or cited. They also describe its deliberate and intelligent use of some of the social aspects of electronic scholarly publishing:

> By publishing polymorphously in paper and electronic media, this journal (JAIR) can offer an electronic edge to authors while appearing traditional to those who do not know its workings. Its authors and readers are part of a scholarly community where there is strong consensus on a computerized typesetting format (in this case Post-script), and in which every research lab has free (or subsidized) electronic access to Internet services. And JAIR is allied with a (commercial) publisher that routinely markets and sells books to libraries, scholars and professionals. One other key feature of JAIR's stealth approach is that it doesn't broadcast its e-journal status in its name. It is a fascinating model. [Kling and Covi, 1995]

JAIR publishes full-length research articles, short technical notes, and survey and expository articles. JAIR's editorial board is dedicated to the rapid dissemination (publication online within two months of submission) of important research results to the global AI community. The journal's scope encompasses all areas of Artificial Intelligence, including automated reasoning, cognitive modelling, knowledge representation, learning, natural language, neural networks, perception, and robotics.

Access

Access to the online version of the journal is free with no indication of any change in status in the future. JAIR deliberately uses the widest possible range of electronic distribution technologies. It is possible to access JAIR articles using news (from the `comp.ai.jair.announce` and `comp.ai.jair.papers` newsgroups), on the Web, via anonymous ftp and via automated email. The online version of JAIR is supported by USC/ISI, the University of Washington, Carnegie-Mellon University, the University of Genoa, and NASA's Ames Research Center.

The journal is also sold as a printed annual issue by Morgan Kaufmann Publishers at a cost of US$75.00 per volume (typically two volumes per year). The journal’s Website exhorts readers to encourage their library to subscribe to the printed version, as revenues from such sales help support JAIR’s operation.

Navigation

The Web site provides a table of contents by volume, another version by author, a complete list of titles with abstracts and a search engine which supports keyphrase and full-test searching. The search engine (hosted by a server in Canada) also links to an Information Space representation of JAIR hosted by a server at MIT. This is typical of the cooperative model that informs much of the work of JAIR.

The ftp server is organised into directories by volume. The automated email server supports single file retrieval and database searching.

Presentation

The default presentation format for JAIR is Postscript. Each article is a single Postscript file, formatted and paginated to match the printed journal. A printout of the Postscript is therefore of somewhat higher quality than a traditional photocopy of an original. Articles are also pro-

vided in PDF form from the Website only. Many articles are also available in HTML form. JAIR's editors do not encourage citations to include URLs – they apparently want the printed version of the articles to look as much like 'real' print as possible.

Additional features

Because of the editor's desire to have articles look like print journal articles, much of the possibilities for e-journal innovation are removed. Those articles that provide HTML versions offer the usual features: structuring of the article in chunks, Next/Up/Previous navigation buttons, internal hyperlinks for the article table of contents and footnotes, some external links for things like email addresses and personal homepages. One innovative feature that is available regardless of format is the possibility of online appendices. This possibility is sometimes used by authors to provide demonstrations or source code for algorithms and programs discussed in their articles.

6.3.3 Journal of Interactive Media in Education

Overview

The *Journal of Interactive Media in Education* (*JIME*) was launched in September 1996 and is published by the Knowledge Media Institute of the Open University in the U.K. It is available online at <http://www-jime.open.ac.uk>. A representative opening screen is shown in figure 6–3.

Its stated aims are:

- To foster a multi-disciplinary and intellectually rigorous debate on the theoretical and practical aspects of interactive media in education.
- To clarify the cognitive, social and cultural issues raised by the use of interactive media in education.
- To radically improve teaching and learning through better interactive media.
- To publish leading international research on the theories, practices and experiences in the field.
- To link scholars and commercial practitioners
- Through its innovative use of interactive Net-based media, to be an action research project which explores the changing face of journals, and more broadly, scholarly practice in the age of digital publishing and communication. [Anonymous, 1998a]

Figure 6–3: Opening Screen for JIME. Source: JIME Website

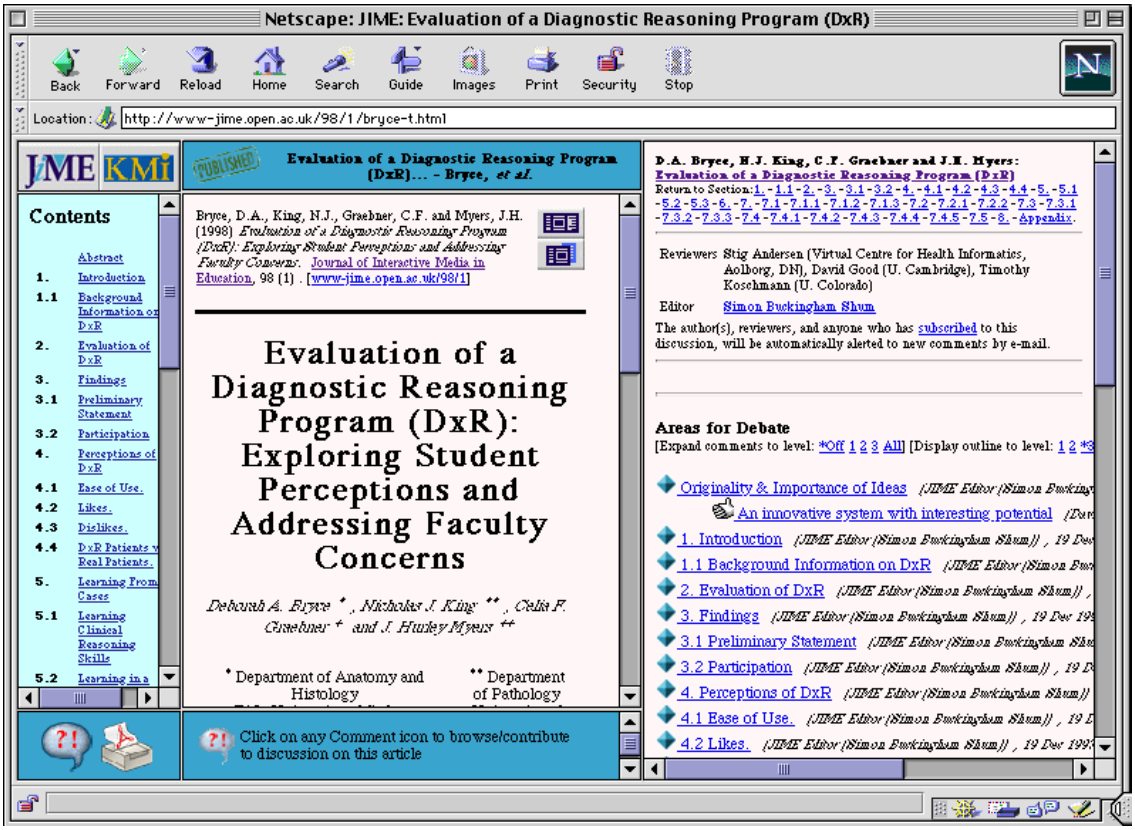
Access

Access to the journal is free (at this stage) and there is no indication on the Website of any plans to start charging. JIME is currently funded by the Open University, through the Knowledge Media Institute's (KMI) research budget, and other internal budgets as an action research project. KMI is in turn sponsored by a range of commercial organisations. JIME have had discussions with many publishers but find that the publishers are all locked into the idea that they're going to charge for access. JIME do not want this because of the effects on their open participatory model [Buckingham-Shum, 1998].

Navigation

JIME provides a heavily frames-based interface (although it can be used by a non-frames capable browser). The interface is now available in two versions: one with a floating comments window and one without. The version without is shown in figure 6–4. This version is optimised for larger screens (832 by 624 pixels or greater), and the screen capture shown is only at 800 by 600 pixels. The panel on the left shows a table of contents for the article. The panel in the middle shows the currently selected section within the article. The panel on the right

Figure 6–4: Sample JIME article. Source: JIME Website



shows the comments for the current article by section (the thumbs up indicates an approving response). Comments can be categorised as Agree, Disagree or None.

Presentation

The presentation of the articles is either in HTML or PDF. The printer icon at the foot of figure 6–4 links to a PDF version of the article displayed. Naturally, the PDF version only contains the formatted text of the article, losing the commentary, live navigation via hyperlinks and embedded simulations. Because of the hypermedia richness of JIME the paper version is significantly inferior.

Additional features

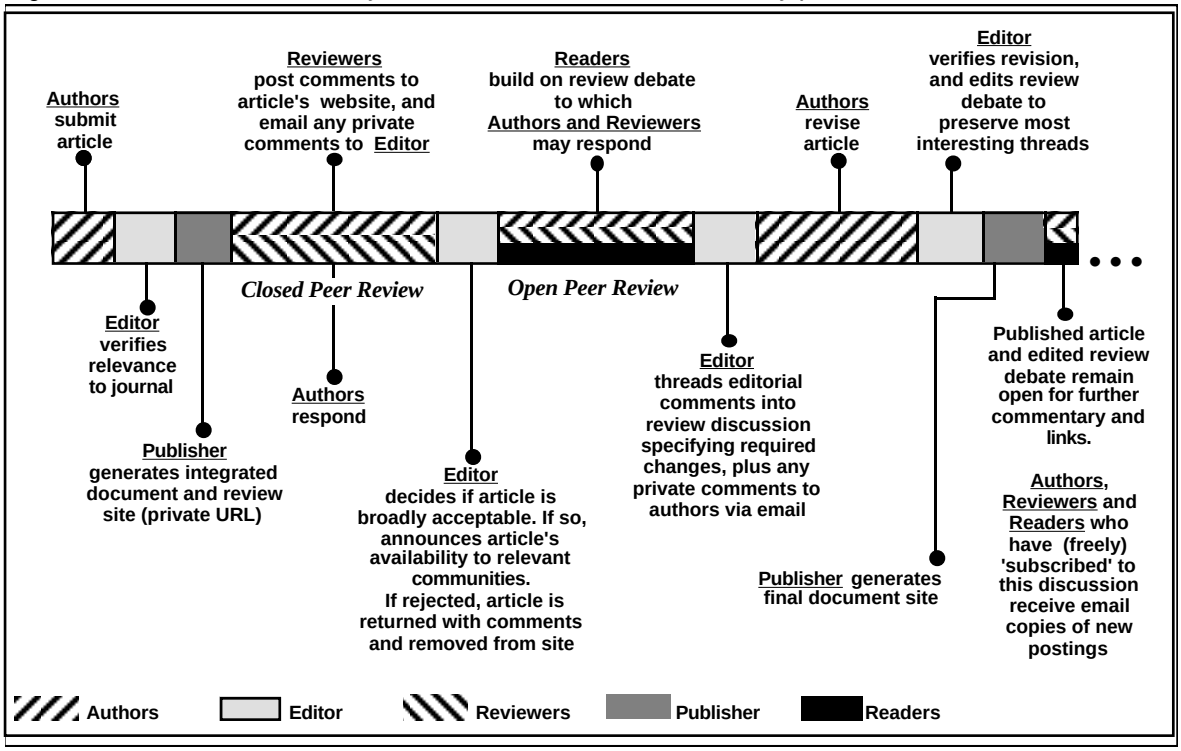
One of the things that distinguishes JIME is its integration of the review and publishing process. JIME has a publishing model that provides for both closed and open peer review as well as initial and ongoing peer commentary.

> All JIME articles are integrated with a structured web discussion space. Reviewers and authors debate a submission, after which open peer review is invited. An edited version of this discussion is preserved with the final publication, providing a forum for subsequent commentary and links to related material. Authors, reviewers, and an-

yone else who has subscribed to that discussion are alerted by email to new postings. [Anonymous, 1998b]

The review process is depicted in figure 6–5.

Figure 6–5: JIME Review LifeCycle. Source JIME Website. Used by permission.



Another of JIME's distinguishing features is its use of Shockwave and Java to support demonstrations or interactive examples embedded within articles. This allows readers to interact with the systems being described and enables the journal to model its own field of application.

6.4 Conclusion

The transforming possibilities of the new technologies have been applied to scholarly journals in a number of different ways. This chapter has reviewed some of these developing responses and shown a range of solutions to the need to support key publishing functions while doing more than is possible in print. The leading-edge example journals described are proof that it is possible to devise workable combinations of e-journal functionality and content. However, [Metz, 1995] compares the current position of e-journals with the early days of television and reminds us that predicting the success of new technologies is notoriously difficult.

IV Surveys & Case Studies

“The habit of basing convictions upon evidence, and of giving to them only that degree or certainty which the evidence warrants, would, if it became general, cure most of the ills from which the world suffers” Bertrand Russell

7 Surveys

7.1 Introduction

The previous chapter examined a range of developing responses from those wishing to publish scholarly journals to the pressures for transformation and the potentials of the new technologies. As part of this discussion a number of types of e-journals were discussed as well as three particular leading-edge e-journals. The development of these e-journals demonstrate an increase in interest in e-journals as a legitimate form of scholarly communication.

Despite this increase in e-journal activity, there has been little research into potential and actual e-journal users (see 1.8: Related work on page 11 for a review of existing survey work). This chapter reports on two related surveys dealing with attitudes and access to e-journals. The first had as its target population the readership of a single e-journal. To provide comparative data, the same survey instrument was also used with members drawn from comparable professional societies.

This chapter first discusses the processes associated with administering these two surveys. First it considers the design of the survey instrument, the choice of target population, the administration of the survey instrument, and the response rate. Next the techniques of analysis (both descriptive and statistical) are outlined, with particular care to document any assumptions made. Finally the results are presented, with the addition of tables and graphs as appropriate. The interpretation of the results and their comparison with related survey work will be dealt with in Chapter 9: Interpretation of findings.

7.2 Survey process

7.2.1 Design

The survey work reported in this thesis occurred as two separate exercises, an email survey and a print survey. The survey instrument went through several pilot stages, and was modified as a result. The instrument was specifically designed around a series of balances between completeness versus time required to fill in, and open versus closed questions. At each stage, users were invited to add additional responses as well as free-form comments at the end. The final set of questions was organised in five groups.

The first group dealt with general demographic issues: the industry category of the respondent and their primary employment role. The intention here was to see if responses would vary according to the respondent's circumstances. A range of general categories was provided and respondents could add their own category if they wished.

The second group asked about their access to computing technology either at home or at work. The specific questions asked whether they had any access to any of a personal computer, a CD-drive (for CD-ROM based publications), a sound card or equivalent, a colour screen, a direct network connection or modem access, or any other form of technology they wished to specify.

The third group asked them how often they used particular forms of electronic publishing: subscribing to an email list, accessing an ftp server, accessing a gopher server, accessing the Web, using a CD-ROM, viewing electronic journals, viewing the e-journal *Psyche* (see below), or publishing electronically.

The fourth group asked them to provide their feelings (using a standard 5-point Likert scale from Strongly Agree to Strongly Disagree) about a series of possible advantages of electronic publishing. They could also add up to two other advantages if they wished.

The fifth group asked them to provide their feelings (using the same scale) about a series of disadvantages of electronic publishing. As before, they could add up to two other disadvantages if they wished.

7.2.2 Target populations

Email survey

The first survey was undertaken using readers of the e-journal *Psyche - An Interdisciplinary Journal of Research on Consciousness*. *Psyche* is an early e-journal that commenced before the arrival of the World Wide Web. It was originally delivered using email combined with an ftp (File Transfer Protocol) archive. Since the arrival of the Web, it has introduced some additional formatting of articles, together with a range of augmented services for its reader community.

This survey was trialled and administered in conjunction with a research project funded by the Australian Vice-Chancellor's Committee (AVCC) through the Standing Committee on Information Resources, Library infrastructure projects, Program 3 - Electronic Publishing of

Full Text Materials. The author was a member of the EPICentre (Electronic Publishing Innovations Centre) project team together with Patrick Wilkens, the editor of *Psyche*, who provided access to the mailing list of subscribers for use in the survey.

The final form of the instrument was formatted to be delivered (and in most cases returned) by electronic mail. This meant that ASCII characters only were used for layout. The survey instrument as administered is attached to this thesis in the Appendices (see 11.1: Email survey instrument on page 238).

Print survey

The original study design also included the collection of comparative data from readers of print-based journals in the field of psychology. Unfortunately, there are no print journals with a directly comparable subject focus. After reviewing the results from the email survey, it was decided that the best way to survey a broadly equivalent population was to target members of psychological societies/associations in the three countries that accounted for nearly 75% of the email survey responses, and in the proportions from that survey. The rationale for this was that there might be significant differences by country. As well, the national societies were obvious survey distribution points. The societies chosen were the U.S. American Psychological Association (APA), the U.K. British Psychological Society (BPS) and the Australian Psychological Society (APS).

The survey instrument used was almost identical to that for the email survey. The differences were in the formatting, the choices for one question, and the addition of an extra choice to one question.

Because the print survey was to be administered in print form, the facilities of a word-processor were used to improve the formatting without major changes to the layout. Providing the original ASCII-formatted survey instrument would have looked both amateurish and unattractive.

The choices for question four were changed slightly and the letter codes altered to reduce errors. In the email survey, about 5% of respondents had found it confusing that they were asked to select the letters A through E (with no mnemonic value) in question 4. These respondents typically wrote in R for Regularly (instead of B) or N for never (instead of E). Because it is impossible to determine if a respondent meant R to mean Regularly or Rarely, such answers had to be coded as incorrect. To try to avoid this problem in the print survey, the

letters required in question four were changed to the first letter of each option, and Rarely changed to Seldom so as not to clash with Regularly. This change should not have affected the response pattern except by reducing incorrect answers.

An additional choice was also added to question 4. It was important to determine if the respondents to the print survey had ever viewed *Psyche*, and if so how often. This enabled later segmentation of the print survey respondents into two sub-populations if required.

The survey instrument as administered incorporating these changes is attached to this thesis in the Appendices (see see 11.2: Print survey instrument on page 240).

It should be noted that because of the time taken to process the email survey instrument and to arrange for the distribution of the print survey instrument, a year elapsed (roughly) between the administration of the email and print survey instruments. This needs to be borne in mind when considering the issue of the rate of technological change and its possible effect on survey responses.

7.2.3 Survey administration

Email survey

As discussed in [Berge&Collins1996], email surveys make it "easy to census the entire subscription list of a discussion or distribution list by sending a single message to the mailing list" rather than relying on a sampling technique. This was the technique used for this research. Psyche has associated with it two electronic mailing lists, Psyche-L and Psyche-D. Psyche-L is used for distribution of the ASCII version of *Psyche*, while Psyche-D is used for discussion based around the themes of the journal. The instrument was originally distributed to Psyche-L on September 14, 1995. In order to increase the response rate, it was redistributed to both Psyche-L and Psyche-D on February 16, 1996. The instrument was also posted on a Web site so that people could download the text, complete it and send it in. Use of a Web forms mechanism to gather responses was not adopted due to lack of time to set up the form and data-handling. Other than the redistribution to the mailing lists (which is by definition non-targeted), there was no specific follow-up of respondents.

Print survey

Those surveyed were provided with:

- a generic covering letter (it was not possible to gain access to the membership data-

bases to create a personalised letter) including contact information and the information required by Monash University's Standing Committee on Ethical Research on Humans (SCERH)

- the survey instrument, printed double-sided on a single A4 sheet
- a reply-paid envelope for return of the survey
- a bookmark (generously provided by RAECO Library Services) as a token appreciation for the respondent's time.

The mailout of the print survey was governed by the procedures of the different societies/ associations.

The American Psychological Association (APA) does not make its mailing lists available to any organization or individual.They provided mailing labels for one time use after approving the materials submitted. A random sample of 2000 labels was ordered from the total membership of approximately 85,000. There was no subgroup that corresponded adequately to the readership of *Psyche*. Assuming a 10% response rate, 2000 surveys would provide around 200 responses (compared with the 190 responses from the U.S. A. in the email survey). The surveys were sent out from Australia in February 1997.

The U.K. British Psychological Society (BPS) does not divulge members' addresses to any third party.They require the materials to be supplied to their head office where they handle the insertion and postage for a fee.The BPS allows members to nominate one of a number of interest areas as part of their membership profile. The closest match to the *Psyche* population was deemed to be Cognitive Psychology. A random sample of 360 labels was ordered from this section. Assuming a 10% response rate, this would provide around 36 responses (compared with the 36 responses from the U.K. in the email survey). The materials were sent from Australia to the U.K. in December 1996. It was agreed with the BPS office that they were not to be posted out to members until mid-January 1997, thus avoiding the Christmas rush.

The Australian Psychological Society (APS) uses a similar system. A random sample of 220 labels from the entire membership was ordered. Assuming a 10% response rate, this would provide around 22 responses (compared with the 22 responses from Australia in the email survey). The materials were sent to the APS head office for distribution in February 1997.

Note that all the above procedures made it impossible to send out follow-up letters. In other words, the mechanisms to improve response rate were little better than for the email survey

(none). The only difference was the ability to include the bookmark with the survey as a minor inducement to feel some sense of reciprocal obligation.

7.2.4 Survey response

Email survey

The last response from the email to Psyche-L and Psyche-D came in on March 3, 1996. 336 responses were received in all, from a mailing list of 2800 participants. This is a response rate of 12% which is not unusual for Internet-based surveys where there is no overt inducement for completion or personalised follow-up. There is as yet little documented experience of response ratios to surveys administered over the Internet. Anderson and Gansneder reported that in their review of the literature that researchers reported response rates between 41% and 76% when administering surveys via email [Anderson&Gansneder1995]. Many of the studies they cite are now some years old and date from the time when email was something of a novelty. There is some anecdotal evidence that professionals of all categories are suffering from 'survey fatigue' with longitudinal researchers in the Information Systems field reporting steadily dropping response rates.

Of the 336 responses to this survey, 314 came directly from subscribers to either of the mailing lists. The remaining 22 were from people who had the survey instrument forwarded to them or who had found the on-line version while searching the Web for Psyche-related topics. For this reason, all the responses have been grouped together and analysed as a block. Of the 336 responses, 332 were received by email. Of the remaining four, three were printed out and mailed in. One was faxed in.Respondents seem to have found it easy to complete and return the survey instrument. Nearly one third of the responses came within the first two days.

Because the overwhelming majority of Psyche responses were in the form of email it was possible to do some analysis of the geographical distribution of the respondents. Of the four print responses, it was possible to assign a domain to three of them.Table 7–1shows an initial division into US domains, and then all other countries. The.us domain was only associated with k12.xx.us email addresses in this survey. Outside the US, only countries with more than 5 respondents have been listed separately. Australia is noteworthy for having the second-largest non-US contingent, despite Canada's larger population. The 52 aggregated responses included respondents from Argentina, Austria, Belgium, Brazil, Chile, Croatia, Denmark, Estonia, Fiji, Finland, Hong Kong, Hungary, Ireland, Israel, Italy, Japan, New Zealand, Nor-

way, Portugal, Singapore, the Slovak Republic, South Africa, Spain, Sweden, Switzerland, and Taiwan.

Table 7–1: Email survey Responses by Country (N=336)

Category	No.	Percent
US Domains:		
.com	46	13.7
.edu	125	37.2
.gov	4	1.2
.net	4	1.2
.org	8	2.4
.us	3	0.9
Countries with more than 5 replies:		
.uk	36	10.7
.au	22	6.5
.ca	17	5.1
.nl	8	2.4
.de	5	1.5
.fr	5	1.5
Other Categories:		
Other countries (aggregated)	52	15.5
Unknown	2	0.6
Total:	**336**	**100.0**

Print survey

The overall response rate for the print survey was 702 responses from 2580 survey instruments which is 27.2%. This is quite acceptable for a one-shot survey with no follow-up and no inducement to respond.

The response rate differed quite markedly by society. For the American Psychological Association (APA), the response rate was 24.3%, with a total of 486 APA responses received. For the U.K. British Psychological Society (BPS), the response rate was 35.8%, with a total of 129 BPS responses received. For the Australian Psychological Society (APS), the response rate was 39.5%, with a total of 87 APS responses received. This may reflect a greater exposure to surveys (of all sorts) in the U.S. and hence a greater reluctance to take part.

The last print survey was received in May 1997.

Table 7–2 shows the number and percentage of surveys received in each print subgroup as well as the email survey.

Table 7–2: Frequency Distribution for Survey Groups (N=1038)

Survey Group	Count	Percent
Email	336	32.370
Print (APA)	486	46.821
Print (APS)	87	8.382
Print (BPS)	129	12.428
Total	**1038**	**100.000**

7.2.5 Error minimisation

The four main error types to be avoided in survey work are sample selection bias, nonresponse error, item nonresponse error and response error.

Sample selection bias

This can occur either when a poor or nonrepresentative sample is selected from a target population or when the list chosen for the survey mailout is the wrong list. In the case of the email survey of the *Psyche* readership, the lists used were those read by the subscribers to *Psyche* (and used for notification of new issues) and were thus clearly the correct lists. No sampling was undertaken - the entire list was used. In the case of the print survey, the random selection of members of the societies was undertaken by the societies themselves (according to their procedures for mailouts). It was not possible to influence this selection process but it can be assumed to be truly random. Where possible (with the BPS) a sub-sample of the membership with an interest in the discipline area of Psyche was requested. This should ensure the best possible match between email and print survey respondents. As far as possible, therefore, causes of sample selection bias were eliminated in the survey design.

Nonresponse error

This relates to the bias inherent in the sub-population that responds to the survey. Selecting the right sample or list is of little use if a biased sub-sample is the only one that responds. Determining if the responses are from such a biased sub-sample is very difficult. The standard advice is to aim for a high response rate (75% or higher according to [Mangione1998]). Such a response ratio could not realistically be expected in the circumstances of the survey and neither the email or print surveys achieved anything near this. In the case of the email

survey, a number of reminders were sent to try to improve compliance. These were not targeted reminders because of difficulties in getting access to the list addresses. In the case of the print survey, the societies were either not prepared to supply the addresses or would only supply addresses for a one-shot mailout. This meant that reminder letters were not an available option. One attempt to improve responses rates in the print survey was the inclusion of a small reward (the bookmark) in an attempt to induce some sense of obligation. Because of the low response rates it is possible that the responses to the surveys are biased, but eliminating this source of error was not achievable given the constraints inherent in the mailing lists used.

Item non-response error

This occurs when respondents fail to answer individual questions, answer them incorrectly or add comments that do not fit the existing categories. Where this occurred, it was coded as Blank or Invalid (or sometimes Blank/Invalid) and is so shown in the tables and graphs. The percentage of invalid responses typically falls between 1 and 5. The percentage of Blank responses typically falls between 5 and 10, and can be interpreted as respondents not answering if they were unsure. Some causes of instructional ambiguity were picked up in the pilot stage and some others identified once the email survey was processed. These causes were remedied once detected.

Item response error

This occurs when respondents misunderstand the wording of the questions as presented. The pilot phase of the survey instrument and a number of valuable comments from colleagues ensured a final survey that was as clear as possible while remaining short and focused.

7.2.6 Data encoding

The results were entered into the Statview package to produce both descriptive and statistical summaries of the information. The final data set comprised 1038 rows by 44 columns for a total of 45672 data points.

The answers to the survey instrument were encoded as a series of variables. These variables appear throughout this chapter as identifiers in tables and graphs. The variables are grouped into categories according to the sections of the survey instrument. Table 7–3 shows the variables used.To assist in grouping variables, each variable has been prefixed with its survey instrument category. Variables starting with T- relate to the Technology questions. F- indi-

Table 7–3: Encoded Variables

Variable/Category	Survey Question
Demographics	
Industry Category	Industry Category
Employee Role	Primary Employment Role
Technology	
T-PC	Access to a Personal Computer
T-CD Drive	Access to a CD-ROM drive
T-Sound	Access to Sound output
T-Colour	Access to a Colour screen
T-Network	Access to a Direct Network connection
T-Modem	Access to a Modem connection
T-Other	Any other form of computing technology (specified by the respondent)
Frequency of use of electronic publishing mechanisms	
F-Subscribe	Subscribe to electronic publishing forums
F-FTP	Use FTP (File Transfer Protocol) to access materials
F-Gopher	Use Gopher to access materials
F-WWW	Use World Wide Web to access materials
F-CDROM	Access materials on CD-ROM
F-Views	View electronic journal(s)
F-Psyche	View the electronic journal Psyche
F-Publishes	Publish electronically
Advantages of electronic scholarly publishing	
A-Speed	Speed of publication
A-24 Hour	24 hours a day access
A-Convenience	Convenience
A-Feedback	The way they encourage feedback
A-Paper	Reduced paper consumption
A-Searching	Ease of searching
A-Multimedia	Multimedia publications
A-Affordability	Affordability
A-US1	User-entered advantage number 1
A-US2	User-entered advantage number 2
Disadvantages of electronic scholarly publishing	
D-Quality	Poor quality
D-Refereeing	Lack of refereeing

Table 7–3: Encoded Variables

Variable/Category	Survey Question
D-Copyright	Concerns about copyright
D-Plagiarism	Increased plagiarism
D-Skills	Special skills needed to use
D-Equipment	Special equipment needed to access
D-Format	Format is not reader friendly
D-Costs	Communication costs to access them
D-US1	User-entered disadvantage number 1
D-US2	User-entered disadvantage number 2

cates the Frequency of use questions. A- and D- stand for Advantages and Disadvantages respectively.

7.2.7 Descriptive analysis

The main descriptive summary produced was frequency distributions for the answers to each question. The answers to each question consisted of Nominal data (choices from a number of names; industry categories, employee roles, Likert values). This meant that it was not possible to calculate descriptive statistical values that are only appropriate for Continuous variables like means, standard deviations and the like.

The presentation of the descriptive survey results treats the email survey and print surveys together. This enables cross-comparisons to be made more easily between these groups. Within the print survey each of the subgroups by professional society has been broken out. This is because statistical work has indicated patterns of relationships within the print subgroups.

7.2.8 Statistical analysis

Because the survey variables are all Nominal, the range of possible statistical tests is restricted. The most appropriate is Contingency table analysis.

> Contingency table analyses determine whether a relationship exists between two nominal variables. Other statistics (t-tests, regressions, means, correlation tests) apply to dependent variables that are continuous, that is, they are capable of taking on many different values with an obvious ordering to them like height, weight, income chemical concentration, sales, etc. Tests applied to continuous variables lose their va-

lidity with nominal variables that do not have an ordered, continuous property. [Abacus1996, p. 81].

Contingency tables (called cross-tabulations in SPSS) structure the data into a two-way table showing the groupings for each of two different variables. For instance, a contingency table for the variables of Society and Industry Category would show all the possible answers for Society on one table axis and all the possible answers for Industry Category on the other table axis. The cells of the table each show the number of observations for one combination of answers. For this survey this contingency table looks like table 7–4.

Table 7–4: Observed Frequencies for Society, Industry Category (N=1038)

	Blank	Combined	Consultant	Education	Government	Industry	Other	Totals
Email	1	4	22	239	10	36	24	**336**
APA	3	12	168	171	43	18	71	**486**
APS	0	1	45	19	10	6	6	**87**
BPS	0	1	10	95	11	3	9	**129**
Totals	**4**	**18**	**245**	**524**	**74**	**63**	**110**	**1038**

Once a contingency table has been constructed it is possible to examine the values to see which combinations of answers show more or less observations than would be expected if the two variables are independent. The statistical test to use in this case is the chi-square test for independence.

> The hypothesis of independence states that the likelihood of an observation falling into one group for one variable is independent of the other group the observation falls into. To calculate this test, Statview finds the expected value for the number of observations for every combination of groups based on the hypothesis of independence and compares the expected with the observed values in each cell. [Abacus1996, p. 82].

The null hypothesis is that the two variables are independent. That is, a respondent who is from the BPS is no more likely to work in an educational institution than in other industry. Conversely, a respondent who works as a consultant is no more likely to be a member of the APA than any other society. If one calculates a low chi-square value (and a corresponding high probability indicated by the letter *p*) for a particular combination of two variables, one would tend to accept the null hypothesis.

If the null hypothesis is rejected (on the basis of a large chi-square value and corresponding low *p* value) then one would have identified a relationship between the two variables. One can then examine the contingency tables in more detail to identify particular combinations of variables where the expected number of observations is significantly different from that observed.

Statview provides both an expected values table and a table of post-hoc cell contributions to the overall chi-square statistic to assist with this. The *expected* values table shows the expected number of observations for every combination of groups based on the hypothesis of independence. Note that the chi-square test is not valid when the minimum *expected* value in any cell is less than five. Observed values in a cell can be lower than five without causing any problems.

The post-hoc cell contributions are

> a form of standardized residual that indicate what each cell in the table contributes to the chi-square statistic. Since they are calculated to follow a standard normal distribution, absolute values greater than, for example, 1.96 for a 0.05 probability level indicate that the cell in question provides significant information about the combinations of groups of the variables whose occurrence is different than what would be expected under the hypothesis of independence. [Abacus1996, pp. 82-3].

