Authors :	Bo-Christer Bjork, Ziga Turk
Title:	How Scientists Retrieve Publications: An Empirical Study of How the Internet Is Overtaking Paper Media
Availability:	This text may be freely searched and displayed. Permission must be received for subsequent distribution in print or electronically. Please contact spo-help@umich.edu for more information.
Source:	*How Scientists Retrieve Publications: An Empirical Study of How the Internet Is Overtaking Paper Media* Bo-Christer Bjork, Ziga Turk vol. 6, no. 2
DOI:	http://dx.doi.org/10.3998/3336451.0006.202

How Scientists Retrieve Publications: An Empirical Study of How the Internet Is Overtaking Paper Media

Bo-Christer Björk and Ziga Turk

The current mainstream scientific-publication process has so far been only marginally affected by the possibilities offered by the Internet, despite some pioneering attempts with free electronic-only journals and electronic preprint archives. Additional electronic versions of traditional paper journals for which one needs a subscription are not a solution. A clear trend, for young researchers in particular, is to go around subscription barriers (both for paper and electronic material) and rely almost exclusively on what they can find free on the Internet, which often includes working versions posted on the home pages of the authors.

A survey of how scientists retrieve publications was conducted in February 2000, aimed at measuring to what extent the opportunities offered by the Internet are already changing the scientific information exchange and how researchers feel about this. This paper presents the results based on 236 replies to an extensive Web-based questionnaire, which was announced to around 3,000 researchers in the domains of construction information technology and construction management.

The questions dealt with how researchers find, access, and read different sources; how many and what publications they read; how often and to which conferences they travel; how much they publish, and criteria for where they eventually decide to publish. Some of the questions confronted traditional and electronic publishing, with one final section dedicated to opinions about electronic publishing.

According to the survey, researchers already download half of the material that they read digitally from the Web. The most popular method for retrieving an interesting publication is downloading it for free from the author's or publisher's Web site. Researchers are not particularly willing to pay for electronic scientific publications. There is much support for a scenario of electronic journals available freely in their entirety on the Web, where the costs could be covered by, for instance, professional societies or the publishing university.

Introduction

The efficiency of the scientific publishing process is an important issue for society as a whole. Research and development spending is in most industrialised countries at least 3% of Gross Domestic Product and the indirect effects of research and development in all sectors of society

are evident. Nevertheless, there has been relatively little R&D work looking at how information dissemination works in this domain. The mechanisms for scientific publishing were shaped in the last two centuries and have so far remained almost unchanged by the Internet revolution. The academic community seems to have great difficulty changing its attitudes and the way the publication system works.

Scientists spend 50-60% of their time communicating, and any innovations that facilitate their communication with peers and with industry have far-reaching benefits. An excellent example on how free access to R&D information can foster progress is the Internet itself. All Internet standards and a large majority of research papers have been available free on the Internet rather than being sold by a standards organisation or through costly subscriptions to paper-based journals.

An important part of the scientific-communication process consists of reporting on research results in the form of drafts, conference papers, and peer-reviewed scholarly articles. On average, academics read more than 100 such publications per year. This also means that they spend weeks every working year searching for and retrieving relevant information. There is also a "cost" of opportunities lost because they may overlook important publications because of not being alerted of their existence, or because it is too tedious or costly to get hold of them.

"As readers, scientists are very progressive and supportive of free electronic journals, but as authors they are very conservative in their choice of where to submit their manuscripts"

Up until a few years ago, the only mechanism for disseminating these papers was in the form of paper publication. Since the 1990s the Internet has changed the situation dramatically. Currently there are a number of options available:

- The traditional paper-based journals and conference proceedings
- The electronic copies of the above, still accessible only through subscription or pay per view
- Electronic-only refereed journals that are free of charge
- Haphazard copies of all kinds of material on the home pages of authors or their institutions
- A handful electronic pre-prints archives.

Of these options, the traditional paper journals and conference proceedings still account for the majority of publications retrieved, but the downloading of free, "haphazard material" is quickly catching on. The big drawback with this is that such material usually escapes efficient indexing and is therefore mostly retrieved as a result of a citation elsewhere, thus side-stepping paying for the "official versions." The market share of the free electronic journals is still very small (a 1999 study found some 400 such journals), (Wells 1999) whereas the total number of scholarly journals has been estimated to be almost 100,000 (Tenopir and King 2000). There are only a handful of successful free electronic preprint archives.

In the world that preceded word processing on personal computers and the communication channels offered by the Internet, commercial scientific publishers obtained their position in the publication life-cycle chain because they provided a clear value-added service. Scientists provided the content, editing, review, and quality control free. Publishers typeset the papers, printed them, and distributed them. Today this is not a problem any more: scientists lay out their papers themselves, while the Internet offers a marketing and distribution channel operating at almost zero marginal cost.