The main application of the chi-square test to contingency tables was to determine whether it was possible to treat all the print survey subgroups as an aggregate. In order to test this, the dataset was first restricted to print replies only. Contingency tables were then calculated for each variable. The null hypothesis for each successive variable was that the Society variable (used to distinguish print subgroups from each other, as well as from the email survey) and the variable under consideration were independent. Chi-square values and probabilities were then calculated. (The Society variable as restricted to responses to the print survey only will be referred to as Society-Print from now on as a convenient shorthand). There are three possible outcomes from this contingency table analysis for the Society-Print responses.

If none of the expected cells in the contingency table for a given variable are less than 5 and the chi-square p-value is > 0.05 then the print subgroups show independence. The subgroups were therefore aggregated and treated as a block (relative to the email responses) for these questions. Variables in this category are T-CD Drive, T-Sound, and T-Colour.

If none of the expected cells in the contingency table for a given variable are less than 5 and the chi-square p-value is < 0.05 then the print subgroups show dependence with the Society variable. The variables in this group are T-Network, T-Modem, F-Subscribe, F-Web, F-CDROM, and F-Views. In this case, the response patterns for each subgroup are significantly different and they could not be treated as an aggregate for these questions.

For the remainder of the survey questions, at least one expected cell in the contingency table had a value of less than five. This means that the chi-square p value calculated for the disaggregated print survey subgroups was not valid and could not be used as the basis for analysis. For these variables, the print subgroups was also aggregated and compared as a block with the email survey responses. Variables in this category are Industry Category, Employee Role, T-PC, F-Ftp, F-Gopher, F-Psyche, F-Publishes, all the A(Advantages) variables and all the D (Disadvantages) variables.

For those cases where the contingency table analysis indicated it was reasonable or necessary to group the print subgroups together, there was a possibility that there might be statistically significant differences between the email and print surveys. To further explore this, contingency tables were constructed between the Survey variable and each of the technology variables. This provided a much better picture of the distribution of results between the various groups.

Thus, when discussing the results of the contingency tables, the variable being compared to the survey question under discussion will be one of two possibilities. If the analysis of the print subgroups has supported the hypothesis of independence, then the print subgroups can be aggregated and compared as a whole to the email survey. In this case the variable referred to will be Survey. If the contingency table analysis has not supported the hypothesis of independence between the three different psychological societies within the print survey, then the print subgroups cannot be aggregated and treated together. In this case the variable referred to will be Society, which distinguishes the three psychological societies and the email survey as four separate groups. In other words, the email survey is being treated as another ‘society’ for convenience of analysis.

7.3 Survey Results

7.3.1 Basic demographics

Industry Sector results

This question asked respondents to indicate the industry sector for their employing institution. They were asked to choose one category and mark it with a cross. For students, it was assumed that their 'employing institution' was the one where they were studying. No-one indicated confusion with this interpretation. Table 7–5[1] shows the responses ordered alphabetically for each of the separate surveys and sub-surveys.

Table 7–5: Industry Sector breakdown by Society (N=1038)

Survey	Blank	Combined	Consultant	Education	Government	Industry	Other	Total
Email Count	1	4	22	239	10	36	24	**336**
Email Percent	.30	1.19	6.55	71.13	2.98	10.71	7.14	**100.00**
APA Count	3	12	168	171	43	18	71	**486**
APA Percent	.62	2.47	34.57	35.19	8.85	3.70	14.61	**100.00**
APS Count	0	1	45	19	10	6	6	**87**
APS Percent	0.00	1.15	51.72	21.84	11.49	6.90	6.90	**100.00**
BPS Count	0	1	10	95	11	3	9	**129**
BPS Percent	0.00	.78	7.75	73.64	8.53	2.33	6.98	**100.00**
Overall Count	4	18	245	524	74	63	110	**1038**
Overall Percent	.39	1.73	23.60	50.48	7.13	6.07	10.60	**100.00**

The *Blanks* columns is for those who did not answer the question. The *Combined* column shows those respondents who chose more than one option under Industry Sector (the survey instrument asked for primary industry sector only). These were a very small percentage.The *Others* column allowed respondents to write in a sector different to the choices given. The APA respondents were on average twice as likely to take this option.

1. In order to make it easier to read this table (and successive tables like it), a different shade has been used for the count and percent rows. In other words, the percent rows are always the same shade. As this table shows percentages, it is important to indicate clearly if these are row or column percentages. Three things indicate this: the totals are in bold, they are at the end of the row (i.e. row percentages), and the percent totals always add up to 100.

Because some of the cells in the expected frequencies table (not shown) contain less than 5 expected responses no statistics can be calculated for the independence of survey group/subgroup and industry category. It can however be observed that the majority of the APA subgroup (some 70%) work in the Consultant or Education sectors. More than half of the APS subgroup work as Consultants. More than two thirds of the BPS respondents and the email respondents identified their 'industry' as being Education. The survey instrument did not ask for a breakdown by education sector, but it seems reasonable to assume the majority are from the tertiary sector. The email addresses for the email group certainly reflect this interpretation.

Employment Role results

This question asked respondents to indicate their primary employment role. They were asked to choose one category and mark it with a cross. Table 7–6 shows the responses ordered alphabetically.

Once again, the column labelled Combined is for those respondents who marked more than one category. As a noticeable number of those who marked more than one category in the email survey indicated that they felt they were unable to separate their teaching from their research, a separate category of Teaching/Research was created. Respondents who selected both the Teaching and Research categories in the email survey were recoded into this category. This category was then added to the print survey to avoid the need for recoding.

Because some of the cells in the expected frequencies table (not shown) contain less than 5 expected responses no statistics can be calculated for the independence of survey group/subgroup and employee role. However, the breakdowns of answers within the subgroups are very different and need to be considered separately.

Table 7–7 shows the contingency table for Industry Sector against Employee Role for the APA subgroup. Within this subgroup, the predominant category is clinical practitioner (55% of all responses). Of these clinical practitioners, over half work as consultants. The education sector is responsible for over half the research employees, and almost all the teaching or teaching/research employees (not surprising perhaps). As one would expect, almost all the consultants list their industry sector as consultant as well.

Table 7–8 shows the same breakdown for the APS. Here the overall numbers are much smaller, but consultant is still the largest employment role.

Table 7–6: Employment Role breakdown by subgroup and total

	Administration	Blank	Combined	Consultant	Information Services	Other	Practitioner	Research	Retired	Student	Teaching	Teaching/Research	Totals
Email Count	10	2	13	16	17	14	24	75	4	69	77	15	**336**
Email Percent	2.98	.60	3.87	4.76	5.06	4.17	7.14	22.32	1.19	20.54	22.92	4.46	**100.00**
APA Count	33	0	12	20	0	12	267	25	17	2	41	57	**486**
APA Percent	6.79	0.00	2.47	4.12	0.00	2.47	54.94	5.14	3.50	.41	8.44	11.73	**100.00**
APS Count	7	3	3	36	0	6	14	6	1	0	3	8	**87**
APS Percent	8.05	3.45	3.45	41.38	0.00	6.90	16.09	6.90	1.15	0.00	3.45	9.20	**100.00**
BPS Count	3	0	6	10	0	5	9	16	5	16	18	41	**129**
BPS Percent	2.33	0.00	4.65	7.75	0.00	3.88	6.98	12.40	3.88	12.40	13.95	31.78	**100.00**
Overall Count	53	5	34	82	17	37	314	122	27	87	139	121	**1038**
Overall Percent	5.11	.48	3.28	7.90	1.64	3.56	30.25	11.75	2.60	8.38	13.39	11.66	**100.00**

Table 7–9 shows the contingency table for Industry Sector against Employee Role for the BPS subgroup. The APS respondents are much more evenly distributed by both employment role and industry sector.

7.3.2 Access to technology

Respondents were asked to indicate which pieces of technology they had access to, either at home or at work. The intention was to determine what facilities an e-journal could assume

Table 7–7: Industry Sector (X-axis) against Employment Role (Y-axis), APA (N=486)

	Blank	Combined	Consultant	Education	Government	Industry	Other	Totals
Administration	1	0	0	14	7	4	7	**33**
Blank	0	0	0	0	0	0	0	**0**
Combined	1	3	2	2	0	0	4	**12**
Consultant	0	0	19	0	0	1	0	**20**
Information Services	0	0	0	0	0	0	0	**0**
Other	0	0	0	3	1	4	4	**12**
Practitioner	1	9	143	34	27	8	45	**267**
Research	0	0	1	16	3	1	4	**25**
Retired	0	0	2	8	4	0	3	**17**
Student	0	0	0	2	0	0	0	**2**
Teaching	0	0	0	40	0	0	1	**41**
Teaching/Research	0	0	1	52	1	0	3	**57**
Totals	**3**	**12**	**168**	**171**	**43**	**18**	**71**	**486**

Table 7–8: Industry Sector (X-axis) against Employment Role (Y-axis), APS (N=87)

	Blank	Combined	Consultant	Education	Government	Industry	Other	Totals
Administration	0	0	0	5	1	1	0	**7**
Blank	0	0	3	0	0	0	0	**3**
Combined	0	0	3	0	0	0	0	**3**
Consultant	0	1	30	0	0	3	2	**36**
Information Services	0	0	0	0	0	0	0	**0**
Other	0	0	0	0	3	1	2	**6**
Practitioner	0	0	8	1	3	0	2	**14**
Research	0	0	1	2	2	1	0	**6**
Retired	0	0	0	0	1	0	0	**1**
Student	0	0	0	0	0	0	0	**0**
Teaching	0	0	0	3	0	0	0	**3**
Teaching/Research	0	0	0	8	0	0	0	**8**
Totals	**0**	**1**	**45**	**19**	**10**	**6**	**6**	**87**

from its readership. There is little point in adding sound or full-colour video if a majority of readers cannot benefit from this.

Table 7–9: Industry Sector (X-axis) against Employment Role (Y-axis), BPS (N=129)

	Blank	Combined	Consultant	Education	Government	Industry	Other	Totals
Administration	0	0	0	2	1	0	0	3
Blank	0	0	0	0	0	0	0	0
Combined	0	0	0	4	0	1	1	6
Consultant	0	0	10	0	0	0	0	10
Information Services	0	0	0	0	0	0	0	0
Other	0	0	0	1	0	2	2	5
Practitioner	0	0	0	1	5	0	3	9
Research	0	1	0	14	1	0	0	16
Retired	0	0	0	4	1	0	0	5
Student	0	0	0	11	2	0	3	16
Teaching	0	0	0	18	0	0	0	18
Teaching/Research	0	0	0	40	1	0	0	41
Totals	0	1	10	95	11	3	9	129

Table 7–10 shows a summary of the responses to this set of questions, ordered by the sequence of questions in the survey instrument. It is clear from table 7–10 that over two-thirds

Table 7–10: Access to Technology (N=1038)

Technology	Access count	Access%
Personal computer	999	96.24
CD-ROM drive	759	73.12
Sound output	714	68.79
Colour screen	904	87.09
Direct network connection	617	59.44
Modem connection	626	60.31
Other	77	7.41

of the respondents have access to various components of technology to support multimedia (sound + CD-ROM + colour output) on their computers. A more detailed analysis of the raw data reveals that of the 999 who have access to personal computers, 756 also have access to CD-ROM drives, 710 have access to sound output and 899 have access to colour screens. The set of those respondents who have access to a personal computer with sound output, CD-ROM, and colour output is 641 of the 1038 total respondents or 61.75%

What therefore is the detailed technology picture arising from these surveys?

Access to a personal computer

There was some confusion about what constituted a personal computer (PC). Some respondents indicated that they considered a Unix workstation as a PC, while others listed this in the Other section. Because of the difficulty in reinterpreting the data, the responses have been left as received.

Of the total pool of respondents, 96.24% have access to a personal computer (however they define that). There are no statistically significant correlations between this variable and either the Society-Print variable (which distinguishes between the different print survey subgroups) ($p = 0.4090$) or the Survey variable (which distinguishes between the email and print surveys) ($p = 0.2061$). In other words, the entire survey population can be treated as single group with respect to this variable.

Access to a CD-ROM drive

Nearly three-quarters of the respondents (73.12%) have access to a CD-ROM drive. As with the previous variable, there are no statistically significant correlations with either Society-Print ($p=0.2774$) or Survey ($p=0.6875$).

Access to sound output

Slightly more than two-thirds of the respondents (68.79%) have access to sound output. There are no statistically significant correlations with either Society-Print ($p=0.8061$) or Survey ($p=0.3247$).

Access to colour

Nearly as many respondents (87.09%) have access to colour screens as have PC's. There are no statistically significant correlations with Society-Print ($p=0.3455$). That is, the differences in responses between print subgroups probably arise by chance. There is, however, an correlation between Colour and Survey ($p=0.0045$). In other words, the chance is less than 5% that the differences in access to colour output between the email and print groups is due to chance. The email group shows slightly greater but statistically significant access to colour output than would be expected by chance (and the print group shows correspondingly less

access to colour output). Table 7–11 shows the frequency distribution and expected versus observed values.

Table 7–11: Frequency Distribution for T-Colour, split by Survey (N=1038)

	No	Yes	Total
Email Expected	43.38	292.62	**336.00**
Email Count	29	307	**336**
Email Percent	8.63	91.37	**100.00**
Print Expected	90.62	611.38	**702.00**
Print Count	105	597	**702**
Print Percent	14.96	85.04	**100.00**
Overall Expected	134.00	904.00	**1038.00**
Overall Count	134	904	**1038**
Overall Percent	12.91	87.09	**100.00**

Access to a direct network connection

Nearly three-fifths of the respondents (59.44%) had access to a direct network connection to the Internet. Table 7–12 shows the frequency distribution for the answers to this question, broken down by Society. If one performs a contingency analysis then there is a statistically

Table 7–12: Frequency Distribution for T-Network, split by Society (N=1038)

Category	No	Yes	Total
Email Expected	136.28	199.72	**336**
Email Count	78	258	**336**
Email Percent	23.21	76.79	**100.00**
APA Expected	197.12	288.88	**486**
APA Count	246	240	**486**
APA Percent	50.62	49.38	**100.00**
APS Expected	32.29	51.71	**87**
APS Count	57	30	**87**
APS Percent	65.52	34.48	**100.00**
BPS Expected	52.32	76.68	**129**
BPS Count	40	89	**129**
BPS Percent	31.01	68.99	**100.00**
Overall Expected	421	617	**1038**
Overall Count	421	617	**1038**
Overall Percent	40.56	59.44	**100.00**

significant difference ($p<0.0001$) between the responses from the different subgroups. One

way to examine these differences is via the observed and expected frequencies. Table 7–12 also shows these values for this variable.

Table 7–13: Post Hoc Cell Contributions for Direct Network connection, split by Society

Society	No	Yes
Email	-7.874	7.874
APA	6.193	-6.193
APS	4.953	-4.953
BPS	-2.361	2.361

Another way to see the relative contributions of each cell is the post-hoc cell contributions table. Table 7–13 shows the contributions of the different subgroups to the contingency table chi-square p-value. A value of 1.96 is significant at the 0.05 significance level, and thus all of the societies show significant differences between expected and observed values (assuming the hypothesis of independence). The APA respondents show much less access than expected with 246 negative observed answers versus 197.12 expected. The APS respondents demonstrate the same picture: 57 negative observed answers versus 32.29 expected. The BPS respondents on the other hand show significantly more access to a direct network connection than expected with 89 positive observed answers, as against 76.68 expected. The largest post-hoc contribution comes from the email survey respondents who have a contribution of 7.874 and an observed positive count of 258 versus an expected value of 199.72. The email survey respondents had the greatest access to direct network connections, followed in order by the BPS, APA and then APS respondents.

Access to a modem connection

Just over three-fifths of the respondents (60.308%) had access to a modem connection to the Internet. As with the network question, there is a statistically significant difference ($p<0.0001$) between the responses from the different subgroups. The post-hoc cell contributions of each of the different subgroups to the contingency table chi-square p-value are significant at least at the 0.05 probability level. Table 7–14 shows observed and expected frequencies broken down by Society. The APA respondents show more access to a modem connection than expected. The APS respondents show less access (as with the network question). The BPS respondents also show significantly less access than expected. The email survey respondents again have the highest proportion of positive answers (although the APA respondents are very close).

Table 7–14: Frequency Distribution for T-Modem, split by Society (N=1038)

Category	No	Yes	Total
Email Expected	133.36	202.64	**336**
Email Count	110	226	**336**
Email Percent	32.74	67.26	**100.00**
APA Expected	192.90	293.10	**486**
APA Count	170	316	**486**
APA Percent	34.98	65.02	**100.00**
APS Expected	34.53	52.47	**87**
APS Count	47	40	**87**
APS Percent	54.02	45.98	**100.00**
BPS Expected	51.20	77.80	**129**
BPS Count	85	44	**129**
BPS Percent	65.89	34.11	**100.00**
Overall Expected	412	626	**1038**
Overall Count	412	626	**1038**
Overall Percent	40.56	59.44	**100.00**

7.3.3 Use of electronic publishing technologies

The next set of questions was intended to determine the frequency with which the respondents used various forms of electronic publishing technologies and therefore their familiarity with those technologies. The possible activities that respondents might take to work with various forms of electronic publishing are:

- subscribing to various electronic publishing forums (such as mailing lists, network news groups, bulletin board systems and the like)
- using FTP to access materials stored on remote ftp servers
- using Gopher to access materials stored on remote gopher servers
- using the World Wide Web to access materials
- accessing materials stored on CDROMs
- viewing electronic journals (however retrieved)
- publishing electronically themselves.

Electronic publishing fora

Many early e-journals were distributed via email lists; *Psyche* still is. This question was specifically targeted at respondents' use of one-many computer-mediated communications fora rather than one-one email.

Table 7–15: Frequency Distribution for F-Subscribe, split by Society (N=1038)

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid	Total
Email Expected	67.01	38.52	43.38	24.28	126.24	36.58	**336.00**
Email Count	147	68	71	25	7	18	**336**
Email Percent	43.75	20.24	21.13	7.44	2.08	5.36	**100.00**
APA Expected	96.92	55.72	62.74	35.12	182.60	52.99	**486.00**
APA Count	43	31	41	36	267	68	**486**
APA Percent	8.85	6.38	8.44	7.41	54.94	13.99	**100.0**
APS Expected	17.35	9.97	11.23	6.29	32.69	9.47	**87.00**
APS Count	3	4	5	6	59	10	**87**
APS Percent	3.45	4.60	5.75	6.90	67.82	11.49	**100.0**
BPS Expected	25.73	14.79	16.65	9.32	48.47	14.04	**129.00**
BPS Count	14	16	17	8	57	17	**129**
BPS Percent	10.85	12.40	13.18	6.20	44.19	13.18	**100.0**
Overall Expected	207.00	119.00	134.00	75.00	390.00	113.00	**1038.00**
Overall Count	207	119	134	75	390	113	**1038**
Overall Percent	19.94	11.46	12.91	7.23	37.57	10.89	**100.00**

Table 7–15 shows the frequency distribution of the print respondents for the F-Subscribe variable split by Society. What is striking about this table is the preponderance of responses at the right-hand (or less frequent) end of the Likert scale. For each of the print subgroups, the combination of the Never and Seldom responses is always greater than 50%. As might be expected for a survey distributed via an email subscription list, the email survey respondents had a much higher frequency of usage with the majority of responses appearing towards the left-hand or more frequent end of the Likert scale. Nearly two-thirds of the email respondents subscribe to email lists frequently or regularly. For the print respondents, over half have never subscribed. Somewhat anomalously, 2% of the email respondents report never having subscribed to an email list (this of course includes the list used to distribute the survey instrument); perhaps they were subscribed by someone else?.

The difference between Society subgroups is significant ($p = 0.0378$). Not all of the cells in the contingency table contribute equally to this figure. Table 7–16shows the post-hoc cell contributions. The APA and APS respondents show significantly less frequency of use than the Email respondents, whereas the Email respondents have Never used an email list (or equivalent) significantly less than either the APA or APS respondents. The BPS respondents contribution is only significant for the Frequently cell.

Table 7–16: Post Hoc Cell Contributions for F-Subscribe, split by Society

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid
Email	13.281	6.138	5.465	.185	-16.333	-3.957
APA	-8.394	-4.826	-4.033	.212	10.840	3.014
APS	-4.023	-2.100	-2.082	-.124	6.085	.190
BPS	-2.761	.358	.097	-.480	1.657	.893

FTP

Table 7–17: Frequency Distribution for F-FTP, split by Survey (N=1038)

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid	Total
Email Expected	33.99	33.02	62.15	43.05	139.19	24.60	**336.00**
Email Count	82	64	92	39	19	40	**336**
Email Percent	24.40	19.05	27.38	11.61	5.65	11.90	**100.00**
Print Expected	71.01	68.98	129.85	89.95	290.81	51.40	**702.00**
Print Count	23	38	100	94	411	36	**702**
Print Percent	3.28	5.41	14.25	13.39	58.55	5.13	**100.00**
Overall Expected	105.00	102.00	192.00	133.00	430.00	76.00	**1038.00**
Overall Count	105	102	192	133	430	76	**1038**
Overall Percent	10.12	9.83	18.50	12.81	41.43	7.32	**100.00**

FTP was another distribution (technically access rather than distribution, but never mind) mechanism used by early e-journals. It has largely been overtaken by delivery of e-journals over the Web, either from ftp archives using the ftp protocol with a Web front end or directly via the http protocol.

Table 7–17 shows the breakdown of responses to the question about use of FTP (File Transfer Protocol). The print responses have been aggregated because some of the expected value cells in the Society-Print contingency table for this question were less than 5. Not all the contingency table cells contribute significantly to this value. Table 7–18. shows the post hoc cell contributions. Cells that do not contribute at the 0.05 percent level are shown as un-bolded to allow the significant cells to stand out.

Reading these two tables together reveals that the print respondents make frequent, regular or occasional use of FTP less than expected in contrast to the email respondents.

Table 7–18: Post Hoc Cell Contributions for F-FTP, split by Survey

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid
Email	10.56	6.9	5.10	-.80	-16.18	3.92
Print	-10.56	-6.9	-5.10	.80	16.18	-3.92

Gopher

Table 7–19: Frequency Distribution for F-Gopher, split by Survey (N=1038)

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid	Total
Email Expected	13.27	20.39	66.36	71.54	145.02	19.42	**336.00**
Email Count	25	30	110	104	41	26	**336**
Email Percent	7.44	8.93	32.74	30.95	12.20	7.74	**100.00**
Print Expected	27.73	42.61	138.64	149.46	302.98	40.58	**702.00**
Print Count	16	33	95	117	407	34	**702**
Print Percent	2.28	4.70	13.53	16.67	57.98	4.84	**100.00**
Overall Expected	41.00	63.00	205.00	221.00	448.00	60.00	**1038.00**
Overall Count	41	63	205	221	448	60	**1038**
Overall Percent	3.95	6.07	19.75	21.29	43.16	5.78	**100.00**

Gopher is now a relatively old technology, largely superseded by the Web. This transition was well under way when the email survey took place and had nearly completed by the time of the print survey. It is reasonable to assume that the email respondents were likely to be early adopters of new technologies and that their transition from Gopher to the Web (see next section) would have occurred earlier than the print respondents in any case. In other words, for the purposes of this question it would have been preferable if the surveys had been sequenced with the print survey first and email survey later.

Based on the actual data rather than the ideal data, there are still some useful things to say about the picture revealed in Table 7–19: Frequency distribution for F-Gopher split by Survey. The print responses have been aggregated on statistical grounds. The differences in responses between the print and email respondents are significant ($p < 0.0001$). The email respondents are using Gopher more than expected for responses of Frequently to Seldom and less than expected for Never. The print respondents are using Gopher less than expected, and not using it (the Never response) more than expected.

WWW

Table 7–20: Frequency Distribution for F-WWW, split by Society (N=1038)

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid	Total
Email Expected	87.72	57.29	79.31	33.34	67.01	11.33	**336.00**
Email Count	153	73	61	16	13	20	**336**
Email Percent	45.54	21.73	18.15	4.76	3.87	5.95	**100.00**
APA Expected	126.88	82.87	114.71	48.23	96.92	16.39	**486.00**
APA Count	86	67	124	64	136	9	**486**
APA Percent	17.70	13.79	25.51	13.17	27.98	1.85	**100.00**
APS Expected	22.71	14.84	20.53	8.63	17.35	2.93	**87.00**
APS Count	7	13	22	12	33	0	**87**
APS Percent	8.05	14.94	25.29	13.79	37.93	0.00	**100.00**
BPS Expected	33.68	22.00	30.45	12.80	25.73	4.35	**129.00**
BPS Count	25	24	38	11	25	6	**129**
BPS Percent	19.38	18.60	29.46	8.53	19.38	4.65	**100.00**
Overall Expected	271.00	177.00	245.00	103.00	207.00	35.00	**1038.00**
Overall observed	271	177	245	103	207	35	**1038**
Overall Percent	26.11	17.05	23.60	9.92	19.94	3.37	**100.00**

The Web has of course become most people's front-end to much of the Internet. Most respondents would probably use the Web for a range of uses unconnected to e-journals. While this question asked about the Web in an electronic publishing context, frequency of Web usage will naturally have been affected by these other uses. Nevertheless, the question does provide a good picture of the respondent's familiarity with a key electronic publishing technology.

The contingency table analysis for this combination of variables is valid as no expected cell has a value less than 5. The differences between the different Society subgroups are significant ($p < 0.0001$). The table of post-hoc cell contributions (Table 7–21) reveals that only some of the cells contribute significantly to this chi-square figure. The APA and BPS respondents show significantly less frequent or regular use of the Web than expected and significantly more non-use. A surprisingly high proportion of the print survey respondents (between 20 and 40%) have never used the Web. The email respondents show significantly more frequent or regular use and less non-use

Table 7–21: Post Hoc Cell Contributions for F-WWW, split by Society

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid
Email	9.86	2.77	-2.86	-3.85	-8.97	3.19
APA	-5.79	-2.63	1.36	3.28	6.08	-2.55
APS	-4.01	-.55	.39	1.26	4.39	-1.82
BPS	-1.86	.50	1.67	-.57	-.17	.86

CD-ROM

Table 7–22: Frequency Distribution for F-CDROM, split by Society (N=1038)

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid	Total
Email Expected	37.87	48.23	88.69	49.85	93.87	17.48	**336.00**
Email Count	51	41	92	71	54	27	**336**
Email Percent	15.18	12.20	27.38	21.13	16.07	8.04	**100.00**
APA Expected	54.78	69.76	128.29	72.10	135.78	25.28	**486.00**
APA Count	36	65	127	57	182	19	**486**
APA Percent	7.41	13.37	26.13	11.73	37.45	3.91	**100.00**
APS Expected	9.81	12.49	22.97	12.91	24.31	4.53	**87.00**
APS Count	4	10	25	12	36	0	**87**
APS Percent	4.60	11.49	28.74	13.79	41.38	0.00	**100.00**
BPS Expected	14.54	18.52	34.05	19.14	36.04	6.71	**129.00**
BPS Count	26	33	30	14	18	8	**129**
BPS Percent	20.16	25.58	23.26	10.85	13.95	6.20	**100.00**
Overall Expected	117.00	149.00	274.00	154.00	290.00	54.00	**1038.00**
Overall observed	117	149	274	154	290	54	**1038**
Overall Percent	11.27	14.35	26.40	14.84	27.94	5.20	**100.00**

CD-ROM was also used as a delivery platform in the early days of e-journals, and some researchers were still arguing until recently that they are the best solution to quality problems with on-line delivery [Jasperse1996]. This question tries to identify frequency of use of CD-ROM to access materials (not further defined).

The contingency table analysis for this combination of variables is valid as no expected cell has a value less than 5. The differences between the different Society subgroups are significant ($p < 0.0001$). The table of post-hoc cell contributions (Table7–23) reveals that only some of the cells contribute significantly to this chi-square figure. The picture is not as clear

Table 7–23: Post Hoc Cell Contributions for F-CDROM, split by Society

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid
Email	2.754	-1.368	.498	3.947	-5.895	2.844
APA	-3.694	-.845	-.182	-2.643	6.407	-1.760
APS	-2.056	-.795	.517	-.286	2.919	-2.283
BPS	3.409	3.886	-.865	-1.360	-3.783	.546

cut as for some of the previous variables but some patterns can be discerned. The APA and APS respondents make less frequent use of CD-ROMs than expected while the BPS and Email respondents make more frequent use than expected. The BPS and Email respondents make no use of CD-ROMs less than expected and the APA and APS respondents make no use more than expected. As a general summary, for this variable the BPS and Email respondents show greater patterns of use than the APA and APS respondents.

E-journals

Table 7–24: Frequency Distribution for F-Views, split by Society (N=1038)

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid	Total
Email Expected	21.36	43.70	74.45	52.12	127.86	16.51	**336.00**
Email Count	51	108	112	41	5	19	**336**
Email Percent	15.18	32.14	33.33	12.20	1.49	5.65	**100.00**
APA Expected	30.90	63.21	107.69	75.38	184.94	23.88	**486.00**
APA Count	9	13	80	79	283	22	**486**
APA Percent	1.85	2.67	16.46	16.26	58.23	4.53	**100.00**
APS Expected	5.53	11.32	19.28	13.49	33.11	4.27	**87.00**
APS Count	0	3	13	18	52	1	**87**
APS Percent	0.00	3.45	14.94	20.69	59.77	1.15	**100.00**
BPS Expected	8.20	16.78	28.58	20.01	49.09	6.34	**129.00**
BPS Count	6	11	25	23	55	9	**129**
BPS Percent	4.65	8.53	19.38	17.83	42.64	6.98	**100.00**
Overall Expected	66.00	135.00	230.00	161.00	395.00	51.00	**1038.00**
Overall observed	66	135	230	161	395	51	**1038**
Overall Percent	6.36	13.01	22.16	15.51	38.05	4.91	**100.00**

The previous questions in the Electronic Publishing section of the survey focused on the underlying technologies and the respondents' usage of these. The final three questions are more

concerned with respondent's activities with respect to electronic publishing. This question asks about viewing electronic journals in general to try and get some baseline measure of activity in this area.

The contingency table analysis for this combination of variables is valid as no expected cell has a value less than 5. The differences between the different Society subgroups are significant ($p < 0.0001$). The table of post-hoc cell contributions (Table7–25) reveals that only some of the cells contribute significantly to this chi-square figure. In particular, the email re-

Table 7–25: Post Hoc Cell Contributions for F-Views, split by Society

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid
Email	8.057	12.681	5.998	-2.037	-16.787	.765
APA	-5.583	-9.285	-4.147	.622	12.563	-.541
APS	-2.539	-2.769	-1.693	1.394	4.359	-1.697
BPS	-.849	-1.616	-.812	.777	1.145	1.159

spondents are much more likely than the APA or BPS respondents to view e-journals; over double the expected viewing rate in the case of Frequent or Regular usage. (The BPS post-hoc cell contributions are not statistically significant). The APA and APS respondents are much more likely than expected (nearly 60% of each population) to have never viewed an e-journal. This is a very clear-cut difference between the email and print survey populations

Psyche

This question was only included on the print survey and asked how often the respondents viewed a specific e-journal: *Psyche*. There was no point in asking this question in the email survey as one could reasonably assume that everyone was a frequent or regular user of the e-journal. This variable makes it possible to see how many *Psyche* readers there are in the print survey population, and if there is any significant difference in responses between *Psyche* readers and nonreaders in the print survey population.

Table 7–26 speaks for itself: the overwhelming majority of the print respondents have never viewed Psyche. Only 5.4% of respondents have *ever* viewed *Psyche*, and so the print and email survey populations can effectively be regarded as almost entirely disjoint. From the point of view of the research design this is a good outcome because it means that the print survey population can be compared directly with the email survey population without having

Table 7–26: Frequency Distribution for F-Psyche, split by Society (N=1038)

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid	Total
APA Count	1	1	8	15	434	27	**486**
APA Percent	.21	.21	1.65	3.09	89.30	5.56	**100.00**
APS Count	0	2	2	1	81	1	**87**
APS Percent	0.00	2.30	2.30	1.15	93.10	1.15	**100.00**
BPS Count	2	3	4	6	104	10	**129**
BPS Percent	1.55	2.33	3.10	4.65	80.62	7.75	**100.00**
Overall observed	3	6	14	22	619	38	**702**
Overall Percent	0.43	0.86	1.99	2.12	88.18	5.41	**100.00**

to worry about the overlap set. Given these numbers, there is also little point in using this variable to break down the print survey responses. No statistics can be calculated for the print subgroups because of the small expected cell sizes.

Publish electronically

Table 7–27: Frequency Distribution for F-Publishes, split by Survey (N=1038)

	Frequently	Regularly	Occasionally	Seldom	Never	Invalid	Total
Email Expected	5.83	6.15	22.01	36.25	242.45	23.31	**336.00**
Email Count	14	15	49	80	147	31	**336**
Email Percent	4.17	4.46	14.58	23.81	43.75	9.23	**100.00**
Print Expected	12.17	12.85	45.99	75.75	506.55	48.69	**702.00**
Print Count	4	4	19	32	602	41	**702**
Print Percent	.57	.57	2.71	4.56	85.75	5.84	**100.00**
Overall Expected	18.00	19.00	68.00	112.00	749.00	72.00	**1038.00**
Overall Count	18	19	68	112	749	72	**1038**
Overall Percent	1.73	1.83	6.55	10.79	72.16	6.94	**100.00**

This variable was designed to find out how many of the survey population were actually publishing electronically themselves as opposed to reading other people's electronic publications. Clearly it is necessary for authors as well as readers to migrate to on-line electronic publishing in order for it to become a success.

Table 7–27: Frequency Distribution for F-Publishes, split by Survey reveals that the print respondents (aggregated on statistical grounds) publish electronically much less frequently

than expected and also less than the email respondents. The differences are significant at $p<0.0001$. Two data points in particular are noteworthy. The overwhelming majority of the print respondents have never published electronically (less than 10% have ever done so). In contrast, nearly half of the email respondents (47%) have published electronically, and nearly 10% do so frequently or regularly.

7.3.4 Advantages of electronic scholarly publishing

These questions asked respondents to rate a number of possible advantages of electronically published scholarly articles on a five-point Likert scale running from Strongly Agree through to Strongly Disagree (SD). The advantages were gathered from a close reading of the critical literature dealing with electronic scholarly journals. Respondents were also encouraged to add other advantages if they wished to. In fact only 17 of the 1038 respondents (1.6%) added one additional advantage and only 4 added a second.

All of the contingency tables for the print subgroups for this set of questions contained cells with expected values of less than 5 and so the print responses were aggregated and compared with the email responses for each question. Even then, it was only possible to derive valid statistical results for two of the advantages questions: A-Paper and A-Affordability. This was because most of the contingency tables by Survey also had expected cells with values less than 5. Therefore, for the majority of the advantages questions only the observed values will be reported and no statistical results are available.

Speed of publication

Figure 7–1: Frequency Histogram for A-Speed, split by Survey (N=1038)

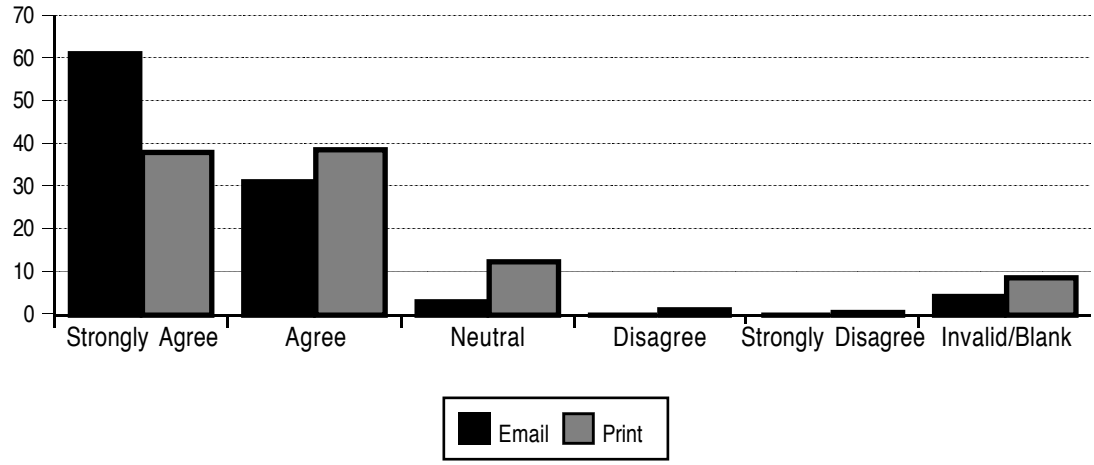


Table 7–28: Frequency Distribution for A-Speed, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Count	205	106	10	0	0	17	**336**
Email Percent	61.01	31.55	2.98	0.00	0.00	4.47	**100.00**
Print Count	267	274	87	8	3	63	**702**
Print Percent	38.03	39.03	12.39	1.14	0.43	8.98	**100.00**
Overall Count	472	380	97	8	3	80	**1038**
Overall Percent	45.47	36.61	9.34	0.77	0.29	7.52	**100.00**

One of the advantages of e-journals is that they can be published as soon as the articles have been accepted. There is no need to wait for a gap some months hence in a crowded printing schedule. In some e-journals there is not even a wait for an issues to become ready; articles are published on their own as soon as they are ready. With the increasing pace of research delay to publication in the print literature is becoming a serious problem.

Both groups of respondents are very positive about this advantage, with almost 90% of the responses at Neutral or better. The email respondents are definitely more positive than the print respondents, strongly agreeing in a ratio of 3:2. Notably nearly 10% of the print respondents left this question blank. Figure 7–1 shows the pattern of responses of declining positive responses very clearly.

24 hour access

Figure 7–2: Frequency Histogram for A-24 Hour, split by Survey (N=1038)

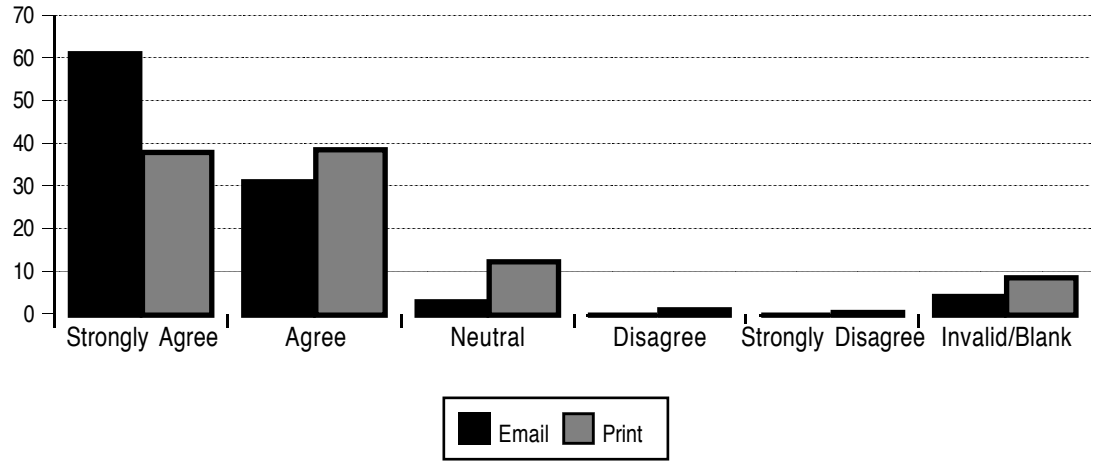


Table 7–29: Frequency Distribution for A-24 Hour, split by Survey

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Count	163	124	27	4	0	18	**336**
Email Percent	48.51	36.90	8.04	1.19	0.00	5.36	**100.00**
Print Count	345	255	43	5	1	53	**702**
Print Percent	49.15	36.32	6.13	0.71	0.14	7.54	**100.00**
Overall Count	508	379	70	9	1	71	**1038**
Overall Percent	48.94	36.51	6.74	0.87	0.10	6.84	**100.00**

Unless one subscribes personally to a journal (or has a large collection of copied articles/reprints) one is restricted to the hours in which the library is open. On-line journals are available all the time (provided one has the necessary equipment, the licensing arrangements are in place and the network is working!). This is still likely to be significantly better than the library hours.

In this case the percentages for each Likert rating are nearly identical for each group of respondents, with the exception of a slightly greater number of blank answers. Both groups are again very positive about this advantage with nearly 90% of answers at the level of Strongly Agree or Agree.

Convenience

Figure 7–3: Frequency Histogram for A-Convenience, split by Survey (N=1038)

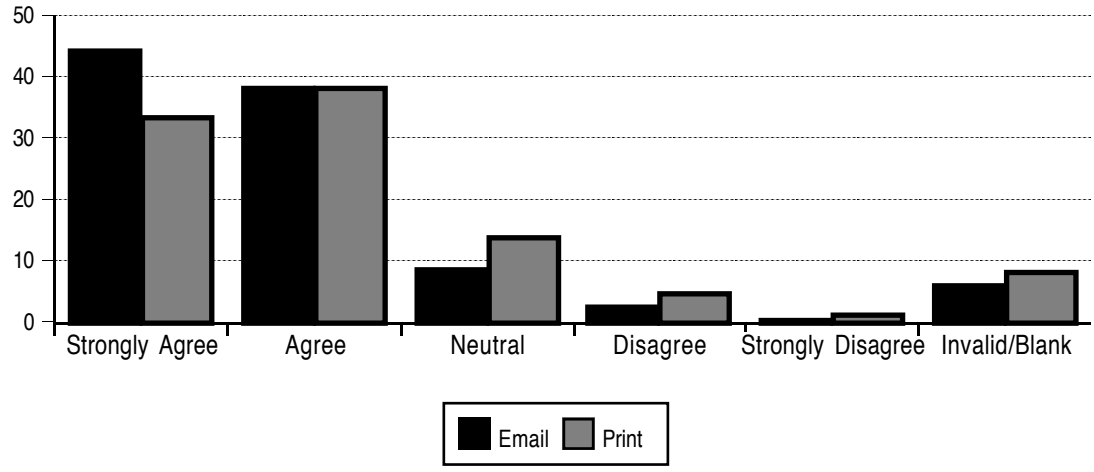


Table 7–30: Frequency Distribution for A-Convenience, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Count	149	128	29	9	1	20	**336**
Email Percent	44.35	38.10	8.63	2.68	0.30	5.95	**100.00**
Print Count	236	270	97	34	8	57	**702**
Print Percent	33.62	38.46	13.82	4.84	1.14	8.11	**100.00**
Overall Count	385	398	126	43	9	77	**1038**
Overall Percent	37.09	38.34	12.14	4.14	0.87	7.42	**100.00**

By convenience this question meant that e-journals were easy to access. An example might be that e-journals are potentially available from the researcher's desk rather than requiring a trip to the serials section of the library.

The answers to this question show a very similar pattern of responses to the A-Speed questions (Table 7–28). In this case the email respondents do not exceed the level of strong agreement of the print respondents to the same extent. In fact on the Agree measure the reverse is the case.

Feedback

Figure 7–4: Frequency Histogram for A-Feedback, split by Survey (N=1038)

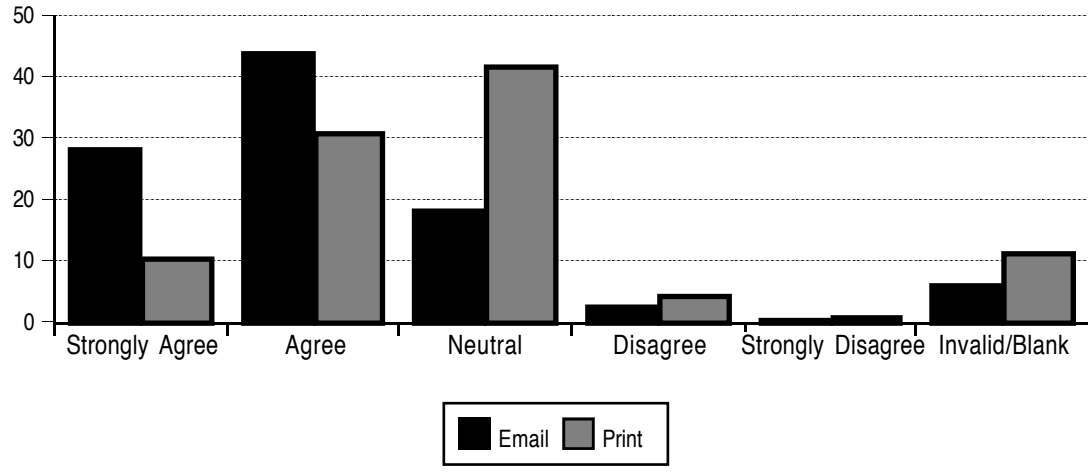


Table 7–31: Frequency Distribution for A-Feedback, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Count	95	147	62	9	2	21	**336**
Email Percent	28.27	43.75	18.45	2.68	0.60	6.25	**100.00**
Print Count	74	217	294	32	6	79	**702**
Print Percent	10.54	30.91	41.88	4.56	0.85	11.25	**100.00**
Overall Count	169	364	356	41	8	100	**1038**
Overall Percent	16.28	35.07	34.30	3.95	0.77	9.64	**100.00**

The feedback processes for the print literature can be divided into public and private. Public feedback is largely restricted to letters to the editor which often appear one (or more often two) issues of the journal later and are necessarily restricted in space. Private feedback occurs via letter or email (if an email address is provided). E-journals allow for immediate private feedback via clickable email links. Public feedback can occur via on-line comments (as used by JIME), an associated mailing list or newsgroup or a Web-based conference. The reader of the article is much more aware of the discourse around the article and the community of scholars actively debating the ideas contained in it.

The email respondents were markedly more in agreement (nearly 70% strongly agreeing or agreeing) with this advantage than were the print respondents (only 40% at the same level of agreement). Over 40% of the print respondents were neutral about this advantage.

Reduced paper.

Figure 7–5: Frequency Histogram for A-Paper, split by Survey

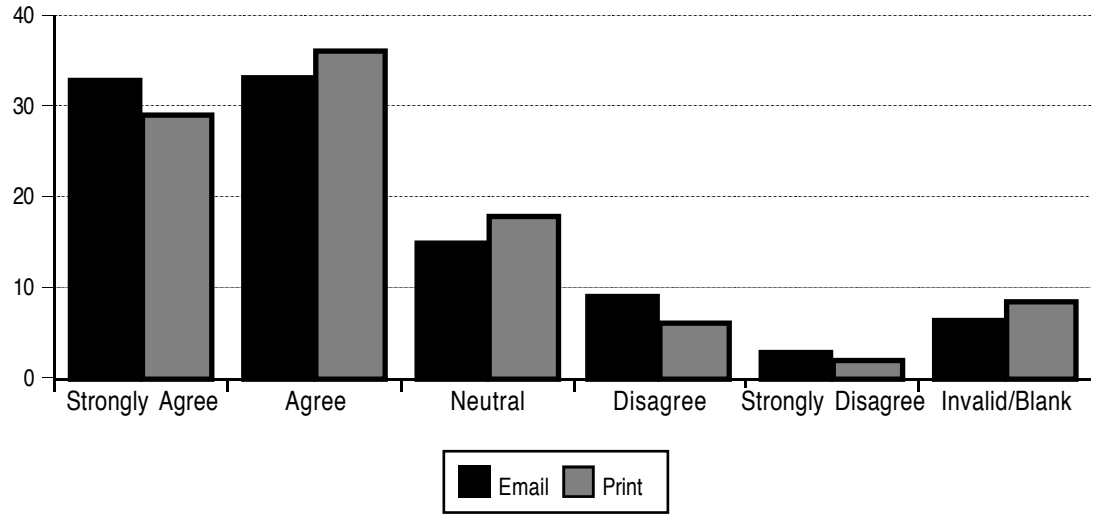


Table 7–32: Frequency Distribution for A-Paper, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Expected	102.29	118.47	56.97	23.95	8.09	26.22	**336.00**
Email Count	111	112	50	31	10	22	**336**
Email Percent	33.04	33.33	14.88	9.23	2.98	6.55	**100.00**
Print Expected	213.71	247.53	119.03	50.05	16.91	54.78	**702.00**
Print Count	205	254	126	43	15	59	**702**
Print Percent	29.20	36.18	17.95	6.13	2.14	8.4	**100.00**
Overall Expected	316.00	366.00	176.00	74.00	25.00	81.00	**1038.00**
Overall Count	316	366	176	74	25	81	**1038**
Overall Percent	30.44	35.26	16.96	7.13	2.41	7.8	**100.00**

Despite the always imminent but apparently illusory advent of the paperless office, a shift to scholarly e-journals holds out the promise of some reduction in paper consumption. There is some anecdotal evidence that most readers of a journal are interested in only about 10% of the articles in any given issue. Having the articles available on-line allows a reader to view them onscreen before printing them if necessary. John Warnock, CEO of Adobe Systems, talks about moving from a Print then Distribute model to Distribute then Print. E-journals provide this electronic distribution function very well. It is also possible to keep e-journal articles in electronic form as part of a scholar's personal document database for later use and reuse.

There is a statistically significant (p = 0.0011) difference between email and print respondents for this question. The email respondents strongly agree more than expected, but agree less than expected. Overall, agreement with this advantage is lower (more neutral, disagree and strongly disagree responses) than for the advantages discussed already. The respondents are still positive towards this advantage (65% agree or strongly agree overall).

Searching

Figure 7–6: Frequency Histogram for A-Searching, split by Survey

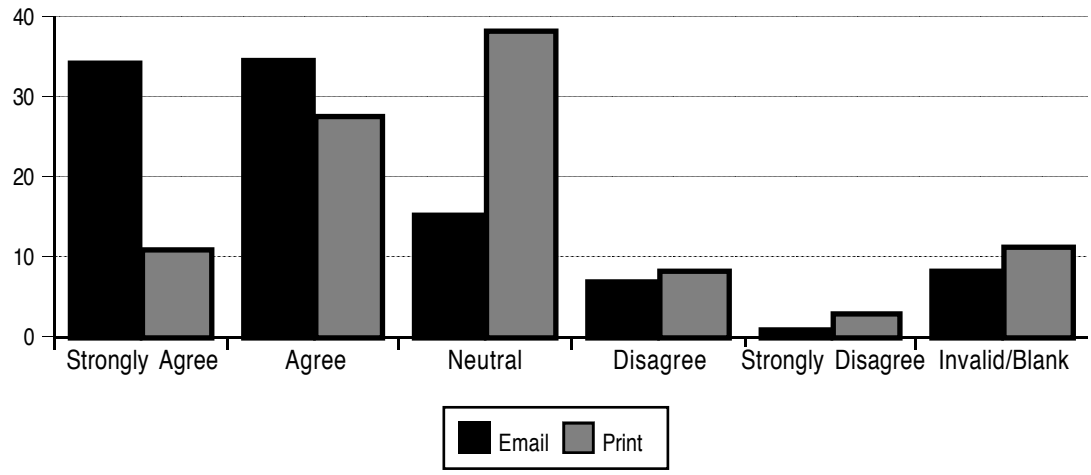


Table 7–33: Frequency Distribution for A-Searching, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Count	104	118	66	21	4	23	**336**
Email Percent	30.95	35.12	19.64	6.25	1.19	6.84	**100.00**
Print Count	160	261	144	64	6	67	**702**
Print Percent	22.79	37.18	20.51	9.12	0.85	9.54	**100.00**
Overall Count	264	379	210	85	10	90	**1038**
Overall Percent	25.43	36.51	20.23	8.19	0.96	8.66	**100.00**

One of the problems with print publications is that they are not directly searchable. For this reason the last twenty years has seen a dramatic growth in the availability of searchable document 'surrogates' like indexes and abstracts (initially via a proprietary on-line service, then on CD-ROM, and now probably via the Internet). One of the attractions of e-journal articles is the ability to search the entire text of a journal (or even set of journals) directly.

Over 50% of the respondents either strongly agreed or agreed that this was an advantage. Nearly 20% of each population was neutral. Less than 10% of the respondents disagreed or strongly disagreed that ease of searching was an advantage.

Multimedia

Figure 7–7: Frequency Histogram for A-Multimedia, split by Survey

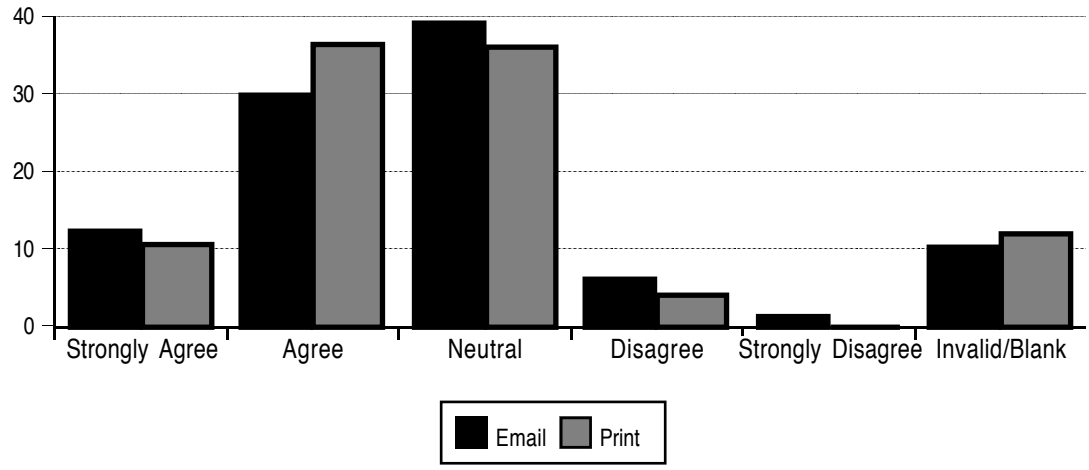


Table 7–34: Frequency Distribution for A-Multimedia, split by Survey

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Count	42	101	132	21	5	35	**336**
Email Percent	12.50	30.06	39.29	6.25	1.49	10.42	**100.00**
Print Count	76	257	255	29	1	84	**702**
Print Percent	10.83	36.61	36.32	4.13	0.14	11.96	**100.00**
Overall Count	118	358	387	50	6	119	**1038**
Overall Percent	11.37	34.49	37.28	4.82	0.58	11.47	**100.00**

One of the advantages of moving away from print and towards electronic delivery of scholarly journals is the ability to include more multimedia content (more because paper is still a medium). Such content might just be colour (which was too expensive in print) or it might be dynamic content like video clips or animations. It might even be sound files or embedded simulations. These are all things that are unfamiliar within the context of a scholarly journal but which could obviously add to the quality of the communicative transaction.

Both print and email respondents were more tentative about this advantage. A higher proportion than for the earlier advantage questions (around 10% of both populations) left this question blank and neutral was the most popular response for both populations. The level of

strong agreement was also very low. This perhaps reflects users' uncertainty about (and lack of familiarity with) such content.

Affordability

Figure 7–8: Frequency Histogram for A-Affordability, split by Survey

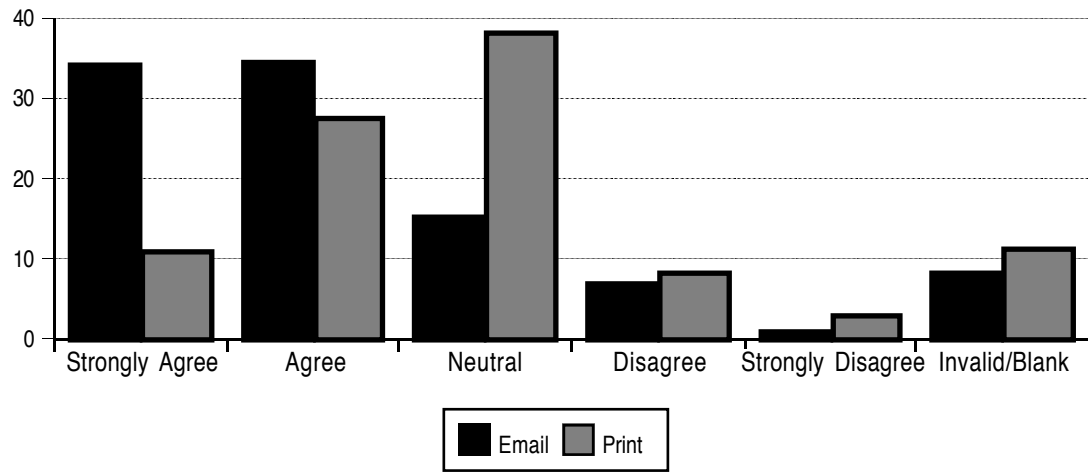


Table 7–35: Frequency Distribution for A-Affordability, split by Survey

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Count	115	116	51	23	3	30	**336**
Email Percent	34.23	34.52	15.18	6.85	0.89	8.33	**100.00**
Print Count	78	194	269	59	22	80	**702**
Print Percent	11.11	27.64	38.32	8.40	3.13	11.4	**100.00**
Overall Count	193	310	320	82	25	108	**1038**
Overall Percent	18.59	29.87	30.83	7.90	2.41	10.4	**100.00**

One of the ongoing debates around e-journals has been the effect of removing the need for physical distribution of paper on the overall cost. A number of e-journals are still available free on-line but some are only available on a subscription or site license basis. This question sought to determine if respondents felt that affordability of e-journals was an advantage.

This question reveals a large difference between the two respondent populations. The differences are significant ($p < 0.001$). The email respondents were much more optimistic about the advantages of e-journals with respect to affordability, with nearly 70% of responses at Strongly Agree or Agree. The print respondents only had less than 40% of responses at the same level. The print respondents also had a much larger percentage of neutral responses.

7.3.5 Disadvantages of electronic scholarly publishing

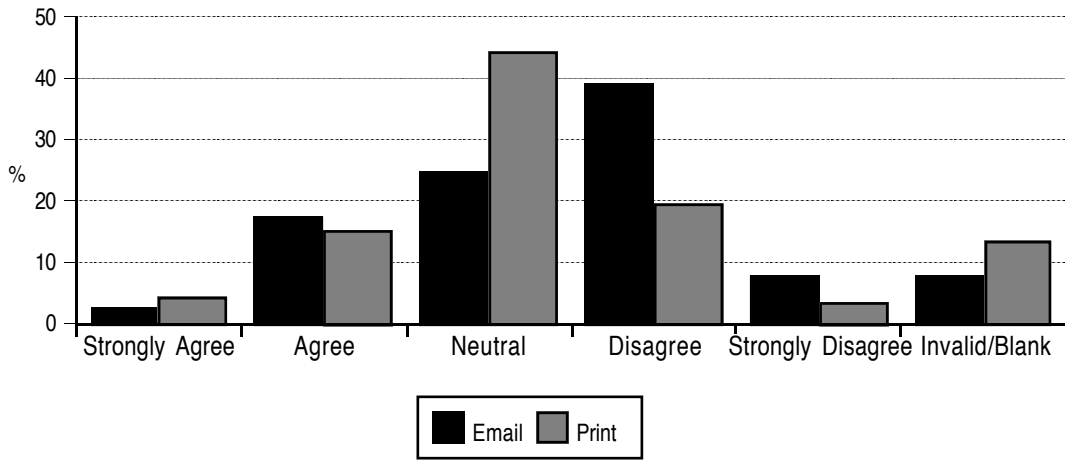
These questions asked respondents to rate a number of possible perceived disadvantages of electronically published scholarly articles on a five-point Likert scale running from Strongly Agree through to Strongly Disagree (SD). As with the advantages, the disadvantages were gathered from a close reading of the critical literature dealing with electronic scholarly journals. Respondents were encouraged to add other disadvantages if they wished to. Only 35 of the 1038 respondents (3.4%) added one additional disadvantage and only 8 added a second.

All of the contingency tables for the print subgroups for this set of questions contained cells with expected values of less than 5 and so the print responses were aggregated and compared with the email responses for each question. This provided valid statistical results for all of the Disadvantages questions; in each case the differences between the print and email responses were significant at $p < 0.0001$.

A number of respondents had problems with this set of questions. For most of the questions a little over 10% of the print respondents and a little under 10% of the email respondents provided either blank (assumed to mean insufficient knowledge to provide an answer or no opinion) or invalid (most often a question mark) answers.

Quality.

Figure 7–9: Frequency Histogram for D-Quality, split by Survey (N=1038)



One of the standard criticisms of on-line publishing (of all sorts) is the lack of quality. This is perhaps influenced by the number of student homepages visible on the Web in its early days. Of course, quality is medium independent [Harnad1995a] and there is no reason to expect that there is anything inherent in on-line publishing that makes quality more difficult to achieve in such an environment.

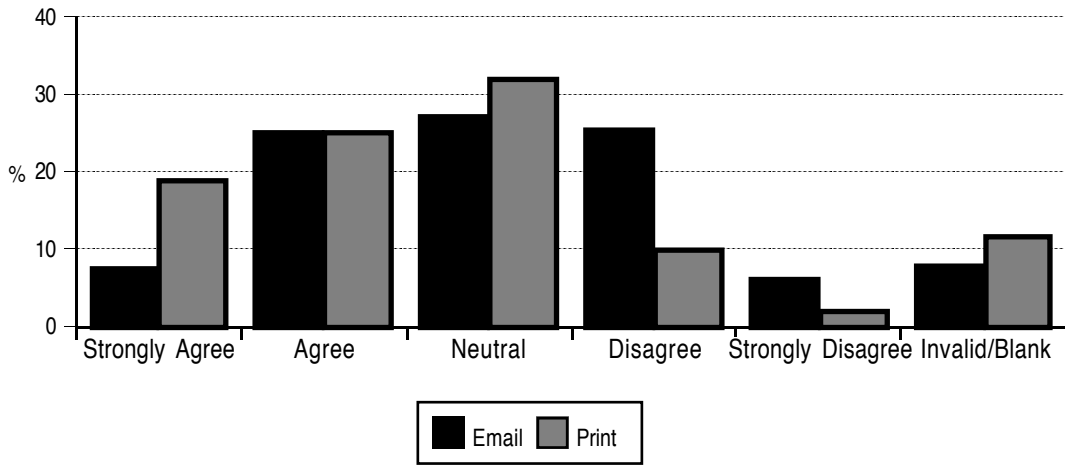
Table 7–36: Frequency Distribution for D-Quality, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Totals
Email Expected	12.95	53.41	127.21	87.08	16.19	39.17	**336.00**
Email Count	9	59	83	131	27	27	**336**
Email Percent	2.68	17.56	24.70	38.99	8.04	8.04	**100.00**
Print Expected	27.05	111.59	265.79	181.93	33.82	81.83	**702.00**
Print Count	31	106	310	138	23	94	**702**
Print Percent	4.42	15.10	44.16	19.66	3.28	13.39	**100.00**
Overall Expected	40.00	165.00	393.00	269.00	50.00	121.00	**1038.00**
Overall Count	40	165	393	269	50	121	**1038**
Overall Percent	3.85	15.90	37.86	25.92	4.82	11.66	**100.00**

This question asks respondents to assess poor quality as a disadvantage of electronically published scholarly articles. In contrast to most of the Advantages questions, the responses to this question are clustered much more closely around the Neutral position in the middle of the Likert scale. The (statistically significant) differences between the email and print respondents are most noticeable for the Neutral (where the print respondents agree at almost double the rate of the email respondents) and Disagree (where the situation is almost reversed) dimensions. Clearly the email respondents are less concerned about this perceived disadvantage.

Refereeing

Figure 7–10: Frequency Histogram for D-Refereeing, split by Survey (N=1038)



Another often expressed concern about on-line publishing is the lack of the quality control processes inherent in most print publishing, and usually expressed as ‘refereeing’ or ‘peer review’. In contrast to the previous question, the responses to this question are spread across

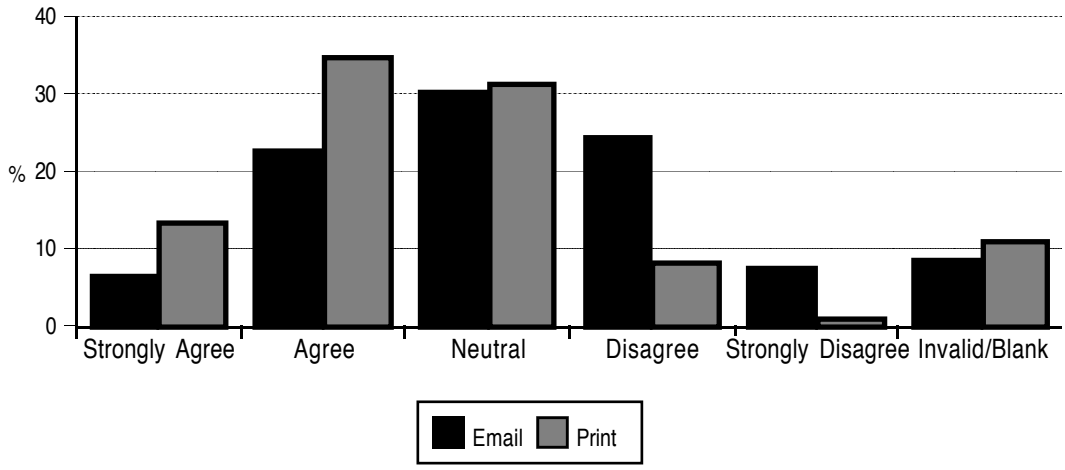
Table 7–37: Frequency Distribution for D-Refereeing, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid/ Blank	Total
Email Expected	51.79	84.49	102.29	50.17	11.65	35.61	**336.00**
Email Count	26	85	91	86	21	27	**336**
Email Percent	7.74	25.30	27.08	25.60	6.25	8.04	**100.00**
Print Expected	108.21	176.51	213.71	104.83	24.35	74.39	**702.00**
Print Count	134	176	225	69	15	83	**702**
Print Percent	19.09	25.07	32.05	9.83	2.14	11.82	**100.00**
Overall Expected	160.00	261.00	316.00	155.00	36.00	110.00	**1038.00**
Overall Count	160	261	316	155	36	110	**1038**
Overall Percent	15.41	25.14	30.44	14.93	3.47	10.60	**100.00**

the Strongly Agree, Agree and Neutral positions on the Likert scale. The differences between the email and print respondents are most evident for Strongly Agree (where the print respondents take a much stronger position) and Disagree or Strongly Disagree (where the email respondents are much more prevalent). While both groups are concerned, the email respondents are much less so.

Copyright

Figure 7–11: Frequency Histogram for D-Copyright, split by Survey (N=1038)



A range of concerns are often expressed about copyright in an on-line environment. In the experience of the author, most electronic publishing conferences seem to always end up discussing copyright and economics on the last day. Naturally, providers of commercial content of all sorts are concerned that they may lose control of that content, and therefore of the revenue stream that should flow from it. This applies also to journal publishers whether for prof-

Table 7–38: Frequency Distribution for D-Copyright, split by Survey (N=1038)

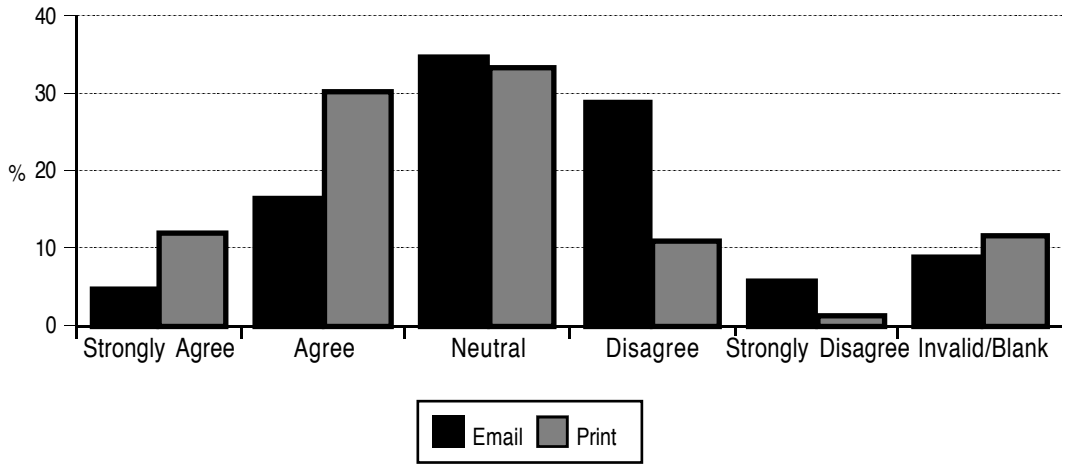
	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid Blank	Totals
Email Expected	37.55	103.91	104.23	45.32	10.36	34.64	**336.00**
Email Count	22	76	102	82	25	29	**336**
Email Percent	6.55	22.62	30.36	24.40	7.44	8.63	**100.00**
Print Expected	78.45	217.09	217.77	94.68	21.64	72.36	**702.00**
Print Count	94	245	220	58	7	78	**702**
Print Percent	13.39	34.90	31.34	8.26	1.00	11.11	**100.00**
Overall Expected	116.00	321.00	322.00	140.00	32.00	107.00	**1038.00**
Overall Count	116	321	322	140	32	107	**1038**
Overall Percent	11.18	30.92	31.02	13.49	3.08	10.31	**100.00**

it or otherwise. Loss of control of on-line journal articles can potentially translate into loss of sales of subscriptions or individual articles.

This question sought to test respondents' concerns about copyright. The patterns of answers are broadly similar to those for the Refereeing question. The print respondents are more concerned than are the email respondents. They also agree with this disadvantage more than do the email respondents.

Plagiarism

Figure 7–12: Frequency Histogram for D-Plagiarism, split by Survey (N=1038)



From the point of view of the scholar (rather than the publisher), plagiarism is a more serious concern than uncontrolled distribution of content. In fact, most scholars would probably *prefer* the widest possible distribution for their ideas! Of course, plagiarism is not restricted to electronic information; it is just that the ability to easily copy and paste is inherent in most

Table 7–39: Frequency Distribution for D-Plagiarism, split by Survey (N=1038)

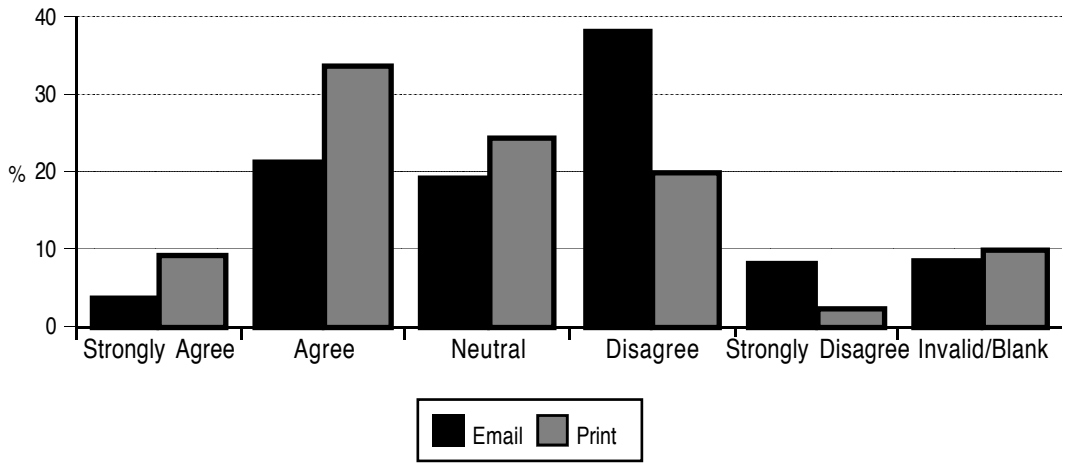
	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid Blank	Total
Email Expected	32.69	87.40	113.62	56.65	9.39	36.25	**336.00**
Email Count	16	56	117	97	20	30	**336**
Email Percent	4.76	16.67	34.82	28.87	5.95	8.93	**100.00**
Print Expected	68.31	182.60	237.38	118.35	19.61	75.75	**702.00**
Print Count	85	214	234	78	9	82	**702**
Print Percent	12.11	30.48	33.33	11.11	1.28	11.68	**100.00**
Overall Expected	101.00	270.00	351.00	175.00	29.00	112.00	**1038.00**
Overall Count	101	270	351	175	29	112	**1038**
Overall Percent	9.73	26.01	33.82	16.86	2.79	10.79	**100.00**

forms of the medium. This question asked whether respondents felt that plagiarism would increase in an environment of electronically published articles.

With the exception of a small difference around the neutral choice, the pattern of responses for this question is almost identical for the Copyright question. It is tempting to assume that the same respondents were responsible for identical answers to both questions. In fact, of the 116 who selected Strongly Agree for the D-Copyright question, only 63 also selected Strongly Agree for D-Plagiarism. This ratio of 2 to 1 is repeated for each level of the Likert scale. The respective numbers are Agree: 321 and 171, Neutral: 322 and 219, Disagree: 140 and 82 and Strongly Disagree: 32 and 14.

Skills

Figure 7–13: Frequency Histogram for D-Skills, split by Survey (N=1038)



We all learn how to read at school and then learn how to work with the literatures of our professions at university or in the workplace. No-one needs to attend half-day courses in ad-

Table 7–40: Frequency Distribution for D-Skills, split by Survey (N=1038)

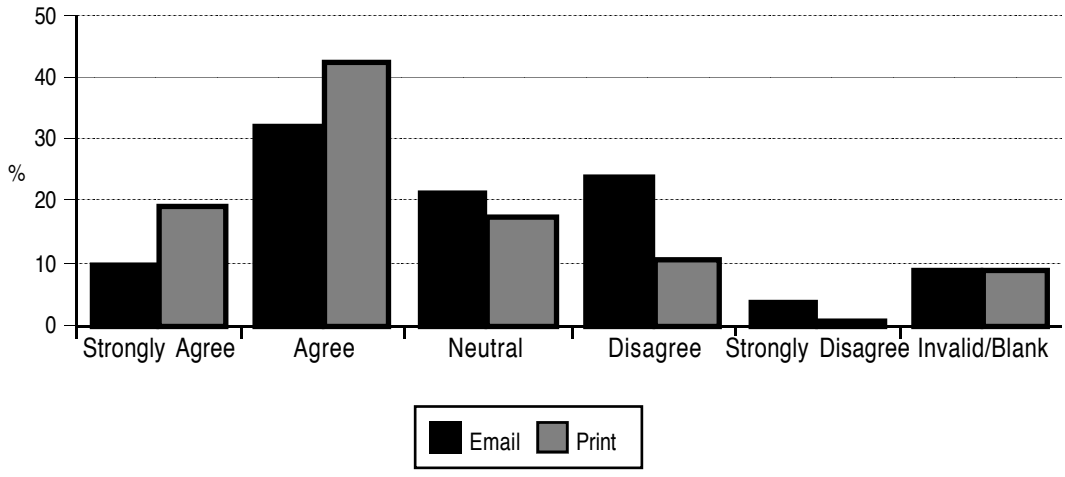
	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid Blank	Totals
Email Expected	25.57	100.02	77.04	87.08	14.57	31.72	**336.00**
Email Count	13	72	65	129	28	29	**336**
Email Percent	3.87	21.43	19.35	38.39	8.33	8.63	**100.00**
Print Expected	53.43	208.98	160.96	181.92	30.43	66.28	**702.00**
Print Count	66	237	173	140	17	69	**702**
Print Percent	9.40	33.76	24.64	19.94	2.42	9.83	**100.00**
Overall Expected	79.00	309.00	238.00	269.00	45.00	98.00	**1038.00**
Overall Count	79	309	238	269	45	98	**1038**
Overall Percent	7.61	29.77	22.93	25.92	4.34	9.44	**100.00**

vanced journal reading. And yet the need for special skills to access electronic journals is sometimes cited as a barrier to uptake. This question asked respondents to rank this need as a potential disadvantage.

Once again, the email respondents were much less inclined to regard this as an issue, with the largest single response (nearly 40%) being one of disagreement. On the other hand, over 40% of the print respondents agreed or strongly agreed that this was a disadvantage.

Equipment

Figure 7–14: Frequency Histogram for D-Equipment, split by Survey (N=1038)



Assuming no physical disability, no-one needs special equipment to read print based scholarly journals. This is not the case for on-line journals. Depending on the level of sophistication of the journal, the reader may require a colour screen, a fast direct network connection, sound output, lots of Random Access Memory (RAM) and a fast Central Processing Unit (CPU) to receive the full effect. For those concerned about access to information, this can be

Table 7–41: Frequency Distribution for D-Equipment, split by Survey (N=1038)

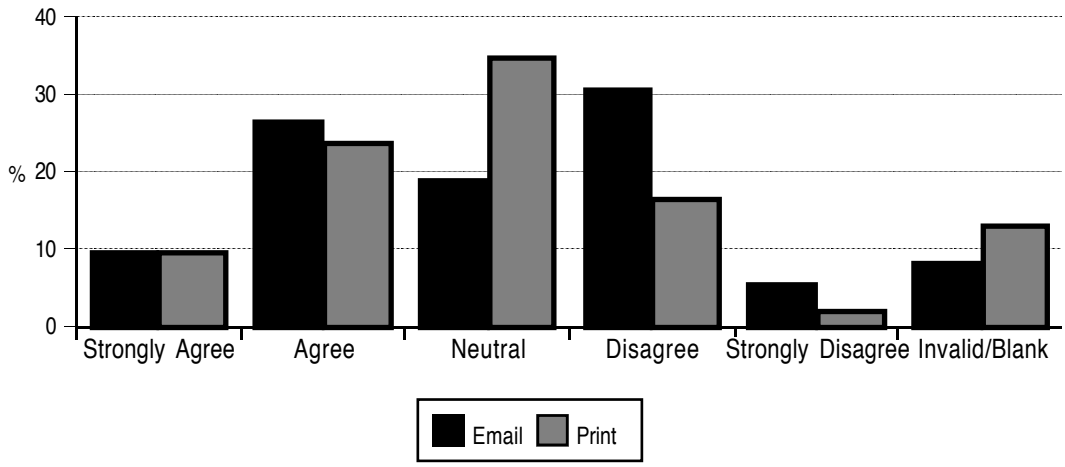
	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid Blank	Total
Email Expected	54.38	131.42	62.80	50.82	6.47	30.10	**336.00**
Email Count	33	108	72	80	13	30	**336**
Email Percent	9.82	32.14	21.43	23.81	3.87	8.93	**100.00**
Print Expected	113.62	274.58	131.20	106.18	13.53	62.90	**702.00**
Print Count	135	298	122	77	7	63	**702**
Print Percent	19.23	42.45	17.38	10.97	1.00	8.97	**100.00**
Overall Expected	168.00	406.00	194.00	157.00	20.00	93.00	**1038.00**
Overall Count	168	406	194	157	20	93	**1038**
Overall Percent	16.18	39.11	18.69	15.13	1.93	8.96	**100.00**

serious issue. This question asked respondents to rate the need for special equipment to access electronic journal articles as a potential disadvantage.

For this question both groups of respondents had patterns of answers that moved towards the strongly agree end of the Likert scale. Over 60% of the print respondents and over 40% of the email respondents either agreed or strongly agreed that this was a disadvantage. The numbers of respondents who strongly disagreed with this disadvantage were the lowest for any of the disadvantage questions. Clearly this is an issue of concern to both groups (although not unexpectedly more so for the print respondents).

Format

Figure 7–15: Frequency Histogram for D-Format, split by Survey (N=1038)



The format of e-journals has sometimes been criticised for not being reader friendly. In addition to the possible negative effects of reading large amounts of text onscreen, readers miss the ability to pick up a journal issue and quickly flick through the table of contents and article

Table 7–42: Frequency Distribution for D-Format, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid Blank	Totals
Email Expected	32.37	83.19	99.70	70.89	11.01	38.84	**336.00**
Email Count	33	89	64	103	19	28	**336**
Email Percent	9.82	26.49	19.05	30.65	5.65	8.33	**100.00**
Print Expected	67.63	173.81	208.30	148.11	22.99	81.16	**702.00**
Print Count	67	168	244	116	15	92	**702**
Print Percent	9.54	23.93	34.76	16.52	2.14	13.11	**100.00**
Overall Expected	100.00	257.00	308.00	219.00	34.00	120.00	**1038.00**
Overall Count	100	257	308	219	34	120	**1038**
Overall Percent	9.63	24.76	29.67	21.10	3.28	11.56	**100.00**

pages looking for items of interest. E-journals can (and do) provide alternate navigation mechanisms but these are still different. Once a desired article has been located on-line, the Web provides limited formatting options and a screen resolution significantly lower than print. Adobes' Portable Document Format is an excellent way of distributing 'electronic paper' but is also hard to read on screen. This question sought to probe respondents' attitudes to this issue.

Both groups of respondents had very similar percentages of responses in the strongly agree and agree categories. Their responses in the neutral and disagree categories are almost mirror-images: print respondents are more likely to be neutral (almost a 7:4 ratio) and email respondents more likely to disagree (almost a 2:1 ratio).

Communications Costs

Figure 7–16: Frequency Histogram for D-Costs, split by Survey (N=1038)

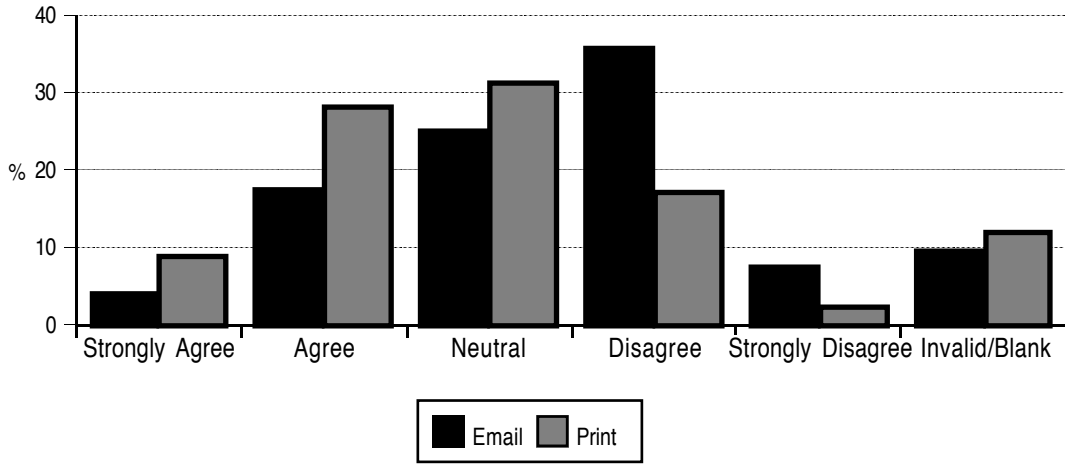


Table 7–43: Frequency Distribution for D-Costs, split by Survey (N=1038)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Invalid Blank	Total
Email Expected	24.60	83.51	99.05	77.69	13.27	37.87	**336.00**
Email Count	14	59	85	120	25	33	**336**
Email Percent	4.17	17.56	25.30	35.71	7.44	9.82	**100.00**
Print Expected	51.40	174.49	206.95	162.31	27.73	79.13	**702.00**
Print Count	62	199	221	120	16	84	**702**
Print Percent	8.83	28.35	31.48	17.09	2.28	11.97	**100.00**
Overall Expected	76.00	258.00	306.00	240.00	41.00	117.00	**1038.00**
Overall Count	76	258	306	240	41	117	**1038**
Overall Percent	7.32	24.86	29.48	23.12	3.95	11.27	**100.00**

For readers of on-line scholarly journals who are using a modem to access the Internet, communications costs are a significant issue. Even readers using an internal network at their place of work may be liable indirectly for costs if their institution is using some form of chargeback arrangement. These costs are incurred each time the journal is accessed remotely and will probably have some time-based component if a modem is used. This is in contrast to print journals where access is free once the print publication is acquired and where there is no time-based charge for use. This question sought to assess respondents attitudes to this potential disadvantage.

For all of the ratings from Strongly Agree through to Neutral the print respondents were more likely to be concerned about this disadvantage. However, for the Disagree rating, the email respondents provided twice as many responses as the print respondents.

7.4 Conclusion

The surveys discussed in this chapter provided a clear picture of the general demographics for the two target populations as well as their access to relevant technology for delivery of hypermedia on-line journals (personal computer, a CD-drive (for CD-ROM based publications), a sound card or equivalent, a colour screen, and a direct network connection or modem). The surveys also quantified their frequency of use of particular forms of electronic publishing: subscribing to an email list, accessing an ftp server, accessing a gopher server, accessing the Web, using a CD-ROM, viewing electronic journals, viewing the e-journal *Psyche*, or publishing electronically themselves. Finally the surveys elicited their feelings about a series of possible advantages and disadvantages of electronic publishing.

The two survey sub-populations (the email and print survey) show significant differences on a number (but not all) of the measures. They are also drawn from the same scholarly discipline (psychology). If one considers the data on access to, and use of, information technology, one interpretation consistent with pattern of answers is that the email survey sub-population are early adopters of information technology for scholarly communication (relative to the print survey sub-population). It is therefore possible that they can be used to indicate early trends in use for the print survey sub-population.

8 Library Case Studies

8.1 Introduction

The previous chapter presented the results of the survey phase of the research. This phase quantified the responses of two comparable groups of scholars to questions about their demographics, access to technology, use of various electronic publishing technologies, and attitudes to the new scholarly publishing possibilities. Scholars, while important, are only one of the stakeholder groups within scholarly publishing and only one component of the publishing ecology (in their dual roles as producers and consumers of journal articles).

To investigate another component of the ecology, a series of in-depth case-studies examined libraries who were responsible for a significant and innovative electronic publishing project, rather than just acting as an access point for material provided by a publisher. The intention of this phase of the research was to get a sense of how five very different libraries were approaching the challenges of the new technologies and rethinking the role they could play in the system of scholarly communication.

This chapter describes the results of this research. Firstly, the overall research process is outlined, with an explanation of the process of project selection, data collection, and analysis. Next, each project (listed in alphabetical order) is described with particular emphasis on its origin and organisation, the products it is responsible for, the lessons learned by the project participants during its life to date, and finishing with an examination of future prospects.

8.2 Case study research

8.2.1 Overview

Case study research involves ongoing interactions between design, data collection and analysis [Yin, 1998]. It also requires the researcher to be able to deal with a range of potential sources of evidence: documentation, archival records, interviews, direct observations, participant observation, and physical artefacts. The case study “attempts to thoroughly assess a cluster of factors by focusing on a small number of cases” [Adams and Schvaneveldt, 1985]. Case studies can be either explanatory, descriptive or exploratory. Obviously, each of these have different orientations. What then are the processes in case study research and what decisions were made for this thesis (and why)?

8.2.2 Case study issues

According to [Yin, 1998], exploratory case studies are the preferred orientation:

> When the available literature or existing knowledge base is poor, offering no clues for conceptual frameworks or notably propositions (p. 236).

While there is a large literature on electronic journals and electronic publishing, there is little dealing with academic libraries as facilitators for electronic scholarly publishing (the focus of the case study research). In this case the exploratory approach seemed most appropriate.

Case studies have often been criticised on the grounds of lack of representativeness, difficulties in generalisability and potential for investigator bias [Adams and Schvaneveldt, 1985, p. 114]. However, [Rose, 1991] argues that such criticism is based on a degree of methodological confusion:

> In case study research … it is considered more appropriate to treat representativeness in terms of a qualitative logic for the selection of cases for study, rather than a quantitative logic of sampling from a population (p. 192).

Such a careful selection was the process followed for this phase of the research.

8.2.3 Designing the case study

Unit of analysis

The first decision was to define the unit of analysis - the level of the case. In classic case study research a single individual is the object of study (and thus the unit of analysis). More common nowadays is case study research that defines organisations or programs as the basic unit. Some of the candidate academic libraries were responsible for multiple electronic scholarly publishing projects. It was decided to select a specific project (usually the lead project) within a library as the basic unit. This is because the libraries are responsible for a wide range of activities and trying to cover all of these would dilute the focus on the electronic publishing projects.

Project selection

A case study may be about single or multiple cases. It is possible to generalise from single cases (in some analytic way) but multiple-case studies can strengthen or broaden such gen-

eralisations (similar to the advantages of multiple experiments). [Yin, 1998] distinguishes between *literal replication* (where the cases are designed to corroborate each other) and *theoretical replication* (where the cases are designed to cover different theoretical conditions). In the latter case, one might expect different results but for predictable reasons. Because of the exploratory mode of the research, it was not possible to determine before the data collection process the most appropriate theoretical base to use to guide project selection. The decision was therefore made to select projects that differed on a range of measures: organisational structure, geographical location (still important in this increasingly wired world), and type of published product. This provided the greatest coverage and best chance of identifying patterns of difference or similarity. The selection technique was based on first investigating candidate projects via a structured literature and Web search.

The research sites were selected to provide a sample of leading-edge projects (details of the projects are discussed below). The geographical range of the selection ended up encompassing both coasts of the United States, the United Kingdom and Europe. No Australian projects (of which there are a number) were included in this phase; the organisation funding this phase (the Victorian Association for Library Automation) required that projects studied be located outside Australia. Together the projects selected provide a good picture of the diversity in the field. Little relevant activity seems to be occurring outside the countries selected. Given the extremely fluid nature of the electronic publishing field, specific comments about the projects are only valid as of early December, 1997 (the end of the data collection process).

Quality control

[Yin, 1998] recommends that researchers continually judge the quality of their case study design. Four tests that are commonly used are to assess if the study has *construct validity*, *internal validity*, *external validity* and *reliability*. Yin argues that these tests should be applied throughout the case study process: during design, data collection, data analysis and reporting Following these recommendations will "increase the quality of your case study tremendously, and overcome traditional criticisms of the weakness of case study research" [Yin, 1998, p. 242]. Table 8–1 summarises 11 recommended tactics covering theses four tests and also indicates the ways in which the research design and conduct for this case study responded to these recommendations.

Table 8–1: Case study tactics and responses (Source: [Yin, 1998])

Tests	Case Study Tactic	Research Phase in which tactic occurs	Action taken in this research
Construct validity	Use multiple sources of evidence	Data collection	Use of interviews, documentary evidence and physical artefacts
	Establish chain of evidence	Data collection	Interview data both taped and transcribed in real time; multiple evidence sources entered into customised object-oriented database
	Have key informants review draft case study report	Composition	Two conference papers and one journal article based on case studies reviewed by key informants before publication
Internal validity	Do pattern matching	Data analysis	Patterns identified across cases
	Do explanation building	Data analysis	Some causal links identified
	Do time series analysis	Data analysis	Not performed in this research, but under consideration as part of follow-up work
	Do logic models	Data analysis	Not performed- requires time series data
External validity	Use rival theories within single cases	Research design	Not used because of exploratory nature of research and lack of existing theory
	Use replication logic in multiple-case studies	Research design	Multiple cases investigated using replication logic
Reliability	Use case study protocol	Data collection	Same data collection procedure followed for each case; consistent set of initial questions used in each interview
	Develop case study database	Data collection	Interview transcripts, other notes and links to online and physical artefacts entered into database

8.2.4 Data collection

Case study protocol

Once the candidates had been identified, a consistent protocol was followed. The most suitable personnel (typically the project director and if possible their manager in the library) were identified by using the project web pages. These potential informants were contacted via electronic mail which outlined the research and requested their cooperation. These initial contacts were also asked if anyone else involved in the project should also be contacted. Between April and September 1997 each site was visited and a semi-structured oral interview

took place with the informants. The interview questions are included in the Appendices to this thesis (see 11.3: Library Case-Study Questions on page 243). Each informant was asked each of these questions unless they indicated that they were not appropriate. Other lines of enquiry were also pursued in the interviews as seemed appropriate. Each interview was transcribed in real time onto a notebook computer and audio-taped, following recommendations in [Jones, 1991a]. Tapes were reviewed the same day to ensure that nothing had been missed and any follow-up questions were asked in a separate interview session on the next day. Some 20 hours of tapes in all were recorded.

Sources of evidence

The case study research relied on multiple sources of evidence. Documentary evidence was important as corroboration and augmentation of interview evidence. Available documentation included the online project documentation and any published articles or conference presentations either by or about the projects. The interview process has already been discussed. Physical artifacts included both print and online products of the projects.

Case study database

An object oriented database package (WebArranger from CESoftware) was used to create a customised case study database. This was used to store the full text of interview transcripts, URLs for project homepages and related documents, contact addresses, and the dates and times of interviews and follow-up contacts. This information was all fully searchable and viewable in a number of ways.

8.2.5 Data analysis

In one sense, data analysis occurs throughout the case study research process as the researcher continually interacts with the collected data and their informants. One of the advantages of case study research is its flexibility, allowing the researcher to pursue new lines of enquiry that are suggested by an informant's evidence or a piece of documentation. In another sense, the major task of analysis occurs once the bulk of the data has been collected and can be inspected, categorised and manipulated. The main analysis tools used were pattern matching and explanation building.

The pattern matching phase involved identifying particular themes in the respondents' responses to questions. It became clear that there were significant clusters of responses around the questions of the nature of publishing and the role of libraries. The analyses of these re-

sponses sought to compare and contrast these responses. The evidence is only presented later in the thesis. The explanation building phase revolved around the respondents' answers to the questions about the critical success factors for their projects to date. The evidence is presented in the *Lessons Learned* section of each case study. The results of both analyses are discussed later in this thesis (see 9: Interpretation of findings on page 200).

8.3 Highwire Press

8.3.1 Overview

Highwire Press is an initiative of Stanford University Libraries/Academic Information Resources. It was selected because it is implementing leading-edge Web-based e-journal technologies and because it is commercially successful (i.e. not just a pilot). Stanford also has a number of interesting projects including the Digital Libraries Initiative, and I was interested in possible synergies. Highwire Press is available on-line at <http://highwire.stanford.edu>.

The stated mission of Highwire [Highwire Press, 1997] is to:

- Foster research and instruction by providing a more direct linkage between the writers and readers of scholarly materials.
- Use innovative network tools for capture, publishing, retrieval, reading and presentation.
- Affect the economics of provision of scholarly information to researchers, especially science, technology and medical (STM) research information.
- Ensure that the nascent marketplace for electronic communication among scholars does not develop along the semi-monopolistic lines of current STM publishing.
- Build new technological, economic and programmatic partnerships with others investigating related problems.

The main informant at Highwire Press was Vicky Reich, Assistant Director and Digital Librarian at Stanford University's Green Library. The author also attended a workshop presented by Mike Keller, Director of Highwire Press when he was visiting Australia in late 1996.

8.3.2 Origins and organisation

Librarians have complained for a long time about the rising cost of STM materials and the need to regain some control of the scholarly literature. Highwire grew out of that frustration.

When Mike Keller came to Stanford four years ago, he was active in some committees with Robert Simoni, Professor of Biological Sciences and 'Godfather' of the *Journal of Biological Chemistry (JBC) Online*. The decision to create *JBC Online* came out of their shared interests and discussions. *JBC* (the print version) and the American Society for Biochemistry and Molecular Biology shared equally in the development costs of bringing it up. Senior university management knew what Highwire were doing and were (and are) actively supportive. At Stanford entrepreneurial ventures are encouraged, and the 'Silicon Valley startup' nature of Highwire fitted well into this. The origins of Highwire were summarised by Vicky Reich as 'Right person, right time, right technology'. It should be emphasised that Highwire operated from the start on very short development cycles and has rewritten its systems three times in the last two and a half years to stay current.

8.3.3 Financial sustainability

Highwire currently operates as a separate cost centre within the library, with the Publisher of Highwire, Mike Keller, also being the University Librarian, and Director of Academic Information Resources. At the time of writing the Highwire Press team listed on their homepage consists of 26 people (including support staff). The more senior Highwire staff also fulfil other positions within Stanford University. The project is currently commercially sustainable, in line with Stanford's policy of extensive chargeback for services. It is not however seen as a way for the university or library to make a profit. Rather, it is viewed as a cost-recovery exercise with both tangible and intangible benefits for the university. Because of its organisational location within the library and physical location in Silicon Valley, the infrastructure was already in place to ease the startup process.

8.3.4 Products

Implicit in the Highwire Press mission statement is their intention to provide a model for re-engineering scholarly communication. To this end, they are working in partnership with scholarly societies to bring existing print journals online. The first of these, *The Journal of Biological Chemistry* has now been joined by 24 largely biomedical publications which are available in both print and electronic form. An additional five or six *dozen* more titles will be available online by the end of 1998 (please consult their homepage for details). For many of these titles there was a personal or professional connection between the journal personnel, which is what drove the development of successive journals within particular discipline areas.

These journals are at the leading edge of Web-based journal publishing and are progressively adding a number of additional value features that are only possible in an online environment. These include:

- "direct searching by author, title keywords or text words, both within journals and across journals
- display in PDF (best for printing) or HTML (best for navigation)
- automatic creation of hyperlinks to MEDLINE citations provided by the National Library of Medicines PubMed service
- links from Genbank accession numbers to full Genbank records
- bidirectional links between citing articles and cited references (where available)" [Newman, 1997]
- "'toll-free' links between the references from one journal article to the full text of the cited article" [Reich, 1997].

8.3.5 Lessons learned

The critical factors for Highwire's continuing success were stated to be a mix of the academic setting, physical setting, technology and support of the university administration. The academic setting within Stanford provides rich connections to subject domain experts and academic staff. The involvement of library staff at all levels ensures a deep understanding of the requirements of all the participants. As librarians, it is their job to deliver information directly to university researchers and understand their needs. Highwire also helps the publishers deliver intellectual property through librarians to researchers. Finally, as librarians themselves, they understand the special needs and constraints that librarians have; they are their own clients. The physical location of Stanford within Silicon Valley facilitates a range of cooperative projects with leading-edge computer companies, as well as 'grapevine' access to the latest technology news. The right technology mix at Stanford includes a very talented staff and an outstanding networking infrastructure, essential for ensuring adequate client access. By the same token, projects like Highwire based on the latest Web technologies need skills that are in high demand throughout Silicon Valley at present. Compared to a computer startup company, a university can't offer the promise of stock options and becoming rich. This can be a problem in recruiting staff.

8.3.6 Future prospects

Vicky Reich hopes that over the next five years the technology will continue to evolve but will also trickle down. Highwire Press will still be doing the leading edge sophisticated online publishing, but making university published literature and technical reports available online will be seen as routine. Of course, as more and more material becomes available, the challenge is to provide enough overall searching tools to access a coherent information space. Another likely outcome is the transition of existing print journals to an electronic-only existence within the next five years, thus providing a significant saving on distribution costs for the societies that provide these journals to their members. The example of The Journal of Biological Chemistry (JBC) (see 6.3.1: Journal of Biological Chemistry on page 117) which is published by Highwire Press is instructive. The print-only subscription price for this journal to institutions is $1600/year. The electronic-only price is $1100/year. In other words, the printing and distribution component of the journal cost is 32%, exactly where publishers say it should lie with respect to the 70/30 debate.

8.4 Internet Library of Early Journals (ILEJ)

8.4.1 Overview

The Internet Library of Early Journals (ILEJ) is a joint project between the Universities of Birmingham, Leeds, Manchester and Oxford. It aims to digitise a critical mass (defined as at least 20 consecutive years) of three eighteenth century journals (*Gentleman's Magazine*, *The Annual Register*, and *Philosophical Transactions of the Royal Society*) and three nineteenth century journals (*Notes and Queries*, *The Builder*, and *Blackwood's Edinburgh Magazine*). While not extremely rare, there are only perhaps 20-25 sets of each journal extant. The digitising will therefore need to be done on a non-destructive basis. ILEJ was selected as an example of a digitisation project working with non-scientific and older serial material. The ILEJ homepage is available online at `<http://www.bodley.ox.ac.uk/ilej/>`.

The project aims to explore the issues associated with making this sort of material available as well as providing access to it. The variables they are particularly interested in are image creation, indexing techniques and Web access to page images. A number of their working decisions have been made with an eye to reducing the cost of doing this sort of work as far as possible. The intention behind providing access to the material in digital form is to facilitate access by researchers (through desktop access and search mechanisms) and to reduce

the need for physical handling of the originals. The project aims to mount 120,000 page images in all.

The main informant for ILEJ was Peter Leggate, Keeper of Science Books, Radcliffe Science Library, Oxford University. Discussions also took place with a number of the technical staff at the Bodleian Library.

8.4.2 Origins and organisation

The project began as a result of a successful bid for funding from the eLib programme, set up after the Libraries Review by the UK Higher Education Funding Councils, chaired by Professor Sir Brian Follett in 1993 (sometimes called the Follett Report). The Bodleian library already had an interest in digitisation projects. They decided to focus on materials from the 18th and 19th century in part because of copyright problems with newer material, and to choose materials that were heavily used. The original bid from the Bodleian included journals, newspapers and slides. The eLib programme identified a number of similar digitisation projects and suggested the eventual consortium with Birmingham, Leeds and Manchester. The project effectively commenced in late 1995.

Peter Leggate is the joint project leader with Hugh Wellesley-Smith, Deputy Librarian, Edward Boyle Library, University of Leeds. The management of the project is being coordinated from Oxford. Scanning the journals from microfilm and keyboarding the index entries also takes place at Oxford. Scanning from hard copy is being undertaken at Birmingham and Manchester. Servers to mount the images are located at Oxford and Leeds.

8.4.3 Financial sustainability

The project is not currently commercially sustainable, and may in fact never be given the nature of the material being digitised. The project team are actively considering the best way to proceed once the initial funding has been allocated. The ILEJ project has been able to extend its operations outside the originally funded period because of early delays caused by waiting for equipment. They have also been able to save some money through higher than expected throughput. At present, the original grant of £338,000 is anticipated to last until August 1, 1998.

Some possible extension options for funding are:

- to use the skills and equipment acquired for this project to provide a service for others

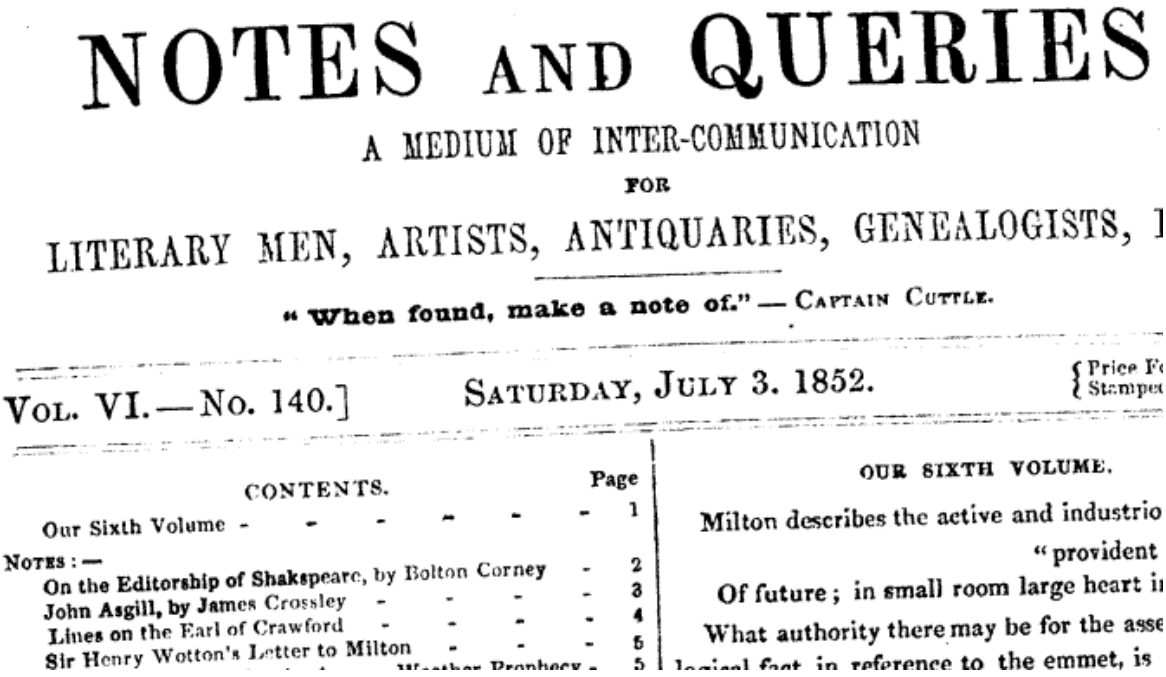
on a fee-for-service basis

- seek funding to digitise a larger collection of similar journals based on the holdings of the libraries involved.

8.4.4 Products

At the time of writing (mid 1998) fifteen years of *Notes and Queries* and *Gentleman's Magazine* have been made available. The selected journal consists of a wide mixture of intermingled article types laid out on the page in a somewhat unstructured manner. This means that providing full-page images is the simplest way to present the material. A segment of a page is shown in figure 8–1. Note the skewing of the original scan (probably to microfilm) and poor definition of the original typeface.

Figure 8–1: Notes and Queries, Vol. 6, No. 140, p. 1. Source: ILEJ Web pages.



NOTES AND QUERIES

A MEDIUM OF INTER-COMMUNICATION

FOR

LITERARY MEN, ARTISTS, ANTIQUARIES, GENEALOGISTS, I

"When found, make a note of." — CAPTAIN CUTTLE.

VOL. VI. — No. 140.] SATURDAY, JULY 3. 1852. {Price Fo Stampe

CONTENTS. Page

Our Sixth Volume - - - - - - 1

NOTES: —

On the Editorship of Shakspeare, by Bolton Corney - 2

John Asgill, by James Crossley - - - - 3

Lines on the Earl of Crawford - - - - 4

Sir Henry Wotton's Letter to Milton - - - 5

OUR SIXTH VOLUME.

Milton describes the active and industrio

"provident

Of future; in small room large heart i

What authority there may be for the asse

Each full page image is between 100 and 200K in size. Pages can be browsed by volume/number/page or searched directly. The input to the search database is OCR text from the page scans. This OCR is done using off the shelf software (designed for modern typefaces and layout) and without manual intervention (because of the prohibitive cost of correcting errors). The result is quite 'dirty' OCR, which would normally present problems in searching. As a partial solution, the Excalibur Technologies EFS search engine is being used for full text fuzzy matching on the dirty OCR text.

8.4.5 Lessons learned

Critical to the success of ILEJ has been the collaboration between sites which provides a wider pool of expertise and technology to draw on. Of course, this has also meant an increased coordination load, taking the time of staff who were paid to do other jobs and who were too busy anyway. This factor was exacerbated by the decision (taken after careful analysis) not to appoint a project manager and to use the funds saved for technical staff instead.

Another difficulty encountered was delays in the availability of necessary scanning technology. This meant a consequential lag in appointing scanning staff and a lack of content in the early stages of the project. Peter Leggate indicated that if the project had ben able to acquire a stock of page images earlier, then they could have modified the servers and programmed functions in a more informed way.

Peter Leggate also made the point that ILEJ is strongly committed to offering a useful service to the scholarly community, not just doing a digitisation project. While this has been a fruitful model, it has also involved a necessarily greater spread of activities.

8.4.6 Future prospects

The issues faced by ILEJ are those that will have to be faced long-term by the new proposed hybrid libraries. Such libraries will have to deal with a spectrum of resources. Very rare and old materials will probably be digitised anyway. Current materials will increasingly become available in electronic form. ILEJ is dealing with the issue of what to do with the vast bulk of material in the middle, although it has been able to avoid the vexed issue of copyright. The most likely solution is to digitise high demand materials first and then do digitisation on demand for the remainder (although funding this on-demand work may be difficult). With respect to material still covered by copyright, Peter Leggate's view is that libraries need to convince journal publishers that there is little money to be made from their older material. One suggestion he made is that after 20 years, perhaps materials could be digitised on demand with a minimal licence cost. Getting the publishers to agree to this may be problematic, of course.

8.5 Project Educate

8.5.1 Overview

Project EDUCATE (End-User Courses in information Access through communication TEchnology) is a joint initiative of Limerick University (Ireland), the École Nationale des Ponts et Chaussées (France), the University of Barcelona (Spain), Chalmers University of Technology (Sweden), the Imperial College of Science Technology and Medicine (United Kingdom) and Plymouth University (United Kingdom). The project homepage is online at <http://educate.lib.chalmers.se>.

The overall aim of the project is to help students, research workers and practitioners to develop their information literacy. EDUCATE was selected because it was publishing online teaching support materials (rather than journals) and because it provided a Nordic/European perspective.

The informants for Project EDUCATE were:

- Jan Rohlin, Library Director, Chalmers University of Technology Library
- Nancy Fjällbrant, project ‘inspirer’, Chalmers
- John Fjällbrant, technical officer, Chalmers.

8.5.2 Origins and organisation

The initial idea came from Nancy Fjällbrant. DG XIII from the European Union had given some talks about possible projects, and Jan Rohlin thought it would be good to have some international projects, despite the EU jargon. Nancy Fjällbrant had already identified a range of weaknesses in user-education in parts of Europe and elsewhere. EDUCATE was originally targeted towards librarians, to replace paper-based materials and to assist them with user-education. The EDUCATE team worked on the idea for one year, developing their own software client, but when the Web came along the decision to move was an easy one. The hyperlinking in the Web environment allowed them to move further and further away from a strictly linear product and thereby to make it more useful.

The decision to seek EU funding arose from a need to raise the amount of money required to do things properly (something very hard to do on a local scale from a small country like Sweden). A successful bid for funding was made under the European Union Telematics for Libraries programme -Third Framework. The original funding was for a three year period from

January 1994 through February 1997. Chalmers University and the Library (as well as some other funding bodies) contributed an equivalent amount of money. The project was in part driven by a desire to do it anyway - the funding just allowed this to happen more effectively and sooner.

Nancy Fjällbrant was able to draw on her International Association of Technology University Libraries (IATUL) contacts to get assistance with the project. Under the grant proposal, Limerick University was to be the overall coordinator. In practice, Chalmers University of Technology Library provides the day to day technical and administrative management. Imperial College are doing some demonstration versions and courseware. Limerick did some of the interface design and Web development. Translation (and adaptation - using and linking to different resources) is being done into French at Ponts et Chaussées and into Spanish at Barcelona.

8.5.3 Financial sustainability

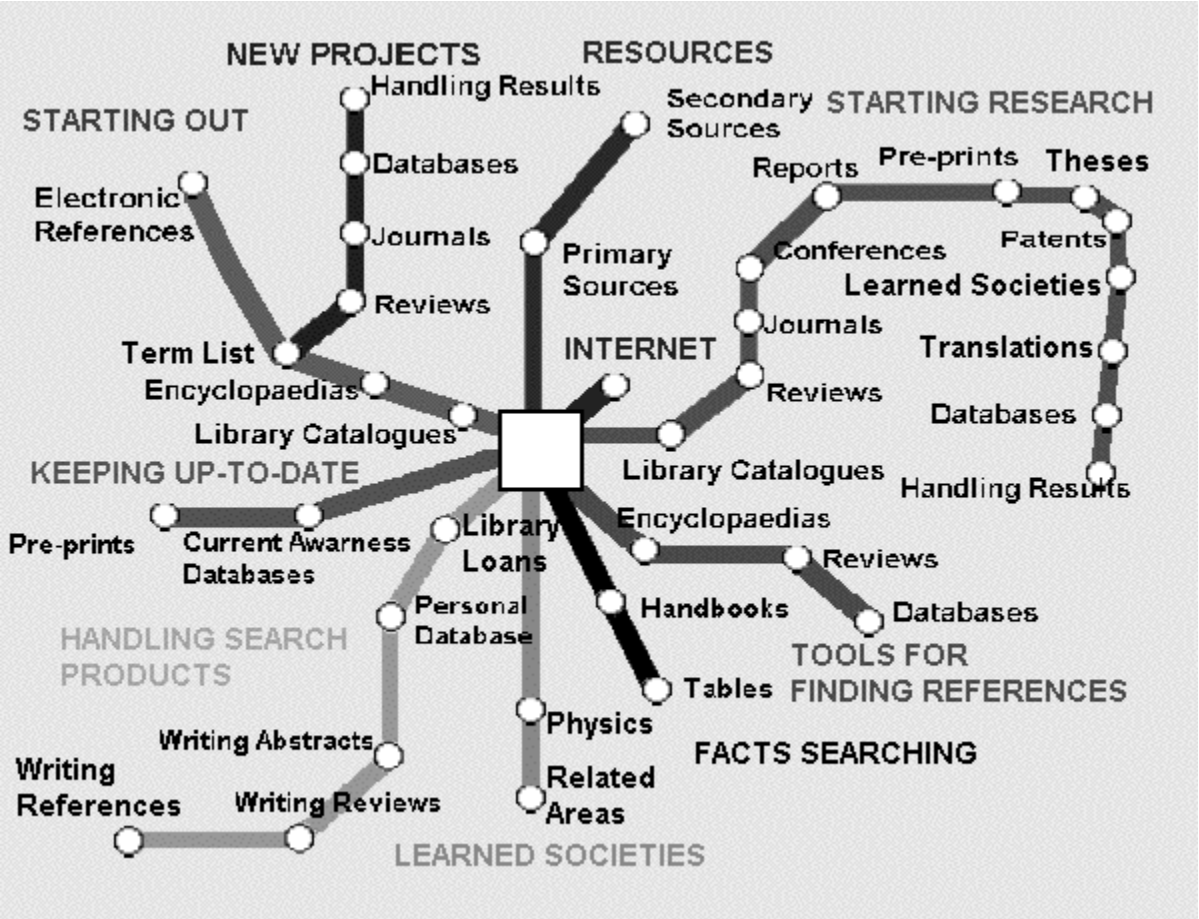
The intention was that EDUCATE should be commercially sustainable immediately. There is a need for a revenue stream for two reasons: maintenance of existing materials and development of new ones. The experience of other electronic publishing projects has been that initial funding is easier to get than maintenance funding. Jan Rohlin wanted to avoid this and has determined that new modules may only be developed with new money. The revenue stream for maintenance of existing materials comes from the licence fees. These have been set as low as possible consistent with getting enough funds. The total from these fees is not yet at the break even point but the trend is looking promising.

8.5.4 Products

The main product of Project Educate has been a series of Web-based self-paced user education courses called Into Info. These provide training in the selection and use of information tools and resources in particular subject areas (chemistry, physics, electrical and electronic engineering, and energy to date). All the Into Info modules are based on a multi-level hierarchical structure with rich internal and external hyperlinking. The first level of the hierarchy offers:

- *Pathfinders*, or routes to follow in different situations;
- *Information Sources* - direct points of access for experienced users;
- *Course in Information Searching* - with goals and objectives, texts, demonstrations, exercises and self-assessment questions - particularly useful for self-study and dis-

Figure 8–2: Pathfinders Navigation map



tance learning;

- *Texts* - available in HTML for downloading;
- *Introduction to the Internet* with links to books and courses about the Internet;
- *IT+++* which provides access to news and general information;
- *A-Z Index* to the main items.

An example Pathfinders map is shown in figure 8–2.

As well as providing users with skills in using the information resources, the Into Info modules act like an annotated bibliography of the highest-quality resources in each discipline area. In this way, users of the modules learn simultaneously how to access resources and which are the best to access.

These courses are both used within the member universities and site-licensed to other universities for use in their own user-education programs.

8.5.5 Lessons learned

The critical success factors for EDUCATE were identified by its director as first of all having a good idea, having the Web (and the idea of the universal browser) come along at the right time, and the importance of checking with the user community throughout the development process. EDUCATE was able to build on the reputation of Nancy Fjällbrant in the user education community and the brand name of Chalmers. The challenges (and benefits) of inter-institution and international cooperation were an issue for EDUCATE and would probably be for any similar project.

8.5.6 Future prospects

Jan Rohlin believes that the project will be of interest for another 3 years, but will have been transformed and reworked completely by that time. At present there are no obvious competitors for provision of this sort of material, but this may not be the case into the future. Modules for different content areas are a possibility with additional funding. The Dedicate project (funding being applied for at present) will be focused on distance education into Eastern Europe and the Baltic states, based on the Educate model and with a focus on the area of information retrieval).

8.6 Project Muse

8.6.1 Overview

Project Muse is an initiative of the Johns Hopkins University (JHU) Press and the Milton S. Eisenhower Library at JHU. It provides worldwide networked access to the full-text of over 40 of the Press's electronic journals. Project Muse was chosen as it is one of the very first Web-based electronic journal projects, and has a predominantly humanities coverage in contrast to Highwire. The Project Muse homepage is available online at <http://muse.jhu.edu/muse.html>.

The informants at Project Muse were:

- Susan Lewis, now at Jossey-Bass Publishing, but who was Online Projects Director for Johns Hopkins Press when Project MUSE was starting;
- Todd Kelley, now Associate Provost & Librarian of the College, St. Mary's College of Maryland, but who was Librarian for Electronic Information Initiatives at the Milton S. Eisenhower Library, Johns Hopkins University when Project Muse was start-

ing;

- James Neal, Director, Milton S. Eisenhower Library, Johns Hopkins University;
- Dawn Hale, Head of Cataloguing, Milton S. Eisenhower Library, Johns Hopkins University.

8.6.2 Origins and organisation

Project Muse arose originally from informal discussions between Susan Lewis and Todd Kelley. They both identified the strengths of the journals published by the press and felt that moving into electronic publication would have significant advantages. The initial development took place in a very bottom-up way, driven by the commitment of these two individuals. The original aims were to:

- reduce the cost of publication to users;
- forge a different approach to intellectual property and copyright
- test new economic models within the scholarly communication environment;
- test new ways of using network technologies.

Both the library and the press recognised that significant seed-funding was required. They made the decision to seek grants from the National Endowment for the Humanities (NEH) and the Andrew Mellon Foundation, and to use the Web (then very new) as the technology platform. These grant applications were both successful, and development commenced.

The library helped enormously with setting up the project and with raising money through fundraising. They developed the university wide relationships between the various players. They also had the technical expertise in online information (but none in publishing). This meant that the logical relationship was for Todd Kelley and Susan Lewis to partner together. Susan Lewis got the files from the press' production system and had them translated into Postscript. The library provided the server, organised the information, and embedded metadata into the articles to assist with searching. In effect, the press provided the content, and the library added value through cataloguing, searchability and dissemination.

8.6.3 Financial sustainability

The original grants from the NEH and the Mellon Foundation finished at the end of 1997. The project is still a joint initiative between the Library and the Press and is now commercially sustainable. The project is starting to provide access to journals that do not come from the JHU Press. It is also now offering individual subscriptions, rather than just site licenses.

8.6.4 Products

Project Muse provides access to the full text of over 40 of the Press's scholarly journals (but not complete runs of each journal) in the humanities, the social sciences, and mathematics. About 3 to 4 outside journals will be added each year from now on. Most of these journals are published in print as well but some are electronic only. No matter what category the journal belongs to, all titles provide:

- "Hypertext links in the tables of contents, endnotes, author biographies, and illustrations
- Text designed for on-screen reading and easy printing
- Boolean searches of either the full text or author/title/keywords from the tables of contents
- Subject indexing of each article using Library of Congress subject headings
- Illustrations that are both larger than those in the print version and often in color
- Subject, title, and author indexes
- Option to create online reference 'shelves' or electronic syllabi
- Ability to print individual articles 'on demand'
- Earlier availability than print edition" [Muse, 1997]

Project Muse prefers to provide access through consortia of universities. This reduces the administrative overheads for the project and enables them to provide discounts to the consortium members. Some of the Muse licences are for single universities, and from the end of 1997 they began supporting individual user subscriptions.

8.6.5 Lessons learned

As one of the earliest projects, Muse is now a mature system whose success can be regarded as established. Critical in getting to this point was the support of senior management in the library (particularly important for a project like Muse that started bottom-up), and the commitment of the early project staff in making it work and dealing with obstacles aggressively. Susan Lewis commented that in the early days they probably spent too much on getting the product right and not enough on marketing and promotion. Other than that, both Susan Lewis and Todd Kelley felt that they got things about right.

8.6.6 Future prospects

When Todd Kelley started out with Project Muse, he thought there were 3 stages (each lasting about 3 years) that they would go through:

i. Startup period

During this phase, on the editorial side (the first copy content costs) he believed that print would subsidise electronic. It has essentially worked out that way. The money coming in from Muse is not going to royalties or editorial costs.

ii. Mature system

This would be characterised by up to 500 sites (or thereabouts) licensing Muse journals. As the project moves into this stage Muse has to pay its own way because grant money has run out and one can start to pay royalties for electronic access on top of print. This is essentially where the project is now. He predicted that as libraries start to drop print, there are costs that the press has which are higher (due to smaller print runs). As these costs change, they should disassociate costs of print and electronic to better reflect their individual costs. This then drives the transition of libraries from print to electronic. Round about year 4 of the project he expects print and electronic subscription rates to cross over. One can then add more and more subscriptions to the electronic product at near-zero marginal costs, which is not the case for print.

iii. All ties between print and electronic are cut

In this phase, print sales start to slip. This puts the Muse model and approach out front rather than tagging onto print. He hopes by end of 1998 that the electronic version will be the journal of record. Todd Kelley has been recommending that the press start to plan for print on demand for those who will still want print. He has also been recommending a document delivery service for those who want it. When asked about the long term future of print, he indicated that he thinks print will be gone altogether by 2003.

8.7 Scholarly Communications Project

8.7.1 Overview

This project, based in the library at Virginia Polytechnic Institute and State University (Virginia Tech), has been working on a range of publishing activities since 1989 [McMillan, 1995]. The project homepage is available online at <http://scholar.lib.vt.edu/>.

The Scholarly Communications Project was selected because of the diversity of its activities, and because it operates in a service rather than cost-recovery environment. The project

> "assists primarily Virginia Tech faculty who are editors of professional journals when they want to also make their publications available to their colleagues in distributed academic communities via the Internet. It assists traditional academic publishers adapt their publications to the Internet and access by the worldwide academic communities. It has also works with a variety of units within the university to extend access to their clients locally, regionally, nationally, and internationally." [Scholarly Communications Project, 1997]

The main informant at the Scholarly Communications Project (SCP) was Gail McMillan, Project Director since 1994. James Powell, Technical Director, and Professor Ed Fox, who has been active in the Electronic Theses and Dissertations initiative were also interviewed.

8.7.2 Origins and organisation

The project was originally established in the autumn of 1989 by the then Vice President for Information Systems. It was decided early on that the most appropriate organisational base was the library. Its current director, Gail McMillan, is also in charge of Special Collections. Over time, the project has added more and more online publishing activities, and is seen by the university community as the experts in this domain.

8.7.3 Financial sustainability

The funding for most project activities is provided as part of the library budget. Some funding for specialist activities comes from outside agencies. Gail McMillan sees her role as a service-oriented librarian to say yes to any reasonable request from within the university and then work out how to resource it. The university administration is very supportive of the ac-

tivities of the project and has allocated significant resources to it in the forward planning process.

8.7.4 Products

The digital products produced through the project now include:

- 12 e-journals (a number of these are electronic only);
- Virginia Tech newsletters and magazines
- some digital image collections (images from the Special Collections department, American Civil War Resources, and a History of Architecture Catalogue)
- the Electronic Theses and Dissertations initiative (providing access to PDF versions of graduate theses which will no longer be accepted by Virginia Tech in paper form)
- regional (the *Roanoke Times* and the *Virginian Pilot*) and international news
- electronic reserve system (also using PDF).

This fairly eclectic portfolio reflects the willingness of the project to respond to requests for assistance from their user community.

8.7.5 Lessons learned

The critical success factors for the SCP include reliable technology and excellent technical support deriving from the strong cooperation between the project and library automation people, realistic expectations from the journal editors and a degree of personal freedom for the project director to pursue new initiatives. The project is free to experiment, make mistakes and risk things. One theme that kept emerging in the discussions was the importance of a service orientation and a commitment to make things happen.

8.7.6 Future prospects

The SCP is one of the things that Virginia Tech likes to show off to new visitors. This commitment is reflected in the allocation of a whole floor of a new building to the project. Gail McMillan has been involved in the design of this floor and has designed space for project managers, programmers, student workers and a conference room. SCP is not seen as a press that goes off-campus. It is rather seen as a facilitator and outlet for content created by faculty and students. They are also thinking about setting up ad-hoc review boards to referee particular papers that don't become a journal but which still get the VT seal of approval. Overall, the future looks very bright.

8.8 Conclusion

This chapter has reviewed five innovative projects that are redefining themselves and their activities in the light of the transforming potentials of the new technologies. These libraries illustrate the possibilities that are inherent in these technologies and point forward to the potential for a significantly changed niche for libraries in the future. The next chapter will consider the results of both the library case-studies and the survey work and interpret these results in the light of the theoretical perspectives presented already (see 2: Theoretical Perspectives on page 22).

V Interpretations & Conclusions

“I don’t mind your thinking slowly, Herr Doktor; I mind your publishing faster than you think.” Wolfgang Pauli

9 Interpretation of findings

9.1 Introduction

Section II of this thesis reviewed the development of the system of print scholarly journals to date and discussed the trends in and capabilities of computer and communication technologies currently available. Section III examined some of the potentials and pressures for transformation of print scholarly journals as well as looking at some developing responses and example e-journals. Section IV reviewed the results of the surveys and library case study work and discussed them (largely in isolation from one another or a body of theory).

This chapter will take all of the information presented so far and interpret it in the light of the research questions (see 1.4: Research questions on page 4). The aim is to see how the various stories told by the literature in the field combine with the data collected for this thesis and the theoretical perspectives chosen to inform the research questions. In considering the published literature, it is important to emphasis that this is a rapidly changing field. The patterns of use and the available technologies have altered considerably over the last five years and the older literature reflects the realities of a very different technology and publishing environment.

This chapter first considers the form and function of the artefacts of scholarly communication (journals and the articles they contain). It then looks at the roles and functions of stakeholders. In each case, the discussion of the findings looks first at the insights provided by the theoretical perspectives adopted. Next the discussion considers the relevant literature, particularly that with a data collection component. These studies have already been introduced and their data collection methods outlined; see 1.8: Related work on page 11). Finally the discussion ties in the research carried out for this thesis: the survey work and the library case studies.

9.2 Transformations in the form of journals

This section deals with the research question:

> *What transformations (if any) will occur in the form of the scholarly journal as it moves into an online hypermedia environment?*

Form is defined as the appearance, organisation, delivery mechanism and type of content associated with the scholarly journal. Some of this form is tied to the delivery technology and hence may or must change if the technology changes.

9.2.1 Insights from theoretical perspectives

Ecology of communicative transactions

Kaufer and Carley's ecology of communicative transactions consists of agents exchanging communicative transactions within an evolving open systems ecology. This idea is an extraordinarily powerful one when discussing a rapidly changing field such as electronic scholarly publishing. Kaufer and Carley's analysis serves as a reminder not to assume that the obvious factors are driving the process, and to be aware of the complex inter-dependencies between technologies, stakeholders and the scholarly environment.

The existing print journal is an example of what Kaufer and Carley call a 'book agent', by which they mean anything that is both printed and mass-circulated. Associated with any form of written communication in their analysis are the concepts of *fixity*, *reach* and *distance*. Distance means the separation between writer and reader caused by any combination of space, time and culture. Communications technologies can extend an author's reach through *asynchronicity*, *durability* and *multiplicity*. Of these concepts, *fixity* and *durability* relate to the form of the scholarly journal. *Reach, asynchronicity*, and *multiplicity* relate to the functions of the scholarly journal (see 9.3: Transformations in the function of journals on page 216). How do these properties inform possible changes to the scholarly article and journal?

Fixity is the extent to which a communication technology enables the communication to be retransmitted without change. Print, of course, is an example of a medium that is inherently highly fixed. Electronic text is much less highly fixed (unless stored on a read-only device). It is quite easy to alter an electronic text accessed across a network, either on the server at source, or in transit between the reader and the client. The user may not be aware that any change had taken place or that what they are reading is not what was originally written by the author. This is particularly a problem if the electronic text is changing frequently without this being marked in some obvious way. In the world of print, articles are defined packages of content and books have editions to indicate significant change. Online publications may change on a daily basis either without any indicator or at best a version number. Some online journals (e.g. *Public Access Computer Systems Review*) now allow the author the option of updating a published article. In this environment, how can a reader know which version has

been cited? Clearly, the current system of scholarly communication requires the fixity principle to be maintained in an online form.

Durability is defined in terms of how long the content of a communication is available for interaction [Kaufer and Carley, 1994, p. 34] and is a property of the medium (i.e. of the form) of the communicative artefact. Durable texts diffuse more widely and for longer and thus increase the author's reach. Durability is a problem for the transmission of texts across time as well as distance in the sense that it may be difficult to access an older electronic text because of technological change. Either the text may have to be migrated to a newer technology (violating the fixity principle by altering it) or it may be unable to access it (violating the durability principle by making it impossible to transmit it onwards). It is essential for electronic texts to be durable so that they can be accessed and cited into the foreseeable future.

In discussing the development of print, Kaufer and Carley point out that new media "coexist with, rather than replace, established media" [Kaufer and Carley, 1993, p. 6]. Rather than pitting print against the previous oral and written cultures they argue a more useful question is "what additional possibilities became available to speakers and writers once they could also rely on print?" [Kaufer and Carley, 1993, p. 6]. One might well extend this to ask what additional possibilities become available once writers can rely on hypermedia? These are ideas that some of the leading edge e-journals are just starting to explore.

On the subject of electronic print, they suggest a profitable research area might be to analyse how things like electronic mail and bulletin boards (and by extension electronic journals) allow the "diffusion of new ideas to the right people, over print alone" [Kaufer and Carley, 1993, p. 393]. Targeted mailing lists exist already, but no one seems to be suggesting targeted e-journals. However, it is possible to list requests for updates on topics of interest with a number of online services. The Amazon.Com online bookstore allows a user to list keywords and to be emailed when new books with these keywords become available. *The Journal of Biological Chemistry Online* allows a reader to indicate they wish to receive email when a specified article of interest is cited. It is quite possible that in the near future a scholar could have a series of alerting requests with a range of services providing a constant stream of messages about new discoveries and references to other scholar's (or their own) work. This is simply a logical extension of the old Strategic Dissemination of Information (SDI) services that libraries have been providing in conjunction with online databases for decades. The difference now is that it can be user initiated and operate on the full text of journals, not just databases of journal surrogates.

In their analysis of academe and the role of print journals Kaufer and Carley argue that in diffusing new ideas journals are simultaneously faster than book publication or face-to-face interaction (due to their frequency of issue and increased reach respectively), and slower than newspapers (due to the gatekeeper function of peer review). The consensus according to Kaufer and Carley is that many scientists regard the speed of journals as too slow, particularly in very fast-moving fields. This is clearly an area where e-journals can have a significant impact. E-journals can diffuse new ideas more quickly because of a range of factors:

- faster refereeing processes through use of email
- faster production processes through electronic submission
- faster appearance because there is no need to wait until a print issue is full before publishing
- faster delivery because there is no need to physically move anything to the reader

Kaufer and Carley argue for a non-deterministic view of the effects of new technologies, taking account of other sociocultural variables as well. They do, however, admit that the benefits of hindsight are much greater when looking at print than when considering the newer technologies. They make the explicit point that "researchers interested in (the newer technologies of communication) might learn something by analogy with our treatment of print" [Kaufer and Carley, 1993, p. 17]. This is a theme echoed by a number of other researchers in the field [Fjällbrant, 1997], [Rowland et al., 1995].

One of the things that such a researcher might learn is not to assume that a new communications technology implies a qualitative difference. Kaufer and Carley argue strongly for first considering the new on a common quantitative continuum with the old:

> To ignore historical continuities across technologies is to reveal more than a historical ignorance of what came before. It is to remain primarily ignorant of the technologies we now take for granted. [Kaufer and Carley, 1993, p. 18].

These insights imply both that changes in journal form will not solely be driven by technology and that the history of the scholarly journal may well have much to say about its future.

Punctuated equilibrium and speciation

According to the punctuated equilibrium model, sudden change in species is driven by environmental change. More specifically, local environmental change on a sub-population of an organism will preadapt this sub-population for the expansion of the changed environment

(Kaufer and Carley's analysis has modelled the process of communication as an ecology. Presumably, scholarly communication is a part of the overall communication ecology. Under this analysis, an individual journal title is analogous to a species, with changes in particular articles being the equivalent of changes at the sub-organism level.What then is the environmental change within this scholarly communication ecology that might drive speciation?

One excellent candidate is changes in the scholarly communication space caused by new technologies. The most important of these technologies for scholars since the middle of this century have been the communication and computing technologies. These have provided the potential for transformations of both professional practices and intermediary processes. The stakeholders who have been affected are publishers, scholars (as both consumers and producers of content), scholarly societies (as representatives of scholars and as publishing intermediaries in their own right) and librarians.

Most recently, the rise of the Internet and its increased possibilities for electronic publishing have provided a new ecological niche that is being colonized by all sorts of communications products, scholarly journals included. Is it possible that the move to the new electronic environment might involve the development of new species of scholarly communication artifacts?

The punctuated equilibrium model suggests that the equivalent of the sub-population adapting might be those journals now making the move into an online existence (with those who are making a break with paper altogether, rather than retaining parallel publishing, being best adapted). Within a relatively short period of time (far shorter in the world of publishing than in geological time) the remaining journals will need to either adapt and move to online or die out. This will appear in the equivalent of the fossil record (library serials holdings) as a quite quick transition from print only to electronic only. The consensus among discussions after electronic publishing conferences is that the transition should be effectively complete by 2010 - 2020.

Another insight provided by the punctuated equilibrium model is that the record of species in the fossil record manifests itself as long periods of stasis punctuated by rapid change or sudden jumps from one species to the next. The last of these punctuary jumps in scholarly communication was the spread of printing, followed by hundreds of years of relative stasis in the form (although not the numbers) of the scholarly journal. The current changes in communication technologies may well be the trigger for another punctuary jump followed by another (although probably shorter given the pace of technological change) period of stasis.

A final insight from the domain of palaeobiology is that patterns of speciation and extinction are chaotic, strongly influenced by the environment around them and very difficult to predict (as opposed to easy to infer with the benefit of hindsight). Stephen Jay Gould argues that if the history of life on Earth were to be replayed from the beginning, the likelihood of *homo sapiens* arising are vanishingly small [Gould, 1989]. The implications of this for any attempt to predict developments in electronic publishing technologies (and by implication the form of e-journals that rely on those technologies) are clear. These developments will also be hard to predict and influenced by the wider computing and communication environment.

One of the critical developments in the understanding of the effects of punctuated equilibrium on speciation and evolution is the fossil fauna of the Burgess Shale described by Stephen Jay Gould in his book *Wonderful Life* [Gould, 1989]. These organisms date back to 530 million years ago, just after the 'Cambrian explosion' - a time of dramatic diversification in evolution. The importance of the Burgess Shale is that it preserves many anatomical designs that existed for a while and then disappeared, never to be seen again. In accounting for this astonishing diversity, Gould points to three main factors:

- the first filling of the ecological barrel, providing empty ecological niches that life could rapidly colonize;
- a directional history for genetic systems, making significant change harder over time;
- early diversification and later locking-in as a property of all systems.

The Internet provides those empty ecological niches for new journals to colonise. The directional history of technology developments is evident in the ways that particular technology patterns (the dominance of Windows, the *de facto* use of Adobe's PDF as a page description standard) become rapidly widespread and resistant to change. The early diversification and later locking in are occurring around us in the e-journal domain at this very moment.

Genres and new media

Three of Agre's [Agre, 1995b] categories of analysis (communities, media and genres) provide a range of useful insights into changes in the form of hypermedia scholarly journals.

Different scholarly *communities* have always had quite different formatting requirements for print journals, traditions of using such journals and access to new technology. The Science, Technology and Medicine (STM) disciplines have always required more from their print publications in the way of the need for diagrams, equations, photographs and non-Roman symbols. There is some evidence ([Zuckerman and Merton, 1979], [Beyer, 1978]) that rejec-

tion rates in STM journals are significantly lower than in the social sciences and humanities. By the nature of their disciplines, the STM communities have tended to have earlier access to high technology of all sorts, including networked desktop computers. It is likely therefore that the STM journals will make the move to electronic delivery first and also use the more innovative delivery and presentation technologies. At present this seems to be the trend, with scholarly societies in the STM field driving the development of (mostly parallel delivery but soon electronic only) e-journals. Among the large commercial journal publishers, it is those with a large STM component (Elsevier, Blackwell Science) that are leading the move to parallel delivery.

With respect to *media* one of the most obvious differences between traditional print journal publishing and electronic journal publishing is the medium in which they are presented. Print has always been used for journals, monographs, conference proceedings, preprints, informal newsletters and the like. Off-line electronic media like CD-ROM and floppy disk have also been used for all of the above. The predominant electronic delivery mechanism at the moment (and into the future) seems to be some form of computer network, either the Internet or whatever it evolves into. Agre points out that the affordances of the medium govern how it is used. The affordances of the e-journal are clearly a concern to a number of readers, as indicated in their survey responses, and the design of new e-journals need to take account of these concerns and support existing usage patterns (where possible) or substitute improved alternatives (where not). The use of a live table of contents at the start of an e-journal as a (partial) replacement for being able to flip through the print pages of the paternal is one such example.

Finally, the scholarly journal as we perceive it today is a mature example of a *genre* that has evolved to be useful to members of a scholarly community. It is perhaps misleading to talk about the scholarly journal as *a* genre. In practice, there are a number of variant sub-genres including less formal newsletters, “letters’-type journals (*Physical Review Letters*), journals with an editor but no peer-review and anonymously peer-reviewed journals. Members of such sub-genres can be placed on continuums of refereeing rigour, type of content and speed of publication. Any replacement for the scholarly print journal will either need to reproduce the functions of this genre or replace them with new functions that are more useful.

Agre’s idea of ‘doing more’ when designing new media is a central one. One of the critical issues is the question of whether hypermedia scholarly journals are a new genre or just an old genre in a new medium. In a real sense, this relates to the question of evolution versus revo-

lution. At what point does an evolving new thing move from being quantitatively different to quantitatively new? Kaufer and Carley argue that researchers should not leap to declare something qualitatively different before first exploring the possibility of mere quantitative difference [Kaufer and Carley, 1993, p. 17].

Scholarly journals are almost always discipline-specific and are therefore intended for a particular scholarly community. It is instructive to consider what the answers to Agre's design questions might be in the context of hypermedia journal design for such a community.

Users: We probably have a fairly good idea of who our scholarly community is, although placing the journal on-line may well open the community up to interested outsiders. Do we want this? Are we prepared for the interactions that may result (particularly in areas like cosmology and evolution where people have strongly held views)?

Purpose: It is probably safe to assume that the purposes of hypermedia scholarly journals will remain the same as their print counterparts, at least for some time.

Activities: In an on-line environment, readers will have a wider range of potential interactions open to them than in print:

- moving directly to a particular known article (rather than reading the whole "issue')
- locating articles containing particular words
- locating articles containing particular media types (video/sound/simulation/image)
- tracking the ongoing peer commentary associated with an article
- using "forward-citation" links to identify all those articles that cite the one they are reading.

Usage: Do we keep the "look and feel" of the traditional journal (as many pilot projects are doing) or do we move beyond this? Will people's expectations of Web-based material (based on the other, increasingly commercial, Web sites they visit) influence what they expect to find in a Web-based e-journal? Is the appearance of the traditional article a means or an end? How might we respond to scientific articles that look like Encarta? What about scientific articles that look like a video game?

Medium: This relates to the 'do more' philosophy; what can we do that we can't already do in print?

Evolution: A Web site associated with a hypermedia journal would certainly add new content as new articles are submitted and as the richness of links in older articles perhaps increases. Considering this aspect of hypermedia journal design more broadly, it is possible to envisage a gradual evolution of the journal's presentation away from print traditions over time, in much the same way as print evolved away from simply copying the older manuscript traditions.

Awareness: For many journals, distribution is associated with membership in a scholarly society. E-journals linked to a print site-license would presumably be promoted internally within the site. If the journal wishes an expanded audience, either for increased subscriptions or to sell non-journal products (such as single articles or special supplements) it may need to advertise more widely.

Access: The increasing availability of technology has already been covered (see 4: Technology Developments on page 57). Over the next decade it is probably reasonable to assume that most university scholars in the developed world will gain access to the Web on desktop. The issues relating to the rest of the world, and to access away from the desktop are a real concern for some technological models, and a real drawback for e-journals relative to print.

The answers to these questions can only be determined through much more research with designers and users of e-journals.

Agre also points out that technology developments do not behave in the same way as developments in other fields because of what economists call path dependencies:

> … media industries are powerfully path-dependent because of effects deriving from the compatibility of different commodities …Once proprietary standards become entrenched in the marketplace, so that compatibility effects create ever-higher barriers to entry for potential competitors, their owners can start to extract rents from a variety of other parties. Moreover, network externalities (costs to individuals that derive from everyone else's choices) mean that dominance over a market tends to expand once it is established. [Agre, 1995b]

A good example of this in the domain of electronic publishing is the way in which Adobe's Acrobat technology is now the *de facto* standard for the distribution of 'electronic print'. As recently as three years ago there were a number of competitor technologies, some with superior technology offerings in some respects [Gruman, 1995]. Acrobat managed to gain a slight edge and that was enough to cause a swing in favour of its PDF technology to such an extent that the alternatives are now longer available on the market.

9.2.2 Insights from research literature

The Butler study [Butler, 1995a] received responses from 199 contributors (authors and editors) to e-journals relating to perceived benefits and disadvantages of electronic publication and to informal recognition and rewards. A number of her benefits and disadvantages are relevant to the form of the scholarly journal. Perceived disadvantages were inadequate graphics (listed by 38% of all respondents selecting this option), archival instability (35%) and potential for text alteration (8%). Advantages related to the form of the journal were speed of publication (71% of respondents) and the ability to publish materials unique to electronic formats (18%).

Part of Harter and Kim's research dealt with the data formats and access methods used in the pool of e-journals they were studying [Harter and Kim, 1996a]. They found that the majority of the e-journals (at that time - 1995) used ASCII text. HTML and Postscript were the second and third most commonly used formats. In this survey, PDF was used less than OCLC's proprietary Guidon format (now obsolete). The main form of access was WWW (66.4% of e-journals) followed by Gopher (42.4%), FTP (41.6%) and Listserv (38.4%) - the numbers do not sum to 1000 because many journals used more than one access method. It is important to realise that these online technologies are quite new and without the centuries of development that underpin the print journal literature. This is visible in access problems: the portion of this study dealing with citations to other online literature tried to follow these references to test their accessibility. Only 51.8% of these references could successfully be retrieved. For the online references that were Web URLs, this percentage only rose to 66%.

Schauder's study questions dealing with the preferred form of electronic scholarly publications [Schauder, 1994b, p.91] found little difference in respondents' attitudes, with CD-ROM, floppy-disk or networked media all deemed either preferred or suitable. His respondents were only moderately concerned that remote printouts appear identical to print originals, although there was a difference in responses by discipline area. Respondents from Physical Sciences and Engineering ranked this as more necessary than those from Arts, Social Scienc-

es, Law and Business, with respondents from Biological Sciences and Medicine being least concerned about print fidelity.

The study by Berge and Collins [Berge and Collins, 1996] of the readership of the *IPCT Journal* found that most respondents (in 1994) preferred ASCII format documents delivered via email. At that time, the preferred retrieval mechanism (40.9%) was via Gopher. They surmised that preferences would tip in favour of retrieval from the Web.

The survey of editors, authors and readers of *Journal of Biological Chemistry* (both print and online) [Gotsch and Reich, 1997] found that scientists regarded the print and online versions of the journal as complementary. Advantages of the online version were rapidity of searching, convenience, timeliness and efficiency of storage. Stated advantages of the print version were ease of reading, quality of figures, serendipity of discovery through browsing (repeatedly referred to), independence from technology while reading and portability. These two sets of affordances do not overlap, and it is very difficult to see how retain all the print advantages while moving to an online only version of the journal. When asked about desirable enhancements to the online version, the two most popular choices were more hypertext links and links to a bibliographic manager (both of which have now been implemented).

Woolfrey's study of the readers of the *Canadian Journal of Communication* found that about half were interested in receiving the journal electronically. Her conclusion was this was so they could search a journal database for articles.

The survey by Hitchcock, et. al. [Hitchcock, 1996] of STM journals available online in 1995 found that at that time Postscript still dominated at the more technical end of the spectrum. ASCII was being mostly used as an archival format. The main changes in the form of the journal that they anticipated were changes to the structure (more broadly-based e-journals) and frequency of publication (articles released as approved rather than waiting for a gap in a printing schedule). Their later survey [Hitchcock et al., 1997] argued that journal formats based on HTML (as the native document language of the Web) and HTML (as the *de facto* standard for printing from digital publication systems) will inevitably dominate e-journal production. They found that PDF (as a dialect of Postscript offering significant compression options) had come to dominate the available formats. In their view, whether a journal adopted HTML or PDF depended on the origins of the journal and its production arrangements. HTML typically was used where the formatting requirements were less critical. PDF was used for journals that had begun as paper only and were now moving to parallel delivery.

Olsen's study [Olsen, 1994] was largely carried out before the Web technologies and used informants with no experience of e-journals. As a result, much of her reported interview material deals less with new things that e-journals could offer and more with the indispensable attributes of print. Her list of key points related to print journals derived from her interviews was:

- the need to view high-quality graphics with text
- the poor ergonomics of reading on screen
- deficiencies of slow screen scrolling relative to fast page flipping
- the importance of serendipity
- the ability to annotate the text
- the 'weight' of text on paper
- the ability to browse

The improvements over print expected of an electronic system included:

- ease of access to the literature
- ease of searching
- reduced lag times in access to recent literature
- creation of a personal database of articles
- reduced space for storage

She argues that:

> In the same way as the printed text on paper has been honed over the centuries to achieve usability, so will the electronic text have to be designed and presented to allow users to achieve their objectives in using the material [Olsen, 1994, p. 32].

The question is, given the list of indispensable attributes (some of which will be very difficult or impossible to provide in an online environment) demanded by the scholars she interviewed how could one proceed to do this design? Would it be possible to sacrifice some of the benefits of print for extra advantages of online publishing? Olsen's conclusion is that:

> Building an electronic journal system cannot be viewed as the same kind of task as building the text systems we have experienced thus in information systems development … (it) will require a new frame of mind, an imaginative approach, and an unconventional configuration of technology [Olsen, 1994, p. 73].

Stewart's analysis of interview data from users of the CORE system shows that all of the respondents ranked most of Olsen's list of essential functions as at least 'important'.Their three areas of concern regarding an ideal e-journal system were period of coverage, presentation of graphics and presentation of text. Interestingly, fewer of Stewart's respondents than Schauder's [Schauder, 1994b] regarded exactly the same presentation of text online (via page images) as important.

Stewart's conclusion (quoting [Schaffner, 1994]) is that

> Electronic journals must, at the start, at least serve the basic functions that print journals have traditionally served. Once the transition has been made, new technologies may allow us to add new roles, to drop some of the traditional roles, or to fill them in intrinsically different ways [Schaffner, 1994, p. 240].

The question is, if it is not possible to provide some of these functions (fast flipping between pages, easy annotation) will added features like full-text searching and hypermedia additions be sufficient compensation?

Schaffner herself argues that enabling technologies may not be sufficient to bring about major change and that new forms of communication are slow to develop and take full advantage of new capabilities [Schaffner, 1994].

The analysis of user responses to the CORE system reported in [Entlich, 1994] provides a number of interesting pieces of data as well as some significant conclusions. They found that the ratio of printing to viewing of articles was 1:4. This was consistent across the two CORE interfaces (Scepter and Pixlook). (It is also consistent with anecdotal evidence about use of PDF documents inside Adobe). The interpretation of the CORE group with respect to printing is that "online full-text systems are used … as a convenient way to discover articles of interest, but not to read them in depth" [Entlich, 1994, p. 110]. This article ends by applying the insights from the CORE project to the form and functionality of the then (i.e. late 1995) current e-journals. Their main conclusions were:

- limitations in HTML's display of special characters remains a problem
- publishers should not overestimate the technical abilities or patience of their users
- there is no reason to expect that the tendency observed in the CORE users to read little on -screen and print articles for closers scrutiny will be any different for Web e-journals
- failure to provide adequate searching is commonplace

- assumptions that users will learn to how use a particular search system effectively are misplaced
- rapid response time is critical and therefore HTML abstracts (at the least) should be provided for PDF articles.

Loch and Huberman in their application of punctuated equilibrium theory to technology diffusion [Loch and Huberman, 1998] argue that in an environment where there are two alternative technology solutions with one being superior both technologies exhibit positive externalities (performance benefits from others using the same technology). The externalities initially produce two stable usage equilibria, one for each technology. These equilibria are subject to sudden punctuatory shifts, with one technology being adopted as the new standard. Critically, they found that the time before such an equilibrium punctuation takes place is dependent on the rate of incremental improvement of both technologies and on the system's resistance to switching between equilibria.

Ann Okerson and James O'Donnell make the excellent point that changes in technology change the way in which we categorise the world. She provides as an example the telephone directory and the novel. Both are currently included (for most of us) in the category 'book'. As efficient on-line indices with free-text searching and hyperlinks begin to replace the telephone directory, it will no longer be thought of as a book. This removal of an item from the category will in turn subtly redefine the category for us. As they delightfully put it:

> Even as we are all reading our Jane Austen on a summer's day in a hammock twenty years from now, the 'book' will have changed by virtue of the things that won't be in book form any longer." [Okerson and ODonnell, 1995', p. 2]

A good way to end this discussion of the effect of changes in the form of the scholarly journal is with a quote from the summary at the end of Appendix A of Gotsch and Reich's survey of scholars associated with *JBC*:

> The results of the survey point to a group very much in transition. Their answers can perhaps best be likened to respondents who were asked about the Iron Horse in the 19th Century. It was obvious that the steam locomotive was something important but no one had an inkling that it would revolutionize transportation to the extent that it did. [Gotsch and Reich, 1997]

9.2.3 Insights from thesis surveys

The surveys did not contain questions specifically targeted at the form of the scholarly journal because the original email survey was originally carried out for another purpose. However, it is still possible to make some useful comments on the results obtained and infer information from the pattern of responses.

One of the standard barriers listed to move to online scholarly publishing is the lack of technology in the reader population. At the beginning of 1997, over 60% of the total respondent pool had access to a personal computer with sound output, CD-ROM and colour display. The rate of technological change has if anything increased and most Wintel personal computers only have a three-year useful life. It is therefore reasonable to infer that the situation will only have improved since these surveys. The technology base for *hypermedia* journals exists.

Arguably, the more interesting electronic publications are those with an online component (either for delivery or feedback or both). Despite large differences between the different survey sub-populations overall slightly under 60% of respondents had access to a direct network connection and slightly over 60% had access to a modem connection to the Internet. Since these surveys, the push to network organisations has continued. The trend towards greater Internet access in the general population has also gained momentum. Again, it is reasonable to infer that the situation will only have got better. The technology base for *hypermedia online* journals also exists. It should be pointed out that the constraints of Internet access speeds would require careful consideration in designing the form of an online hypermedia journal if a significant proportion of users were to be accessing it from home via modem. Such a user population might include any of:

- professionals without access at work
- professionals working from home (either telecommuting or timeshifting)
- students
- interested amateurs.

The vast majority of current e-journals are accessed via the Web [Hitchcock et al., 1997]. Nearly 45% of the survey respondents used the Web frequently or regularly. Therefore they would have the skills to access these e-journals and this would not be a barrier to the journals transformation.

Without wishing to discuss individually each of the proposed advantages often given for e-journals (already discussed - see 7.3.4: Advantages of electronic scholarly publishing on

page 158), the overall pattern was that the respondents were very positive about e-journals. The majority of responses were for the Strongly Agree or Agree ratings in each question (with the exception of Multimedia for reasons already outlined). The survey populations were enthusiastic about the advantages which should encourage those who are advocating change.

With respect to the proposed disadvantages, the overall pattern was that responses clustered around the Neutral choice. In other words, respondents were less concerned about the disadvantages than they were enthusiastic about the advantages. The responses to the proposed Format Unfriendly disadvantage were noticeably lower indicating less concern about e-journal forms. Stronger responses were reported for the Refereeing, Copyright and Need for Equipment questions. Nonetheless, the disadvantages do not appear to be a major barrier to change.

9.2.4 Insights from case studies

A number of the libraries studied as part of the case study phase of the research are engaged in projects that were redefining the form of the scholarly journal in response to new technology. Highwire Press (see 8.3: Highwire Press on page 182) is responsible for *JBC Online*(see 6.3.1: Journal of Biological Chemistry on page 117), *Science Online* and a host of similarly innovative e-journals with a wide range of features impossible in print. Their user surveys are very positive about the features of these journals [Gotsch and Reich, 1997]. The Internet Library of Early Journals (see 8.4: Internet Library of Early Journals (ILEJ) on page 185) is transforming older journals by making them available to a wider audience and searchable (albeit imperfectly) for the first time. Project Muse (see 8.6: Project Muse on page 192) was at the forefront of this trend by taking an early decision to transform a stable of print journals into online existence while adding a host of features impossible in the print versions. Each of these projects demonstrates that the technology does make the transformation of print journals possible and successful.

9.2.5 Other insights

Developments in the field of computer and communications technology can also serve to illuminate some trends. The technologies for printing p-journals are proven over time and well understood. The technologies for performing equivalent activities in an online environment are still being developed. There is a lot of wheel-reinventing going on, and the Not Invented

Here (NIH) syndrome is alive and well. One likely development is the availability of turnkey scholarly publishing solutions.

An early example is already available. The Forequest Company, based in California, offers a server package called JournalOne™, specifically designed for publishing large publications on the World Wide Web. JournalOne is based on a proprietary networked client/server hardware configuration. The software supports delivery of a journal (or other periodical) via the World Wide Web, automates subscription and payment using completely secure, encrypted transactions. Site licenses to publications are also supported. ForeQuest's Web site is at <http://www.forequest.com/>.

A more generally applicable technology for Web publishing is provided by Groupsoft Publisher (available online at <http://www.groupsoft.co.uk/>). Based on Lotus Notes technology, this provides control over the content, structure and appearance of the Web site. The software has been designed to support timely release of time-bound information, a consistent look and feel for the site, and table of contents and full-text searching to assist with navigation. Groupsoft Publisher allows a small editorial team to manage a large and dynamic Web publication. By separating production into three distinct roles - site structure management, page appearance, and content - the online publication can be managed without specialised resources becoming a bottleneck. Its designers claim that assigning individual roles to people involved in production - journalists, editors, graphic designers - and using the powerful workflow capabilities of Lotus Notes, makes Groupsoft Publisher the first industrial-strength collaborative publishing system for the Web.

The Frontier product from Userland Software (available online at <http://www.scripting.com>/) is also designed to automate the production of large and complex Web sites. It has a newsroom management product which also enables a separation between content creators, content designers and content delivery.

9.3 Transformations in the function of journals

This section deals with the research question:

> *Will the move into an online hypermedia environment alter the functions of the scholarly journal?*

Function is defined specifically as those functions provided by the scholarly journal and its articles to the scholarly community. These functions can be summarised as:

- defining a citeable unit for acknowledgement
- fixing a scholar's views at a particular time
- providing gatekeeping or filtering of material
- providing timely communication of results and ideas
- enabling access to information within a given discipline

9.3.1 Insights from theoretical perspectives

Ecology of communicative transactions

The concepts in Kaufer and Carley's model that relate to the functions of the scholarly journal are *reach, asynchronicity*, and *multiplicity*.

Reach for an individual can be defined as the number of people whose mental model is affected by a signatured communication from that individual. Reach is a property of the author, but technology can extend an author's reach. Technology extends communication at a distance through *asynchronicity, durability* (discussed already under journal form) and *multiplicity*.

Asynchronicity removes the requirement that partners in a communicative transaction have to be coexistent in time (and by implication, in space). A move to online hypermedia journals will not alter this existing property of print journals. Such a move does have the potential to enhance discussion around journal articles particularly with innovations like the open peer commentary associated with journals like *JIME*.

Multiplicity is "the number of communication partners that can be communicated with at the same time" [Kaufer and Carley, 1994, p. 35]. Multiplicity implies greater distance and greater speed in spreading information. Network technologies provide for the largest potential asynchronicity, durability and multiplicity of any communications technologies to date. One of the attractions of e-journals is their ability to dramatically increase multiplicity. Large consortium licences allow groups institutions to gain access to journal titles online that they had not previously subscribed to in print at no or little extra cost. There is no (or little) incremental opportunity cost to the publishers in increasing such access. As an example, the author's university now has a site licence for all of the Project Muse journals (see 8.6: Project Muse on page 192), even though it had not subscribed to many of them previously. As a mul-

ti-site institution, having the full text online also simplifies access to titles that had been previously only been easily available within the campus on which they had been located.

Kaufer and Carley argue for a necessary role for print in the activities of scholarly professions in the sense that large diverse professions need to be structured around printed texts. But print is merely a supporting technology, not a deterministic one. The nature of professions depend on the characteristics of a group and not the medium through which they communicate. Like the later technologies of electronic mail, print increased the reach of individuals within a profession and thus supported a wider geographical spread of members. Print also bound the members of a profession more closely together through shared experiences of common printed materials in the forms of journals and newsletters. In diffusing new ideas journals are simultaneously faster than book publication or face-to-face interaction (due to their frequency of issue and increased reach respectively), and slower than newspapers (due to the gatekeeper function of peer review). The obvious question is whether the current system is too fast or too slow. The consensus according to Kaufer and Carley is that many scientists regard the speed of journals as too slow, particularly in very fast-moving fields.

Punctuated equilibrium and speciation

The significance of the punctuated equilibrium model to the transformation of the scholarly journal is that speciation (and hence changes in the scholarly journal) are driven by environmental change. For example, one could (just) imagine the reward structures for scholarship altering so that publishing refereed journal articles was regarded as significantly inferior to excellence teaching. In this case, the pressures for changes in scholarly communication would be quite severe. However, if one makes the reasonable assumption that scholarship is not going to be radically transformed in the next century, then it follows that the functions of scholarship currently embodied in the print scholarly journal system will not dramatically alter either. Instead, technological transformation should allow scholars to carry out their existing functions more effectively.

Genres and new media

Agre's categories of analysis (*activities* and *relationships*) are relevant to the functions of the scholarly journal and its transformation into an online hypermedia environment.

The functions of the scholarly journal are designed to support the *activities* of scholarship. These activities of writing and reading articles, conference papers and books, and taking part in the processes of journal publishing by refereeing and editing are a core part of the scholarly life. They serve as the way in which research is communicated to one's peers and validated as worthy of dissemination. Any functional transformation associated with new hypermedia journals needs to build on this established rich pattern of activities. This is a theme that is reinforced by the survey data (see below).

Scholarly journals also need to support the *relationships* within scholarly communities. Hypermedia journals should do no less than existing print journals, but can potentially do much more through increased possibilities for feedback to the author and interaction between author and reader. The sort of online peer commentary being trialled by *JIME* is an excellent example of how the form of a journal can provide a sense of embedding within an online community.

9.3.2 Insights from research literature

Julene Butler's study [Butler, 1995a] contained a number of questions relevant to journal function. The two most frequently selected disadvantages were the perception that e-journals were not real publications (63%) and that they were less prestigious (54%). These perceptions will inhibit the ability of e-journals to communicate effectively. High-ranking advantages were that e-journals allowed the author to reach the best audience (55%) and enhance scholarly dialogue (48%), both critical journal functions. Importantly, only 22% of Butlers respondents felt that their superiors rated e-journal publication as equal to or better than print. If not changed, this will act as a significant brake on any transformation of the journal.

Harter and Kim's citation study looked by implication at whether e-journals were fulfilling the function of acting as an accessible journal of record and effective means of communication. One of their significant findings was that the citation styles of the online references in the e-journals they were considering was frequently inconsistent, incomplete and/or inaccessible [Harter and Kim, 1996b]. This means that they were deficient in contrast to print publications. As they point out, "clearly the accessibility of cited online resources is potentially a very serious problem in the conduct of research and scholarship" [Harter and Kim, 1996b]. Moreover, even among e-journals, there was very little citation of the e-journal literature. Their conclusion: "e-journals presently play almost no role in scholarly communication, as measured by references cited" [Harter and Kim, 1996b].

One of the assumptions made in the design of *IPCT Journa*l and reported by [Berge and Collins, 1996] was that scholars read articles, not journals. This has implications for the future of the journal itself, and the bundling of articles into issues. If scholars are primarily interested in articles, then what function does the journal now have? Only 10% of their respondents indicated that they would be retrieving all the articles in a specified issue. Only one article was named by over 50% of the respondents. This fits with anecdotal and citation study evidence from the print world that most scholars do not read entire articles and that most articles are read very little. The implications of this will be taken up in the Conclusion to this thesis.

[Hitchcock et al., 1997] cites one example of a publisher that is using the potentials of online publishing to broaden the functions of the journal into wider support for the activities performed by a scholarly community (as discussed by Agre). The Institute of Physics (IoP) is providing a range of online services in conjunction with its delivery of the full text of its journals. These include:

- free access to the Letters and Rapid Communications sections of 12 of their most prestigious journals
- advance notice of abstracts of forthcoming papers
- product finder service
- directory of peers' email addresses
- jobs exchange
- ability for users to customise their login screens

Highwire Press is also working with the American Association for the Advancement of Science (AAAS) to extend their online version of the prestigious journal *Science* (available at <http://www.sciencemag.org/>). *Science Online* now offers Science Now (daily articles that will appear in the next issue of *Science* as well as additional material), Next Wave (a range of resources for the next generation of scientists), Science Careers (with links to employers, job listings and a resumé bank), and Science E-MarketPlace (which provides information about products and advertisers appearing in *Science*). This is a deliberate move to transform the Science World site into a location that can provide a range of additional services for the scientific community. A number of scientific professional societies (the American Society for Biology and Molecular Biology, the American Chemical Society and the American Meteorological Society) have found this additional content sufficiently useful that they have licensed access to Science Now and Next Wave for their own members.

It will be interesting to see if other publishers take up this initiative. It is certainly in line with the current trend for general web directories like Altavista and Yahoo and Web-focused organisations like Netscape to transform themselves into Internet 'portals' for their user communities.

[Hitchcock et al., 1997] also points out the general trend of publishers *changing* the functions of the journal by providing a range of integrative services as part of their offerings. Three of the current eLib projects he describes that as taking this tack are:

- NewsAgent for Libraries (personalised information services, information filtering, software agents, metadata)
- Open Journal Project (derived hypertext links for collections of resources)
- SuperJournal Project (value-added features such as search, display, and multimedia based on clusters of journals in subject areas)

9.3.3 Insights from thesis surveys

The surveys also did not contain questions specifically targeted at the function of the scholarly journal. It is reasonable to infer from the concerns about Refereeing that respondents see the quality control function of journals as important and therefore one that should be retained. The concerns about Copyright suggest that journals need to retain their current advantage of providing protected and branded content.

9.4 Transformations in stakeholder roles

This section deals with the research question:

> *How might technology transform the roles of the stakeholders in the scholarly journal ecology?*

The stakeholders have already been discussed (see 3.3: Stakeholders in the scholarly journal ecology on page 51).

9.4.1 Insights from theoretical perspectives

Ecology of communicative transactions

One of the axioms that arise from the application of Kaufer and Carley's assumptions (see 2.2.1: Initial assumptions on page 23) to communicative transactions is that

> "Individuals exert communicative authority by changing the mental models of others through communications bearing their signature" [Kaufer and Carley, 1994, p. 28]

It would seem that the added speed and reach of electronic journals would increase the ability of authors (and readers who are providing commentaries) to exert this communicative authority and thus subtly enhance their respective roles.

They also argue that the nature of professions depends on the characteristics of a group and not the medium through which they communicate. This would support the argument that moving to online hypermedia publishing would not change the nature of scholarship *per se* (although there is no reason to conclude that it would not alter the roles that scholars can play).

Punctuated equilibrium and speciation

The application of the theory of punctuated equilibria to an analysis of the fossil record supports an interpretation of rapid change rather than 'phyletic gradualism'. Generalising to scholarly publishing would suggest a similarly rapid change in at least the species of scholarly journal. It is not certain that this would mean a similarly rapid change in stakeholder roles, although there are some recent signs of movement in the relative stasis of the last fifty years in scholarly for-profit publishing. One is a recent agreement from Elsevier to allow libraries to make paper copies of requested articles appearing in electronic journals and share them with scholars at other institutions (in the same way as for normal inter-library loan) [Kiernan, 1998a]. Another is the recent decision by the 81 libraries in the Scholarly Publishing and Academic Resources Coalition and the American Chemical Society to produce a new organic-chemistry publication that will be considerably less expensive than its leading competitor, *Tetrahedron Letters*, published by Elsevier Science. The coalition plans to start other low-cost journals in collaboration with other publishers [Kiernan, 1998b]. This initiative by libraries and a scholarly society to work in competition with an existing journal and publisher is a significant shift in role.

Genres and new media

The insights provided by Agre's analysis of the role of genre (and its related terms of community, activity, relationships, and media) in the design of new media [Agre, 1995b] are relevant to every aspect of the transformation of roles. The roles that are being transformed are exercised in Agre's communities, shared forms of activity within a particular institutional

logic. The communities of stakeholders in the scholarly publishing ecology are defined in terms of the activities they perform. The communities are also internally dependent on the patterns of relationships between members as well as being themselves linked via a web of relationships to other communities within the ecology. The entire ecology has been organised around the production of journals in the specific medium of print and is now reorganising around multiple media. All of this is taking place within the genre of the scholarly journal and all the expectations and default settings that this implies. All of this suggests strongly that evolving this system will require interventions and decisions by multiple players at multiple levels and is emphatically not a simple task.

9.4.2 Insights from research literature

Part of the survey by Gotsch and Reich [Gotsch and Reich, 1997] dealt with the role of librarians as institutional subscribers to *JBC*. They found that in smaller organisations the role of the librarian with respect to collection policies was largely clerical. In larger organisations the librarians played a much larger role. This gatekeeper role with respect to electronic publishing acquisition was strongly influenced by cost. They argue that the traditional library role of bringing together the users demand for publications with the publishers supply will continue given current technology and divisions of labour. However, this role is carried out in an environment of budget and institutional constraints which have a profound influence on the outcome. Atkinson agrees that the library has an ongoing role to play between "the information seeker and the information sought" [Atkinson, 1993, p. 211].

[Hitchcock et al., 1997] provides an example of the new technologies allowing a stakeholder to change roles. Catchword Publishing was originally founded on production and subscription services. They now aim to provide a complete scholarly publishing environment on the Internet, based around their proprietary RealPage document format technology. Like Adobe's Acrobat, RealPage converts from Postscript but claims to produce smaller file sizes. The intention of Catchword is to host a significant number of publications (at the time of writing 131) and provide support for managing subscriptions. In effect, they have added publication hosting (a form of publishing) to the services they already provide.

Fytton Rowland argues for little change in roles from the current *status quo*:

> in the long term journals will be a lot cheaper than the print journals of today; that some of them, despite their lower prices, will be sold for profit; that the media conglomerates that dominate publishing today will still be involved in selling some of

them; and that individual academics, groups of academics, learned societies, and university presses will publish others. In other words, I expect to see a diversity similar to today's, but with the competition from academics hostile to publishers having driven prices down [Rowland, 1996]

Atkinson argues on the other hand that:

> It is very unlikely - and it would certainly be very undesirable - for the commercial publishing industry to continue to play the same dominant role in scholarly publication in the online environment that it has in the paper environment: that would be economically unacceptable and technically unnecessary. [Atkinson, 1993, p. 211]

9.4.3 Insights from case studies

As part of the detailed interviews carried out with members of these projects, a cluster of issues emerged that related directly to libraries and their changing roles in an evolving publishing ecology. These can be broadly defined as the nature of the core business of libraries. Themes arising from the interviews have been presented in an aggregated form without identifying the specific source of each comment. This is because it is the total picture that is important, not the source of each colour. These responses are probably not representative of the full diversity of views in the library community. They do, however, serve to provide useful pointers to some future directions.

The interview questions that related specifically to the libraries role were:

- What is a libraries ‘core business’?
- Is this sort of initiative (i.e. the specific project) part of that core business?
- Can electronic publishing (in general) become part of a libraries core business?

In the answers to these questions, the key theme was the role of electronic publishing in fulfilling the university library’s mission to its community. As part of this, once again the questions of preservation and access loomed large.

A number of projects identified what they were doing as explicitly part of their mission to provide access to information or supply information to their community. A number of respondents also emphasized the role of the university library in furthering teaching, learning and research. Providing a service to users was a common thread in the responses.

Based on this service orientation, all of the projects saw what they were doing as a natural outgrowth of their mission to their communities. Comments like '(this project) is a way to provide access outside the library' were typical. Despite the reluctance of a number of projects to identify what they were doing as publishing, all felt it fitted well into their libraries portfolio of activities. A number also stated the need for more such initiatives from other libraries. One respondent explicitly stated the need for libraries to provide alternative models to existing publishers in the new field of electronic scholarly journals. Another talked about library-initiated or facilitated publishing as being one way for libraries to assist their users and take back the initiative from publishers.

All felt that there was no 'in principle' barrier to electronic publishing becoming part of the core business of all academic libraries, although this might depend on particular campus circumstances. One respondent put it best by stating that if it made sense for the library to do something and if users would expect this to live at the library, then the library should go ahead. Only one of the projects was based at a university with an existing press. This no doubt made it easier for the other projects to move into a new area without any precedent of ownership by another stakeholder.

9.5 Transformations in stakeholder practices

This section deals with the research question:

> *How might technology transform the practices of the stakeholders in the scholarly journal ecology?*

The key idea here is that technology enables the transformation of practices because it either makes existing things easy and/or new things possible.

9.5.1 Insights from theoretical perspectives

Ecology of communicative transactions

The coadaptive, coevolving model of Kaufer and Carley suggests that transformation of practices by one stakeholder in response to technology will have effects on all the other stakeholder practices. Examples already discussed have been the move by print publishers to parallel online delivery in response to new e-journal publishing by scholars and libraries. The scholarly publishing community is also seeing gradual change in licensing costs in response to pressure from libraries.

Punctuated equilibrium and speciation

Punctuated equilibrium suggests that changes in practices will occur over a fairly short period of time. A good example has been the shift from very few e-journals available in 1995 to the majority of the large publishers either offering (or promising to offer) electronic versions at the end of 1998.

Genres and new media

Agre's analysis suggests that genres (like the scholarly journal) are relatively stable and expectable forms of communication. It is unlikely therefore such genres will change quickly. However he also points out that genres are usually linked closely to a particular medium: sitcoms to television, plays to the theatre and so on. It will be interesting to see if the scholarly journal genre can successfully make the transition to a new medium without being significantly changed in the process. Examples of such change are journals like *JIME* which offer features like ongoing peer commentary that are not possible in the print-medium version of this genre.

In common with a number of commentators [Gaver, 1992], [Gibson, 1979], Agre notes the dependency between the affordances of a medium and how it will be used. This implies that a new medium will provide new affordances and therefore new usage practices. Scholarly communities and their activities may not change much, but they may well be carrying them out using new media and in new ways.

9.5.2 Insights from research literature

Julene Butler [Butler, 1995a] found that electronic publication was serving to expand the informal contacts of some contributors, as well as to provide more (and more substantive) feedback on their work. This indicates that the combination of electronic publication and communication is increasing information flow between scholars.

One piece of evidence for transformation of citation practices among scholars chiefly comes from Harter and Kim's citation analyses (see 1.8.2: Citation studies on page 13). In their study examining citation practices among authors of e-journal articles ([Harter and Kim, 1996b], [Harter and Kim, 1996a]), they found that only 1.9% of the total references in the sample e-journal articles were to online sources of any type. Moreover, 81.8% of these online references were to just three e-journals (*Public-Access Computer Systems Review* (PACS-R), *Electronic Journals of Virtual Culture* (EJVC), and *E-journal*), and to an influ-

ential subset of all the articles within these e-journals. It is difficult to see how e-journal publishing will become more credible if even those writing in e-journals do not cite the e-journal literature more.

The respondents in Lancaster's survey of academic administrators and library directors [Lancaster, 1995a] found that in 1993 they were not optimistic about the possible advantages of networked publishing being realised. They did see the potential for faster publishing of research activities. With respect to obstacles to transformation of scholarly publishing practices, the greatest ones were "those associated with the academics establishment's ability to implement, manage, and support a publishing network" [Lancaster, 1995a, p. 744]. The academic reward system was not considered an impossible barrier with respondents feeling some hope that epublishing would be acceptable in tenure reviews.

The question in Schauder's [Schauder, 1994b] study that most directly deals with changes in practices also relates to the question of academic rewards. Asked whether they believed that their institutions would rank electronic publishing equally with print publishing, 33% didn't know, 35% answered yes, 19% answered to some extent, and 12% answered no. Positive responses were correlated with seniority. With respect to whether their university should become more active in electronic publishing of professional articles, 30% didn't know, with slightly more (32%) answering yes. In this case, junior academics were the most positive and senior academics least so.

In Berge and Collins study of *IPCT Journal* [Berge and Collins, 1996], their respondents generally regarded its articles as of the same quality (47%, with another 37.7% indicating they had no basis to answer this question) as scholarly refereed print journals. With respect to the attitude of their superiors to electronic publication in refereed journals, only 12.7% indicated that the author would received the same points as for print publications. The other responses were somewhat less points (14.1%), and no points (10.8%). The rest of the responses were either not applicable or don't know. Obviously, very few authors would be prepared to publish electronically if this was to have a negative effect on their publishing record (at least until they had received tenure).

In their final author survey, [Gotsch and Reich, 1997] asked whether the decision to move to an on-line only version of *JBC* would adversely affect their decision to publish in the journal. Approximately 35% replied in the affirmative. They point out that a substantial number of authors did not use the online version at all, many because of access problems. One might conjecture that, in the light of the generally positive comments about *JBC Online* from re-

spondents who did have access, improvements in access for the authors who indicated an adverse response might reasonably be expected to decrease this percentage.

In the section of the survey by Gotsch and Reich [Gotsch and Reich, 1997] that dealt with librarians as institutional subscribers to *JBC*, they found that the practices of acquisition were strongly influenced by cost. Their prediction was a slow and cautious move to cancel print subscriptions and substitute electronic ones if this would have an impact on already constrained serials and monograph budgets. This clearly has implications for journal publishers who are arguing for a subscription increase if libraries want both print and electronic delivery.

9.5.3 Insights from thesis surveys

Overall, the survey respondents showed only moderate use of e-journals. However, the email respondents on the *Psyche* mailing lists (who can be assumed to be early adopters of technology) were much more frequent users, with nearly 50% viewing e-journals frequently or regularly. This contrasts with an equivalent figure of less than 20% for the print survey respondents. One can thus extrapolate that e-journal use is likely to increase over time.

One of the survey questions asked how often respondents published electronically themselves. The answers to this question reflects a significant change in practice among part of the scholarly stakeholder community. The print survey respondents had done almost no electronic publishing. In contrast nearly half (47%) of the email survey respondents had published electronically at some stage, with nearly 10% doing so either frequently or regularly.

This latter finding is consistent with the results from the questions dealing with attitudes to the advantages and disadvantages of electronic scholarly publishing. The email respondents (relative to the print respondents) were simultaneously more in agreement with the proposed advantages of electronic scholarly publishing and less in agreement with the proposed disadvantages. They are therefore more likely to publish electronically than those with less positive views.

9.5.4 Insights from case studies

The cluster of issues arising from the case study interviews that is most relevant to the transformation of library practices relate to the issue of publishing. The questions that were asked under this topic were:

- How do you define publishing?

- Do you regard this just as the initial release of content or does it include keeping the content available long-term?

The key themes running through people's responses to these questions were the importance of access and the need for archiving.

Typical comments on the theme of access were that publishing is 'making information publicly available' or is 'the act of making information public'. One respondent with a commercial publishing background preferred to talk about shaping raw material into something that the market wanted, and then disseminating that shaped material.

Two of the respondents were reluctant to identify what they did as 'publishing', although the question did not require them to. One felt that publishing required some sort of distribution (which their project did not do), but acknowledged that this was an increasingly problematic distinction to make if one was referring to electronic documents on servers. Another respondent preferred to define the library function as providing access to material that might (or might not) be owned by the library. Both of these respondents evinced a disquiet about using a term like publishing that evoked a range of no longer relevant print associations. Both stated the need for a new word or model to describe what they did. One suggested the term 'digital press', but acknowledged that press had inappropriate and over-physical connotations also.

The other key theme was the need to have a changed view of one's commitment to the content in an electronic world. Once print has been produced and sent out, there is no necessary ongoing commitment from the publisher. In an electronic content environment, there is now the need to provide ongoing access and maintenance. Such maintenance might just be keeping the servers up and connected or making sure that URLs still worked. It might also include adding forward references to existing content or improving the interface. In the longer term, it might extend to migration of the content as the presentation and delivery technologies changed. Most of the respondents expressed scepticism that traditional publishers would take on this responsibility, particularly without an ongoing revenue stream. Most also explicitly identified the library as having had this archiving role in the past and needing to now evolve an equivalent as the technologies change. However, a number stated that there should not be an automatic assumption that the library would perform this archiving function, particularly given the possible costs.

9.6 Conclusion

This chapter has interpreted the findings of the research in terms of the theoretical perspectives, the insights from the research literature, the thesis surveys and the case studies. In each case it has considered the relevant findings with respect to transformations in the forms and functions of journals and the roles and practices of stakeholders. It has found support from each of the sources of insights for these transformations. The final chapter will take these interpretations and conclude by considering their implications for the future of the scholarly journal.

10 Conclusions

10.1 Introduction

The previous chapter discussed and interpreted the results of the thesis research in the context of the research questions. The theoretical perspectives used, the literature review and the survey and case study work all have things to contribute to the answers to the research questions. This final chapter will present the conclusions that can be drawn for each of the questions and see what the answers to these questions are saying about the transformation of the scholarly journal.

10.2 Transformations in the form of journals

This section will consider the conclusions with respect to the impact of the transformation of the scholarly journal on its form.

10.2.1Future of the journal as artefact

Changes in communication technologies have always taken time to be adopted. Typically, it takes a while for the implications of the new technologies to be appreciated. The conclusion is that the scholarly journal will continue its transition into an online environment but will do so in two overlapping phases.

The first phase chronologically is the development of new e-journals which have never had a separate print existence (of where the print is clearly secondary). With the exception of a few early projects this commenced in the early 1990's and is continuing today. These e-journals often have features impossible to provide in print but still provide printable versions (often in PDF or HTML) of their articles. This still ties the articles to the lowest common denominator of print. The technology for reading on screen will improve. People will become more accustomed to using the interface features (hyperlinks, searching, zooming, interactivity, dynamic content, etc.) that are impossible in paper. Workable onscreen annotation systems will appear. Gradually, the gap between the online and print versions will become so great that authors will write for the online environment and their articles will not be printable in any usable way. Print will gradually wither and die. People will not have shelves full of journal reprints, but disks full of journal 'rebytes'. Only then will we have the true hypermedia online journal.

The second phase is the move to parallel publishing by the existing print journal publishers, driven by a variety of factors. In time (perhaps as soon as next year), the parallel publishing initiatives will start to drop print and provide the journal online only (with a printable version, probably as PDF) for users who wish to use this. Once this transition starts, the print journals will continue along the same path as the original e-journals with about a five to ten year lag-time.

The transition to online delivery only for the overwhelming majority of scholarly journals will be complete by 2020. The transition will not be linear, but will follow the classic sygmoid (S-shaped) curve. Different disciplines will move at different rates and for different reasons. The drivers for this transition are complex (and have been discussed already) but will certainly include journal economics, speed of communication, and the limitations of static media.

The journal as a genre will continue but in altered form. It has always been a means to an end, not an end in itself. In its current print incarnation it has been (until recently) the best way of performing the various functions necessary for efficient and effective scholarly communication. A better set of technologies should enable the same functions (and probably additional ones) in a better way.

10.2.2A new technology stasis?

Punctuated equilibrium argues for relatively rapid large scale shifts after longs periods of stasis. This has certainly been true in the domain of scholarly publishing. Applying this insight to the new electronic publishing technologies suggests that after the current phase of rapid development in computing and communications hardware and software, there will be a return to another period of relative stasis. Signs that this might be the case are:

- A convergence on the de-facto standards of HTML for onscreen display of articles and PDF for printable versions
- A noticeable slowing in the release timetables of new versions of critical software such as browsers
- The approaching quantum-mechanical barriers to the continual operation of Moore's Law to provide ever cheaper and faster chip technology
- The lack of anything dramatically different (in the way the arrival of Web was, for example) on the technology horizon

These all argue for incremental rather than dramatic change over the next 20 years (which was the case also for the first 50 years of printing in Europe).

10.2.3Archiving the e-journal

Part of a new technology stasis will be the need for a secure and reliable system of archiving. Any system for archiving e-journals has two components: the form in which the e-journal is archived and the system whereby the archiving takes place (covered later - see 10.5.2: Who does the archiving? on page 234).

10.3 Transformations in the function of journals

This section will consider the conclusions with respect to the impact of the transformation of the scholarly journal on its functions.

10.3.1Evolution or revolution?

The scholarly journal in its current form performs a number of critical functions for the scholarly community. These functions are deeply embedded into the nature of contemporary scholarship and will need to continue to be performed into the future. The evidence of the surveys and literature is that these functions are still important. To the extent that e-journals do not fulfil these functions they will be regarded as inferior.

At present, e-journals (as such) are having little impact on the processes of scholarship. Parallel published journals will have slightly greater impact (because of greater reach) than their print equivalents. Over time, provided they can do at least all that p-journals do now, the new e-journals will increase their impact because of the additional features they offer. Moves to make e-journals more closely embedded in the processes of scholarship will support this trend.

The conclusion is that we will see gradual evolution in the functions of the scholarly journal as it becomes transformed (and faster evolution in its form), not revolution.

10.4 Transformations in stakeholder roles

This section will consider the conclusions with respect to the impact of the transformation of the scholarly journal on the roles of stakeholders in the scholarly communication ecology.

10.4.1Interlocking systems and interdependencies

The transforming capabilities of technology may not necessarily cause stakeholders to take on new roles. Academics can now self-publish (via their own journals or through easier distribution of preprints) but does this make them publishers? Libraries can publish, but the case-study participants preferred to talk about technology letting them serve their communities better rather than letting them change roles. There will be some instances of stakeholders shifting roles, but these will be the exceptions rather than the rule for some time to come. The existing webs of genres, communities, activities, relationships and media will serve to constrain the ability of the system to change rapidly with respect to roles (as opposed to with respect to delivery technologies).

10.5 Transformations in stakeholder practices

This section will consider the conclusions with respect to the impact of the transformation of the scholarly journal on the practices of stakeholders in the scholarly communication ecology.

10.5.1Technology as enhancer

The key finding here is that the transforming capabilities of technology are able to enhance the activities of stakeholders. Scholars will be able to work more effectively and access more information. Libraries will be able to increase access to their users. Publishers will be able to provide richer offerings of information. However, once again, system constraints such as recognition of electronic publishing in reward structures seems likely to be a constraint (based on the published research literature – this was not a question on the survey). Use of e-journals (particularly parallel published journals) is likely to run well ahead of a significant move to publish in e-journals (particularly electronic only journals).

10.5.2Who does the archiving?

Libraries are reluctant to trust publishers to either maintain archives or maintain access. Publishers are reluctant to maintain the archives themselves, but also reluctant to allow libraries to obtain the content they have subscribed to (as opposed to access to that content). This is because publishers are concerned that they may lose control of their easily-copied digital content (something that is less of an issue with print information). This suggests at least two possibilities. One is that scholarly societies (who are committed to their members) or librar-

ies (who have always seen archiving as part of their mission) will increasingly act as publishers and ensure continuity of access and migration of content. Another possibility is the emergence of what are effectively scholarly content escrow organisations (such as OCLC's ECO) who can be trusted both by libraries and publishers to behave appropriately and who can develop the expertise in archiving digital content while benefiting from economies of scale.

10.6 Final thoughts

If one considers the current state of electronic scholarly communication, it displays many of the characteristics of the early explosion in the diversity of life on earth. The new ecological niche is that of online publishing, free from many of the constraints of the print world. The analogy should probably not be pushed too far, but in a moment of whimsy one might think of the first e-journal as being a little like the first lungfish pushing its way arduously up some primeval beach on its way to start the colonization of the land.

It is certainly true that the last decade has seen a great diversity of forms of electronic scholarly communication: ftp-based journals, mailing lists, journals on CD-ROM, Web-based journals, MUDs/MUSHs/MOOs as collaborative and publishing spaces, and proprietary SGML-based journals. We are starting to see some of this diversity being narrowed down as particular forms are abandoned and as the scholarly world standardises on a subset of these early experiments. Early proprietary experiments like the OCLC Guidon interface have been abandoned in favour of open standards like the Web. Many electronic publishers are choosing to standardise on Acrobat for parallel print/electronic delivery, or where the formatting requirements make HTML problematic.

In terms of the theory of punctuated equilibrium, it appears that we are in the middle of a punctuary jump after the stasis of the last century (at least) of print journal publishing. New species of communication artifacts are emerging to fill the new online niches. Existing players in the scholarly communication ecology are changing their roles and evolving within the new environment. If the analogy with punctuated equilibrium holds, then a new period of stasis should be anticipated. The prediction of this thesis is that the overwhelming majority of print scholarly journals will have completed their transition to an online existence by 2010 and we will then see another (although perhaps shorter, due to the pace of technological change) period of relative stasis. In the move to this new online status quo, there is the po-

tential for new stakeholders to emerge and existing stakeholders to redefine their roles. This thesis has served to illuminate some of this change.

VI Appendices & References

“Whoever reports a thing in the name of the person who said it brings deliverance into the world” *Pirke Avot* (Ethics of the Fathers), 6:6

11 Survey Instruments

11.1 Email survey instrument

This is the survey that was distributed to the readership of *Psyche* via the PSYCHE-D and PSYCHE-L mailing lists (see 7: Surveys on page 127).

```
QUESTIONNAIRE: Electronic Scholarly Publishing
----------------------------------------------

A.     INTRODUCTION AND PURPOSE

We would like you to take part in this survey because you receive either the
ejournal _Psyche_ or the print journal _The Australian Journal of Chemistry_.
This survey is part of a larger joint project between the Royal Melbourne
University of Technology (RMIT), Monash University and the Australian
Commonwealth Scientific and Industrial Research Organisation (CSIRO). This
project has been funded to investigate electronic publishing by the Australian
Vice-Chancellor's Committee (AVCC).

This  questionnaire is designed to provide a broad picture of the demographics
of the readers of the above journals. A sample of this population will then be
selected for a more detailed follow-up questionnaire. You will be asked if you
wish to be part of such an exercise towards the end of this questionnaire.

Please fill in the blanks and return the questionnaire. A range of return
mechanisms (mail, email, fax) are available. Space is provided near the end of
the questionnaire for additional comments if required. This questionnaire is
also available online at http://www.deakin.edu.au/people/aet/initsurv.txt.

B.     BASIC DEMOGRAPHICS

1.     Industry Category

Please indicate the industry sector of your employing institution. Choose
one category and mark with a cross.

__ Education (Primary/Secondary/Tertiary)
__ Government
__ Industry
__ Consultant/Self-Employed
__ Other (please indicate) _____________________________

2.     Position

Please indicate your primary employment role. Choose one category and mark
with a cross.

__ Practitioner (e.g. clinical psychologist, industrial chemist)
__ Information Services/Publishing
__ Administration
__ Teaching
__ Student
__ Research
__ Consultant/Self-Employed
__ Retired
__ Other (please indicate) _____________________________

3.     Access to Technology

Please mark the pieces of information technology you have ready access to
(either at home or at work). Mark all appropriate items with a cross.
```

__ Personal Computer
__ CD-ROM drive
__ Sound output from computer
__ Colour screen
__ Direct network connection to Internet
__ Modem connection to Internet
__ Other (please indicate) _____

C. ELECTRONIC PUBLISHING

4. Familiarity

Please indicate how often you use the following forms of electronic publishing with the appropriate letter:

a. Frequently b. Regularly c. Occasionally d. Rarely e. Never

__ Subscribe to electronic publishing forums (listservers, BBS, netnews, etc.)
__ Use FTP (File Transfer Protocol) to access materials
__ Use Gopher to access materials
__ Use World Wide Web to access materials
__ Access materials on CD-ROM
__ View electronic journal(s)
__ Publish electronically

5. Advantages and Disadvantages

Below are two possible lists of advantages or disadvantages of electronically published scholarly articles. You may add any items to the lists that you wish. Please indicate your feelings about each advantage or disadvantage to the left of the suggested text according to this scale: SA - Strongly Agree, A - Agree, N - Neutral, D - Disagree, SD - Strongly Disagree.

ADVANTAGES of electronic scholarly journals are:	DISADVANTAGES of electronic scholarly journals are:
__ Speed of publication	__ Poor quality
__ 24 hours a day access	__ Lack of refereeing
__ Convenience	__ Concerns about copyright
__ The way they encourage feedback	__ Increased plagiarism
__ Reduced paper consumption	__ Special skills needed to use
__ Ease of searching	__ Special equipment needed to access
__ Multimedia publications	__ Format is not reader friendly
__ Affordability	__ Communication costs to access them
__ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	__ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _
__ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	__ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

D. FOLLOW-UP SURVEY

6. Do you wish to take part in a more detailed survey?

We are seeking a select group of individuals to trial a number of new electronic and print publishing products. You will be asked a more detailed set of questions as part of this trial. If you are prepared to take part in this exercise later this year, please provide your details below. This will be regarded as an expression of interest only, and does not guarantee that you will be selected.

Name: _____
Address: _____
_____
Phone: _____
Fax: _____
Email: _____

```
******************************************************************************
Any other comments?




******************************************************************************
Please return this questionnaire before December 15, 1995, via any one of:

Post:
Andrew Treloar, School of Computing and Mathematics, Deakin University,
662 Blackburn Road, Clayton, 3168, Australia.

Email:
Andrew.Treloar@deakin.edu.au

Fax:
+61 3 9244 7134

Enquiries to Andrew Treloar at the above addresses or phone +61 3 9244 7461.
******************************************************************************
Based in part on a proposed questionnaire designed by Philip McEldowney,
University of Virginia (philipmc@Virginia.edu). Last revised September 5, 1995
******************************************************************************
```

11.2 Print survey instrument

This is the survey that was distributed via the postal system to members of psychological associations in the U.S.A., the U.K. and Australia (see 7: Surveys on page 127).

ATTITUDES AND ACCESS TO ELECTRONIC SCHOLARLY PUBLISHING

A. BASIC DEMOGRAPHICS

1. Industry Category

Please indicate the industry sector of your employing institution. Choose one category and mark with a cross.

❏ Education (Primary/Secondary/Tertiary)
❏ Government
❏ Industry
❏ Consultant/Self-Employed
❏ Other (please indicate) _____

2. Position

Please indicate your primary employment role. Choose one category and mark with a cross.

❏ Practitioner (e.g. clinical psychologist, industrial chemist)
❏ Information Services/Publishing
❏ Administration

❑ Teaching
❑ Student
❑ Research
❑ Consultant/Self-Employed
❑ Retired
❑ Other (please indicate) _____

3. Access to Technology

Please mark the pieces of information technology you have ready access to (either at home or at work). Mark all appropriate items with a cross.

❑ Personal Computer
❑ CD-ROM drive
❑ Sound output from computer
❑ Colour screen
❑ Direct network connection to Internet
❑ Modem connection to Internet
❑ Other (please indicate) _____

B. ELECTRONIC PUBLISHING

4. Familiarity

Please indicate how often you use the following forms of electronic publishing with the letters:

F - Frequently R - Regularly O - Occasionally S - Seldom N - Never:

	Subscribe to electronic publishing forums (listservers, BBS, netnews, etc.)
	Use FTP (File Transfer Protocol) to access materials
	Use Gopher to access materials
	Use World Wide Web to access materials
	Access materials on CD-ROM
	View electronic journal(s)
	View the electronic journal *Psyche*
	Publish electronically

5. Advantages and Disadvantages

Below are two possible lists of advantages or disadvantages of electronically published scholarly articles. You may add additional items in the spaces provided at the end of the lists.

Please indicate your feelings about each advantage or disadvantage to the left of the suggested text according to this scale:

SA - Strongly Agree, **A** - Agree, **N** - Neutral, **D** - Disagree, **SD** - Strongly Disagree.

	Advantages of electronic scholarly journals are:		**Disadvantages** ofelectronicscholarly journals are:
	Speed of publication		Poor quality
	24 hours a day access		Lack of refereeing
	Convenience The way they encourage feedback		Concerns about copyright Increased plagiarism
	Reduced paper consumption		Special skills needed to use
	Ease of searching		Special equipment needed to access
	Multimedia publications Affordability		Format is not reader friendly Communication costs to access them

C. FOLLOW-UP SURVEY

6. Do you wish to take part in a more detailed survey?

We are seeking individuals who may be asked to trial a number of new electronic and print publishing products. You would be asked a more detailed set of questions as part of this trial. If you are prepared to take part in this exercise, please provide your details below. This will be regarded as an expression of interest only, and does not guarantee that you will be selected.

Name: _____

Address:_____

_____

Phone: _____

Fax: _____

Email: _____

E. **ANY OTHER COMMENTS ?**

Based in part on a proposed questionnaire designed by Philip McEldowney, University of Virginia. Last revised November 2, 1996. © Andrew Treloar (aet@deakin.edu.au), 1996.

11.3 Library Case-Study Questions

These are the questions used to structure the interviews with librarians as discussed in Chapter 8: Library Case Studies. Because this was a semi-structured interview process, not all the questions were used with each participant, and other issues were sometimes addressed as seemed appropriate. However, these questions do reflect the general direction of the interview.

Publishing

1. How do you define publishing?

2. Do you regard this is just the initial release of content or remaining available long-term?

3. Thinking about the long-term sustainability, is this just for existing materials, or allowing the constant addition of new materials?

Commercial sustainability

4. Is there an assumption of commercial sustainability later? Why?

5. What happens if the project becomes commercially viable? Will it be taken away as a result?

6. Will these functions of library be merged into some wider IT/Info Services body?

Role of libraries

7. What is a libraries 'core business'?

8. Is this sort of initiative part of that core business?

9. Can epublishing become part of a libraries core business?

Project origins

10. How did the project start?

11. What is its likely future?

12. Org. context- did it start top down or bottom up?

13. Did it grow and flourish outside mainstream of major strategic mgt of university?

14. Were its origins accidental or deliberate or serendipitous?

15. How does it fit into the universities overall strategic plan?

Critical Success Factors and SWOT

16. What do you regard as the Critical Success Factors (CSF) for Highwire?

17. How are these different to any project of this type?

18. What do you see as the strengths of Highwire?

19. What do you regard as the weaknesses of Highwire?

20. What do you see as the opportunities for Highwire?

21. What do you regard as the threats to Highwire?

22. If you had your time over again, what would you do differently?

Anything else I should have asked?

12 References

12.1 Reference List

(1992). University libraries and scholarly communication. Technical report, Association of Research Libraries. Available online at http://www.lib.virginia.edu/mellon/.

Adams, G. R. and Schvaneveldt, J. D. (1985). *Understanding Research Methods*. Longman, New York and London.

Agre, P. E. (1995a). Community and democracy. *The Network Observer*, 2(7).

Agre, P. E. (1995b). Designing genres for new media: Social, economic, and political contexts. *The Network Observer*, 2(11). Available online at http://communication.ucsd.edu/pagre/tno/november-1995.html#designing.

Agre, P. E. (1995c). Institutional circuitry: Thinking about the forms and uses of information. *Information Technology and Libraries*, 14(4):225–230.

Allen, G. and Skinner, C., editors (1991). *Handbook for Research Students in the Social Sciences*. The Falmer Press, London.

Anonymous (1998a). Jime aims. Available online. Available online at http://www-jime.open.ac.uk/aims.html.

Anonymous (1998b). Jime home page. Available online. Available online at http://www-jime.open.ac.uk/.

Appel, A. W. (1996). How to edit a journal by e-mail. *Journal of Scholarly Publishing*, 27(2):82–100.

ASIS1996 (1996). *Proceedings of ASIS Mid-Year Meeting - The Digital Revolution: Assessing the Impact on Business, Education, and Social Structures*. American Society for Information Science.

Atkinson, R. (1993). Networks, hypertext and academic information services: Some longer-range implications. *College & Research Libraries*, 54:199–215.

Bamford, H. (1972). A concept for applying computer technology to the publication of scientific journals. *Journal of the Washington Academy of Science*, 62:306–314.

Belanger, T. (1994). The materiality of the book: Another turn of the screw. In Sutton, B., editor, *Literary Texts in an Electronic Age: Scholarly Implications for Library Services - Proceedings of the 31st Annual Clinic on Library Applications of Data Processing*, pages 187–195. Graduate School of Library and Information Studies, UIUC.

Benton, M. (1993). *Life and Time*. Random House, Sydney.

Berge, Z. L. and Collins, M. P. (1996). IPCT Journal readership survey. *JASIS*, 47(9):701–710.

Beyer, J. (1978). Editorial policies and practices among leading journals in four scientific fields. *Sociological Quarterly*, 19:77.

Bickman, L. and Rog, D. J. (1998). *Handbook of Applied Social Research Methods*. Sage Publications, Thousand Oaks, CA.

Bickman, L., Rog, D. J., and Hedrick, T. E. (1998). *Applied Research Design: A Practical Approach*, chapter 1, pages 5–37. In [Bickman and Rog, 1998].

Biemer, P. N., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A., and Sudman, S., editors (1991). *Measurement errors in surveys*. John Wiley, New York.

Bijker, W. B. (1987). *The social construction of bakelite: towards a theory of invention*, pages 159–187. In [Bijker et al., 1987].

Bijker, W. B., Hughes, T. P., and Pinch, T. J. (1987). *The social construction of technological systems: New directions in the sociology and history of technology*. MIT Press, Cambridge, Mass.

Bishop, A. P. (1995). Scholarly journals on the net: A reader's assessment. *Library Trends*, 43(Spring):544–570.

Bolter, J. D. (1991). *Writing Space: the Computer, Hypertext and the History of Writing.* Lawrence Erlbaum Associates, Hillsdale, N. J. Also available with a hypertext supplement (order from publisher).

Boyce, P. B. (1996). A successful electronic scholarly journal from a small society. Available online at http://www.aas.org/ pboyce/epubs/icsu-art.html. Paper presented to the ICSU Press - UNESCO Expert Conference on Electronic Publishing in Science and its WG2 Workshop, Paris, France, 12-23 February, 1996.

Boyce, P. B., Owens, E., and Biemesderfer, C. (1996). Electronic publishing: Experience is telling us something. *Serials Review*. Forthcoming.

Braman, S. (1995). Scholarly publishing in the information economy. In [Okerson, 1995], pages 39–49.

Bray, T., Paoli, J., and Sperberg-McQueen, C. M. (1998). Extensible markup language (XML) 1.0. Available online at http://www.w3.org/TR/1998/REC-xml-19980210.

Brent, D. (1995). Stevan harnad's "subversive proposal": Kick-starting electronic scholarship - a summary and analysis. *The Information Society*, 11(4):275–283. Reprinted from Ejournal 5(1), June 1995.

Brown, J. S. and Duguid, P. (1995). The social life of documents. *RELease 1.0*, 95(10):1–19.

Available online at http://www.firstmonday.dk/issues/issue1/documents/index.html.

Buckingham-Shum, S. (1998). Personal communication.

Bush, V. (1945). As we may think. *The Atlantic Monthly*, (176). Available online at http://www.isg.sfu.ca/ duchier/misc/vbush/.

Butler, B. (1986). Scholarly journals, electronic publishing, and library networks: from 1986 to 2000. *Serials Review*, 12:47–52.

Butler, H. J. (1994). The electronic journal: A viable channel for formal scholarly communication? In Andersen, D. L., editor, *Navigating the Networks: Proceedings of the American Society for Information Science mid-year meeting, Portland, Oregon, May 21-25, 1994*, pages 58–70, Medford, NJ. ASIS, Learned Information, Inc.

Butler, H. J. (1995a). Where does scholarly publishing get you? *Journal of Scholarly Publishing*, 26(4):234–246.

Butler, J. (1995b). Research into the reward system of scholarship; where does scholarly electronic publishing get you? In [Okerson, 1995], pages 167–177.

Campbell, R. M., Strange, J. D., and Bannerman, J. (1997). The 'modern' journal goes online. *British Journal of Urology*, 80:Suppl. 3, pp.36–45.

Carley, K. (1995). Communication technologies and their effect on cultural homogeneity, consensus and the diffusion of new ideas. *Sociological Perspectives*, 38(4):547–573.

Coffman, K. G. and Odlyzko, A. (1998). The size and growth rate of the Internet. Available from the author's home page at http://www.research.att.com/~amo/.

Collins, M. P. and Berge, Z. L. (1994). IPCT Journal: A case study of an electronic journal on the internet. *JASIS*, 45(10):771–776.

Conway, P. (1996). Preservation in the digital world. Technical report, Commission on Preservation and Access. Available online at http://www.clir.org/cpa/reports/conway2/.

Covi, L. (1996). Social worlds of knowledge work: How researchers appropriate digital libraries for scholarly communication. In [ASIS1996, 1996]. Available online at http://www.ics.uci.edu/~covi/social-worlds-asis.html.

Covi, L. and Kling, R. (1996). Organizational dimensions of effective digital library use: Closed rational and open natural systems models. *JASIS*, 47(9):672–689.

Crane, D. (1972). *Invisible colleges: diffusion of knowledge in scientific communities*. Chicago University Press, Chicago, Illinois.

Cronin, B. and Overfelt, K. (1995). E-journals and tenure. *JASIS*, 46(9):700–703.

Darwin, C. (1859). *The Origin of Species*. John Murray. Facsimile editionn, E. Mayr (ed.), Harvard University Press, 1964.

Day, C. (1995). Pricing electronic products. In [Okerson, 1995], pages 51–56.

Denning, P. J. (1995). Plagiarism in the web. *CACM*, 38(12):29.

Dillon, A. (1988). Reading from paper versus reading from screen. *Computer Journal*, 31(5):457–64.

Durbridge, N. H. and Stratfold, M. P. (1996). Varying the texture: A study of art, learning and multimedia. *Journal of Interactive Multimedia in Education*, (1). Available online at http://www-jime.open.ac.uk/96/1/durbridge-t.html.

Dyson, E. (1995). Intellectual value. *Wired*, 3(7). Available online at http:// www.hotwired.com/wired/3.07/features/dyson.html.

Eisenhardt, K. M. and Tabrizi, B. N. (1995). Accelerating adaptive processes: Product innovation in the global computer industry. *Administrative Science Quarterly*, (40):84–11.

Eisenstein, E. (1979). *The printing press as an agent of change: communications and cultural transformations in early modern Europe*, volume I-II. Cambridge University Press.

Eisenstein, E. (1983). *The printing revolution in early modern Europe*. Cambridge University Press, Cambridge, U.K.

Eldredge, N. and Gould, S. J. (1972). *Punctuated Equilibrium: an alternative to phyletic gradualism*, pages 82–115. Freeman, Cooper and Company, San Francisco.

Elsevier Science (1996). *TULIP Final Report*. Elsevier Science.

Engelbart, D. and English, W. K. (1968). A research center for augmenting human intellect. In *Proceedings of the Fall Joint Computer Conference*, volume 33(1). ASIPS Press. Role = Author Year = 1968 CrossRef = FJCC1968.

Entlich, R. (1994). Networked delivery of full-text journals: Diverse options, shared limitations. In Bishop, A. P., editor, *Emerging communities: Integrating networked information into library services*, pages 241–252, Urbana Champaign, IL. Graduate School of Library and Information Science, University of Illinois.

Entlich, R. (1995). Electronic chemistry journals: Elemental concerns. *The Serials Librarian*, 25(3/4):111–123.

Entlich, R., Garson, L., Lesk, M., Normore, L., Olsen, J., and Weibel, S. (1996). Testing a digital library: User response to the core project. *Library Hi Tech*, 14(4):99–118.

Fjällbrant, N. (1997). Scholarly communication - historical development and new possibili-

ties. In *Scholarly Communication in Focus: Proceedings of the 1997 International Association of Technology University Libraries (IATUL)*, Trondheim. International Association of Technology University Libraries (IATUL).

Foster, J. (1994). A status report on networked information retrieval: Tools and groups. Technical report, Joint IETF/RARE/CNI Networked Information Retrieval Working Group.

Fox-Wolfgramm, S., Boal, K., and Hunt, J. G. (1998). Organizational adaptation to institutional change: a comparative study of first-order change in prospector and defender banks. *Administrative Science Quarterly*, 43(1):87–127.

Fuller, S. (1995a). Cybermaterialism, or why there is no free lunch in cyberspace. *The Information Society*, 11(4):325–332.

Fuller, S. (1995b). Cyberplatonism: An inadequate constitution for the republic of science. *The Information Society*, 11(4):293–303.

Furuta, R. (1995). Documents in the digital library. Technical report, Center for the Study of Digital Libraries, Texas A&M University. Available online at http://www.cs-dl.tamu.edu/csdl/rickpapers/isdl95/isdl95.ps.

Furuta, R. and Marshall, C. C. (1995). Genra as reflection of technology in the world-wide web. In Fraïssé, S., editor, *Hypermedia Design, Proceedings of the Internatonal Workshop on Hypermedia Design (IWDH'95), Montpellier, France*, Workshops in Computing, pages 182–195. Springer Verlag. Availalbe online at http://cs-dl.tamu.edu/csdl/rickpapers/iwhd95/genre.ps.

Gaines, B. R. (1993). An agenda for digital journals: the socio-technical infrastructure of knowledge dissemination. *Journal of Organizational Computing*, 3(2):135–193.

Garvey, W. D. and Griffith, B. C. (1971). Scientific communication: Its role in the conduct of researcy and the creation of knowledge. *American Psychologist*, 26:349–362.

Gaver, W. (1991). Technology affordances. In Robertson, S. P., editor, *Reaching through technology: CHI '91 Conference proceedings: April 27-May 2, 1991, New Orleans, Louisiana*, pages 79–84, New York. ACM, Addison-Wesley.

Gaver, W. (1992). The affordances of media spaces for collaboration. In Turner, J., editor, *CSCW '92 Sharing perspectives : proceedings of the Conference on Computer-Supported Cooperative Work : October 31- November 4, 1992 Toronto, Canada*, New York, N.Y. ACM, Addison-Wesley.

Getz, M. (1997). An economic perspective on e-publishing in academia. *Journal of Electronic Publishing*, 3(1). Available online at http://www.press.umich.edu/jep/03-01/getz.html.

Gibson, J. J., editor (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, New York.

Ginsparg, P. (1994). First steps towards electronic research communication. *Computers in Physics*, 8(4).

Goldfarb, C. F. (1991). *The SGML handbook*. Oxford University Press, New York.

Goldfarb, C. F. (1997). SGML: The reason why and the first published hint. *JASIS*, 48(7):656–661.

Goldhaber, M. H. (1997). The attention economy: The natural economy of the net. *First-Monday*, 2(4). Available online at http://www.firstmonday.dk/issues/issue2_4/goldhaber/index.html.

Gotsch, C. H. and Reich, V. (1997). Electronic publishing of scientific journals: Effects on users, publishers and libraries. Available online. Available online at http://www-sul.stanford.edu/staff/pubs/clir/Final.html.

Gould, J. D. and et. al. (1987). Reading from CRT displays can be as fast as reading from paper. *Human Factors*, 29(5):497–517.

Gould, S. J. (1989). *Wonderful life : the Burgess Shale and the nature of history*. W. Norton and Company, New York.

Gould, S. J. (1991). *Ever Since Darwin*. Penguin.

Gould, S. J. (1994). The evolution of life on earth. *Scientific American*, 271(4):63–69.

Gould, S. J. and Eldredge, N. (1993). Punctuated equilibrium comes of age. *Nature*, 366:223–227.

Grabinger, R. S., Dunlap, J. C., and Jonassen, D. H. (1992a). Hypertext links. *Performance and Instruction*, pages 36–49.

Grabinger, R. S., Dunlap, J. C., and Jonassen, D. H. (1992b). Sequential links. *Performance and Instruction*, pages 46–49.

Grabinger, R. S., Dunlap, J. C., and Jonassen, D. H. (1993a). Relational links. *Performance and Instruction*, pages 35–40.

Grabinger, R. S., Dunlap, J. C., and Jonassen, D. H. (1993b). Support links. *Performance and Instruction*, pages 33–39.

Graham, G. (1992). The relationship between publishers and academics. *Journal of Scholarly Publishing*, 24(1):13–23.

Graham, P. S. (1997). Building the digital research library: Preservation and access at the heart of scholarship. Available online at http://www.ukoln.ac.uk/services/papers/follett/graham/paper.html. 1997 Follett Lecture, delivered March 19, 1997.

Gruman, G. (1995). The paper chase. *Macworld*, pages 126–131.

Guédon, J.-C. (1995). Research libraries and electronic scholarly journals: Challenges or opportunities? *The Serials Librarian*, 26(3/4):1–20.

Guédon, J.-C. (1996). *The Seminar, the Encyclopedia, and the Eco-Museum as Possible Future Forms of Electronic Publishing*, pages 71–89. In [Peek and Newby, 1996].

Hård, M. (1994). *Machines are frozen spirit: the scientification of refrigeration and brewing in the 19th century - a Weberian interpretation*. Campus, Frankfurt am Main.

Halfhill, T. (1997). Good-bye, GUI, hello, NUI. *Byte*, 22(7):60–72.

Harnad, S. (1990). Scholarly skywriting and the prepublication continuum of scientific inquiry. *Psychological Science*, (1):342–344.

Harnad, S. (1991). Post-Gutenberg galaxy: The fourth revolution in the means of production of knowledge. *The Public-Access Computer Systems Review*, 2(1):39–53.

Harnad, S. (1995a). Electronic scholarly publication: Quo vadis? *Serials Review*, 21(1):78–80.

Harnad, S. (1995b). *Implementing Peer Review on the Net: Scientific Quality Control in Scholarly Electronic Journals*, chapter 6, pages 103–118. In [Peek and Newby, 1996].

Harnad, S. (1995c). The postgutenberg galaxy: How to get there from here. *Times Higher Education Supplement*.

Harnad, S. (1995d). Sorting the esoteric from the exoterica: There's plenty of room in cyberspace. a response to fuller. *The Information Society*, 11(4):305–324.

Harter, S. P. (1996). The impact of electronic journals on scholarly communication: A citation analysis. *The Public-Access Computer Systems Review*, 7(5). Available online at http://info.lib.uh.edu/pr/v7/n5/hart7n5.html.

Harter, S. P. (1998). Scholarly communication and electronic journals: An impact study. *JASIS*, 49(6):507–516.

Harter, S. P. and Kim, H. J. (1996a). Accessing electronic journals and other e-publications: An empirical study. *College & Research Libraries*, 57(5):440–456.

Harter, S. P. and Kim, H. J. (1996b). Electronic journals and scholarly communication: A citation and reference study. In [ASIS1996, 1996]. Available online at http://www-slis.lib.indiana.edu/PrePrints/harter-asis96midyear.html.

Hayes, R. M. (1989). *Who Should be in Control?*, pages 25–33. Syracuse University Press, Syracuse, N. Y.

Hickey, T. B. (1995). Present and future capabilities of the online journal. *Library Trends*, 43(4):528–543.

Highwire Press (1997). About highwire press: Mission. Available online. Cited December 2,

1997. Available online at http://highwire.stanford.edu/about.shtml.

Hitchcock, S. (1996). *A survey of STM online journals 1990-95: the calm before the storm*. Association of Research Libraries, 6 edition. Available online at http://journals.ecs.soton.ac.uk/survey/survey.html.

Hitchcock, S., Carr, L., and Hall, W. (1997). Web journals publishing: a UK perspective. *Serials*, 10(3):285–299. Available online at http://journals.ecs.soton.ac.uk/uksg.htm.

Hunter, K. (1994). Issues and experiments in electronic publishing and dissemination. *Information Technology and Libraries*, 13(2):127–132.

Hunter, K. (1998). Electronic journal publishing: Observations from inside. *D-Lib Magazine*, 4(7). Available online at http://www.dlib.org/dlib/july98/07hunter.html.

Hurd, J. M. (1996). *Models of Scientific Communication Systems*, pages 65–76. ASIS Monograph Series. American Society for Information Science.

Jacobson, T. L. (1994). The electronic publishing revolution is not "global". *JASIS*, 45(10):745–752.

Johnson, J. and et. al. (1989). The xerox star: A retrospective. *IEEE Computer*, 22(9):11–29.

Jones, C. (1991a). *Qualitative Interviewing*, chapter 17, pages 203–214. In [Allen and Skinner, 1991].

Jones, C. (1991b). *Quantitative Research*, chapter 18, pages 215–224. In [Allen and Skinner, 1991].

Kasdorf, B. (1998). SGML and PDF - why we need both. *Journal of Electronic Publishing*, 3(4). Available online at http://www.press.umich.edu/jep/03-04/kasdorf.html.

Kaufer, D. S. and Carley, K. M. (1993). *Communication at a Distance: the influence of print on sociocultural organization and change*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Kaufer, D. S. and Carley, K. M. (1994). Some concepts and axioms about communication. *Written Communication*, 11(1):8–42.

Ketcham, L. and Born, K. (1996). Projecting the electronic revolution while budgeting for the status quo (36th annual report: Periodical price survey 1996). *Library Journal*, 121(7):45–52.

Keyhani, A. (1993). The online journal of current clinical trials: an innovation in electronic journal publishing. *Database*, 16(1):14–23.

Keyhani, A. (1995). Innovations in cost recovery. In [Okerson, 1995], pages 57–65.

Kiernan, V. (1998a). Publisher loosens rules of sharing electronic journals. *Chronicle of Higher Education*, page A20.

Kiernan, V. (1998b). University libraries join with chemical society to create a new, low-cost journal. *Chronicle of Higher Education*, page A20.

King, D. W. and Griffiths, J. (1995). Economic issues concerning electronic publishing and distribution of scholarly articles. *Library Trends*, 43:713–740.

Kling, R. and Covi, L. (1995). Electronic journals and legitimate media in the systems of scholarly communication. *The Information Society*, 11(4):261–271.

Kronick, D. A. (1976). *A History of Scientific & Technical Periodicals: The Origins and Development of the Scientific and Technical Press 1665-1790*. Scarecrow Press, Metuchen, N.J., 2 edition.

Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 2 edition.

Lancaster, F. W. (1995a). Attitudes in academia toward feasibility and desirability of networked scholarly publishing. *Library Trends*, 43(4):741–751.

Lancaster, F. W. (1995b). The evolution of electronic publishing. *Library Trends*, 43(4):519–527.

Landow, G. (1992). *Hypertext: The Convergence of Contemporary Critical Theory and Technology*. Johns Hopkins University Press.

Landow, G. P. (1989). The rhetoric of hypermedia: Some rules for authors. *Journal of Computing in Higher Education*, 1(1):39–64.

Lanham, R. (1994). The economics of attention. In *The Research Library the Day after Tomorrow: Proceedings of 124th Annual Meeting, Association of Research Librarians*, Austin, Texas. Association of Research Librarians, Association of Research Librarians. Available online at http://sunsite.berkeley.edu/ARL/Proceedings/124/ps2econ.html.

Lanham, R. A. (1993). *The Electronic Word: Democracy, Technology and the Arts*. University of Chicago Press, Chicago.

Lesk, M. (1992). Preservation of new technology: A report of the technology assessment advisory committee. Technical report, Commission on Preservation and Access, Washington, D.C. Available online at http://www.clir.org/cpa/reports/lesk2.html.

Levy, D. M. (1994). Fixed or fluid? document stability and new media. In Anonymous, editor, *Proceedings of the European Conference on Hypertext Technology '94*, Edinburgh.

Lewenstein, B. V. (1995). Do public electronic bulletin boards help create scientific knowledge? the cold fusion case. *Science, Technology, & Human Values*, 20(2):123–

149.

Lievrouw, L. A. (1992). Communication, representation, and scientific knowledge: A conceptual framework and case study. *Knowledge and policy*, 5(1):6–28.

Lievrouw, L. A. and Carley, K. (1990). Changing patterns of communication among scientists in an era of "telescience". *Technology in society*, 12(4).

Loch, C. H. and Huberman, B. A. (1998). A punctuated equilibrium model of technology diffusion. *Journal of Management Science*. To appear. Available online at http://www.parc.xerox.com/spl/groups/dynamics/www/punctuated.html.

Lyman, P. (1995). Copyright and fair use in the digital age: Q&A with Peter Lyman. *Educom Review*. 30(1):32-35. Available online at http://www.educause.edu/pub/er/review/reviewarticles/30132.html.

Lynch, C. A. (1994a). The integrity of digital information: Mechanics and definitional issues. *JASIS*, 45(10):737–744.

Lynch, C. A. (1994b). Scholarly communication in the networked environment: Reconsidering economics and organizational missions. *Serials Review*, 20(3):23–30.

Magrini, T. (1995). Ballad and gender: Reconsidering narrative singing in Northern Italy. *Ethnomusicology Online*, (1). Available online at http://research.umbc.edu.ca/eol/magrini/magrini.html.

Mangione, T. W. (1998). *Mail Surveys*, chapter 14, pages 399–427. In [Bickman and Rog, 1998].

Marcoux, Y. and Sévigny, M. (1997). Why SGML? why now? *JASIS*, 48(7):584–592.

Marks, R. H. (1995). The economic challenges of publishing electronic journals. *Serials Review*, 21(1):85–88.

Marshall, C. (1998). Toward an ecology of hypertext annotation. In *Proceedings of ACM Hypertext '98, Pittsburgh, PA (June 20-24, 1998)*, pages 40–49. ACM. Winner of 1998 Engelbart Best Paper Award.

Mason, J. D. (1997). SGML and related standards: New directions as the second decade begins. *JASIS*, 48(7):593–596.

McKie, D. (1979). *The Scientific Periodical from 1665 to 1798*, chapter 1, pages 7–17. In [Meadows, 1979].

McKnight, C., Meadows, J., Pullinger, D., and Rowland, F. (1994). ELVYN - publisher and library working towards the electronic distribution and use of journals. In Schnase, J. L., editor, *Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*. Available online at http://www.csdl.tamu.edu/DL94/paper/mcknight.html.

McMillan, G. (1995). Scholarly communications project: Publishers and libraries. In [Okerson, 1995], pages 135–145.

McNeilly, K. (1995). Ugly beauty: John Zorn and the politics of postmodern music. *Postmodern Culture*, 5(2). Available online at http://muse.jhu.edu/journals/postmodern_culture/v005/5.2mcneilly.html.

Meadows, A. J. (1977). The problem of refereeing. *Scienta*, 112:787–794.

Meadows, A. J., editor (1979). *The Scientific Journal*. ASLIB, London.

Meadows, J., Pullinger, D., and Such, P. (1995). The cost of implementing an electronic journal. *Journal of Scholarly Publishing*, 26:167–173.

Metz, P. (1995). Revolutionary change in scholarly and scientific communications: The view from a university library. *Change*, 27(1):29–33.

Metz, P. and Gherman, P. M. (1991). Serials pricing and the role of the electronic journal. *College & Research Libraries*, 52:315–327.

Moore, G. (1997). Moore says Moore's Law to hit wall. *CNET Computing*. Available online at http://www.dlib.org/dlib/july98/07hunter.html.

Morton, B. (1997). Is the journal as we know it an article of faith? An open letter to the faculty. *PACS-R*, 8(2). Available online at http://info.lib.uh.edu/pr/v8/n2/mort8n2.html.

Mulvaney, J. and Steele, C., editors (1993). *Changes in Scholarly Communication Patterns: Australia and the Electronic Library*, Australian Academy of the Humanities Occasional Paper NO. 15, Canberra. Australia. Australian Academy of the Humanities.

Muse, P. (1997). Muse - general introduction. Available online. Available online at http://muse.jhu.edu/proj_descrip/gen_intro.html.

Mustard, J. (1994). Scientific visualization in earth and planetary science: It looks good but will it publish. *The Serials Librarian*, 24(3/4):3–7.

Nadasdy, Z. (1997). A truly all-electronic journal: Let democracy replace peer review. *Journal of Electronic Publishing*, 3(1). Available online at http://www.press.umich.edu/jep/03-01/EJCBS.html.

Nelson, T. H. (1965). A file structure for the complex, the changing and the indeterminate. In *Association for Computing Machinery: Proceedings of the 20th National Conference*, pages 96–98, New York. Association for Computing Machinery.

Newman, M. (1997). The Highwire Press at Stanford University: A review of current features. *Issues in Science and Technology Librarianship*, Summer. Available online at http://www.library.ucsb.edu/istl/97-summer/journals.html.

Nunberg, G. (1993). The places of books in the age of electronic reproduction. *Representations*, (42):13–37.

Odlyzko, A. (1994). Tragic loss or good riddance? the impending demise of traditional scholarly journals. Full version of paper reprinted in Peek and Newby (Odlyzko1995). Citations will be to this latter version.0.

Odlyzko, A. (1997). Silicon dreams and silicon bricks: the continuing evolution of libraries. Available from the author's home page at http://www.research.att.com/~amo/.

O'Donnell, M. J. (1995). Electronic journals: Scholarly invariants in a changing medium. *Journal of Electronic Publishing*.

Okerson, A. S. (1991). Back to academia? the case for american universities to publish their own research. *Logos*, 2(2):106–112. Available online at http://www.yale.edu/~okerson/case.html.

Okerson, A. S., editor (1995). *Filling the Pipeline and Paying the Piper: Proceedings of the Fourth Symposium*. ARL, ARL Press.

Okerson, A. S. (1996). *University Libraries and Scholarly Communication*, pages 181–199. In [Peek and Newby, 1996].

Okerson, A. S. and O'Donnell, J. J., editors (1995). *Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing*. Office of Scientific and Academic Publishing, Association of Research Libraries.

Olsen, J. (1994). *Electronic Journal Literature: Implications for Scholars*. Mecklermedia Corporation, Westport, CT.

O'Reilly, T. (1995). Publishing models for internet commerce. *OnTheInternet*, 1(4):29–35.

Pantic, D. (1998). Radio B92 and its internet project, opennet: From local to global media. In Smith, J. W. T., editor, *Proceedings of the Second ICCC & IFIP Electronic Publishing Conference*, Canterbury. ICCC/IFIP.

Parang, E. and Saunders, L. (1994b). Electronic journals in ARL libraries: Issues and trends (SPEC flyer 202). Association of Research Libraries Systems and Procedures Exchange Center.

Parang, E. and Saunders, L. (1994a). Electronic journals in ARL libraries: Policies and procedures (SPEC flyer 201). Association of Research Libraries Systems and Procedures Exchange Center.

Peek, R. P. and Newby, G. B., editors (1996). *Scholarly Publishing: The Electronic Frontier*. MIT Press, Cambridge, Massachusetts.

Peters, J. (1995). The hundred years war started today : an exploration of electronic peer review. *Internet Research*, 5(4):3–9.

Price, D. J. d. S. (1963). *Little Science, Big Science*. Columbia University Press, New York.

Rayward, W. B. (1994). Visions of Xanadu: Paul Otlet (1868-1944) and hypertext. *JASIS*, 45(4):235–250.

Rayward, W. B. (1997). The origins of information science and the International Institute of Bibliography/International Federation for Information and Documentation (FID). *JASIS*, 48(4):289–300.

Reich, V. (1997). Personal communication. Via email.

Roistacher, R. C. (1978). The virtual journal. *Computer Networks*, 2:18–74.

Rose, H. (1991). *Case Studies*, chapter 16, pages 190–202. In [Allen and Skinner, 1991].

Rose, S. (1993). *The Making of Memory*. Bantam Books.

Rothenburg, J. (1995). Ensuring the longevity of digital documents. *Scientific American*, 272(1):24–29.

Rousseau, R. and Hooydonk, G. V. (1996). Journal production and journal impact factors. *JASIS*, 47(10):775–780.

Rowland, F. (1994). Electronic journals: Neither free or easy. *Ejournal*, 4(2). Available online at http://www.hanover.edu/philos/ejournal/archive/ej-4-2.txt.

Rowland, F. (1996). Re: Future of electronic publishing–reply. Posting to hyper-journal mailing list, 13/2/1996. Available online at http://www.mailbase.ac.uk/lists/hyperjournal-forum/1996-02/0024.html.

Rowland, F., McKnight, C., and Meadows, J., editors (1995). *Project Elvyn: An Experiment in Electronic Journal Delivery. Facts, Figures and Findings.* Bowker-Saur.

Rubenstein, E. (1995). Punctuated equilibrium in scientific publishing (several science journals change top editors). *Science*, 268(5216):1415.

Schaffner, A. C. (1994). The future of scientific journals: Lessons from the past. *Information Technology and Libraries*, 13(4):239–247.

Schamber, L. (1996). What is a document? rethinking the concept in uneasy times. *JASIS*, 47(9):669–671.

Schauder, D. (1994a). *Electronic Publishing of Professional Articles: Attitudes of Academics and Implications for the Scholarly Communication Industry*. PhD thesis, Institute of Education, University of Melbourne.

Schauder, D. (1994b). Electronic publishing of professional articles: Attitudes of academics and implications for the scholarly communication industry. *JASIS*, 45(2):73–100.

Scholarly Communications Project (1997). About Scholarly Communications Project: VPI&SU. Available online at http://scholar.lib.vt.edu/about/scpabout.html#his-

tory.

Senders, J. (1976). The scientific journal of the future. *The American Sociologist*, 11:160–164.

Senders, J. (1977). An on-line scientific journal. *Information Scientist*, 11(1):3–9.

Senders, J. W. (1980). I have seen the future and it doesn't work: the electronic journal experiment. In Langlois, E. G., editor, *Scholarly Publishing in an Era of Change: Proceedings of the Second Annual Meeting, Society for Scholarly Publishing*, Minneapolis, Minnesota. Society for Scholarly Publishing.

Singleton, A. (1994). The scientific journal: Present trends and likely futures - their impact on library and information services. *Information UK Outlooks*, 7:1–30.

Smith, P. N., Brailsford, D. F., Evans, D. R., Harrison, L., Probets, S. G., and Sutton, P. E. (1993). Journal publishing with acrobat: the CAJUN project. *Electronic Publishing–Origination, Dissemination and Design*, 6(4):481–493.

Snyder, I. (1996). *Hypertext: The Electronic Labyrinth*. Melbourne University Press.

Spender, D. (1992). Performing academics. *Australian Author*, 24(2):18–22.

Spink, A., Robins, D., and Schamber, L. (1998). Use of scholarly book reviews: Implications for electronic publishing and scholarly communication. *JASIS*, 49(4):364–374.

Stewart, L. (1996). User acceptance of electronic journals: Interviews with chemists at Cornell university. *College & Research Libraries*, 57(4):339–349.

Stodolsky, D. S. (1995). Consensus journals: Invitational journals based upon peer review. *The Information Society*, 11(4):247–260.

Task Force on Archiving of Digital Information (1996). Preserving digital information: Final report and recommendations. Available online at http://www.rlg.org/ArchTF/.

Thompson, J. C. (1988). Journal costs: Perception and reality in the dialogue. *College & Research Libraries*, 49:481–482.

Treloar, A. (1994). Working in a worldwide invisible college: the internet and the global scholar. In *Proceedings of CAUSE '94*, Melbourne, Australia. CAUL/CAUDIT. Available online at http://www.deakin.edu.au/~aet/Research/Publications/CAUSE94/.

Treloar, A. (1996). Better than print? Hypermedia scholarly publishing and the world wide web. In Anonymous, editor, *Proceedings of the Victorian Association for Library Automation Bi-ennial Conference (VALA '96)*. Victorian Association for Library Automation. Available online at http://www.deakin.edu.au/~aet/Research/Publications/VALA96/.

Trott, B. (1998). Microsoft showcases new text technology. *Infoworld Electric*. Available

online at http://www.infoworld.com/cgi-bin/displayStory.pl?981114.wht-ext.htm.

Turoff, M. and Hiltz, S. (1982). The electronic journal: a progress report. *JASIS*, 33:195–202.

Valauskas, E. J. (1994). Paper-based or digital text: What's best? *Computers in Libraries*, 14(1):44–47.

Valauskas, E. J. (1997). Waiting for Thomas Kuhn: First Monday and the evolution of electronic journals. *FirstMonday*, 2(12). Available online at http://www.firstmonday.dk/issues/issues2_12/valauskas/.

Varian, H. R. (1996). Pricing electronic journals. *D-Lib Magazine*, 2(6).

Wegener, A. (1962). *The Origin of Continents and Oceans*. Dover Publishing, New York, 4 edition. Translated from the fourth revised German edition by J. Biram. Dover Pub.; New York.

Weinberg, B. H. (1997). The earliest hebrew citation indexes. *JASIS*, 48(4):318–300.

Wilkinson, R. and Robinshaw, H. W. (1987). Proof-reading: VDU and paper text compared for speed, accuracy and fatigue. *Behaviour and Information Technology*, 6(2):125–33.

Woolfrey, S. (1995). Subscriber interest in electronic copy of printed journals. *Journal of Scholarly Publishing*.

Wyly, B. J. (1998). Competition in scholarly publishing? what publisher profits reveal. *ARL Newsletter*. Available online at http://www.arl.org/newsltr/200/wyly.html.

Yates, S. J. and Sumner, T. R. (1997). Digital genres and the new burden of fixity. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS '97), Special Track on Genres in Digital Documents, Maui, HA*, pages Vol. VI, pp. 3–12. Available online at http://kmi.open.ac.uk/~tamara/articles/trs-97-HICSS.pdf.

Yin, R. K. (1998). *The Abridged Version of Case Study Research*, chapter 8, pages 229–259. In [Bickman and Rog, 1998].

Zuckerman, H. and Merton, R. K. (1979). *Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System*, pages 112–146. In [Meadows, 1979].