As publishers of a free electronic refereed journal, the *Electronic Journal of Information Technology in Construction,* we have followed the general development of scientific publishing for some five years and have experienced both the euphoria of taking a small part in achieving this huge paradigm shift as well as the disappointment over the reluctance of many of our colleagues to change their publishing habits. It was, in fact, the apparent contradiction of the readers' preferences for the free Internet journals and the authors' habits to submit to the conservative, but highly rated paper-based journals, that inspired us to investigate a bit more thoroughly, using empirical methods, how our peers actually identify and retrieve the publications that they read.

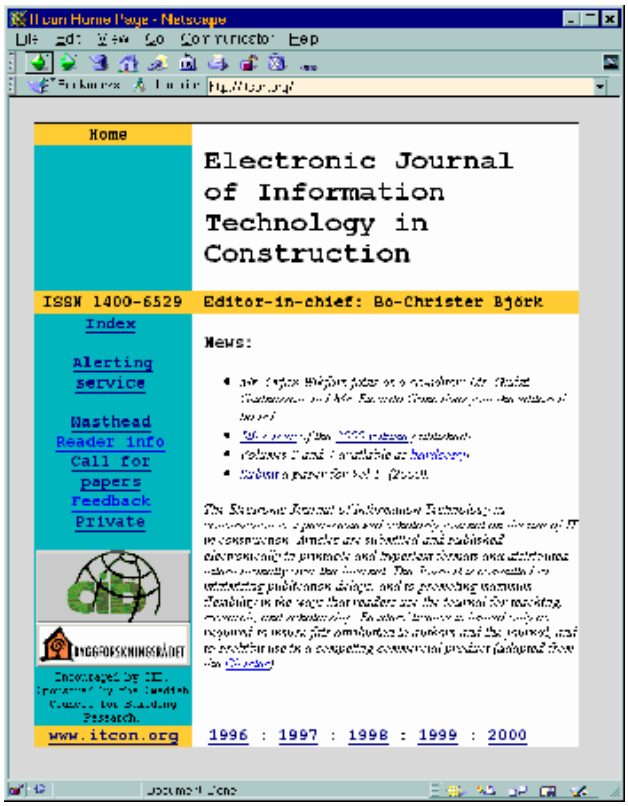


Figure 1: The *Electronic Journal of Information Technology in Construction,* an example of an electronic-only journal.

The general situation of scientific publishing

For a number of years commercial publishers have been forced to raise the subscription prices to periodic journals faster than inflation which has tended to reduce the number of subscribers. This vicious circle has been named the "periodicals crisis" (Walker 1998). At the same time, a number of idealistic researchers have founded electronic journals that are available free on the Internet. Wells (1999) found 387 free electronic scientific journals in her research. She also found that 25% of the journals that at one stage or another had advertised in the NewJour Web site could be considered dead. This high mortality rate is quite typical, since new journals have a hard time attracting authors without the backing of established commercial publishers or learned societies.

A survey done in 1998 by the International Council for Scientific and Technical Information on how researchers use Internet-based services showed that 61% of respondents felt that electronic journals or trade magazines are easy to use or user friendly, but that only 14% of respondents publish in such journals (Anon 1999). Thus it seems that as readers of journals, scientists are very progressive and supportive of free electronic journals, but as authors of articles they are very conservative in their choice of where to submit their manuscripts. In a discussion of the economics of electronic journals, Odlyzko (1998) suggests one future scenario: "scholars will continue submitting their papers to the most prestigious journals they can find, no matter how small their circulation, since prestige is what counts in tenure and promotion decisions, and since everybody that they want to read their papers will get them electronically from preprint servers in any case."

"This set of answers clearly shows the strong preference for the just-in-time search"

In the literature on scientific electronic publishing, there is clear disagreement on whether electronic journals are much cheaper to produce than paper-based ones. (For good discussions see King and Tenopir 1998, Odlyzko 1998). For a commercial publisher, the cost of running an electronic-only, subscription-based journal may not differ much from paper publishing. For the low subscription numbers typical for the majority of scientific journals, the fixed costs of marketing, editing, typesetting, general overhead, etc. are much higher than the variable cost of printing and distribution. If, on the other hand, one takes a broader view of a process, starting with an author preparing a draft manuscript to

another researcher retrieving and reading the same draft manuscript 10 years later, there are strong indications that the electronic process is much cheaper. An important issue here is that the costs to libraries for archiving journals and for readers of identifying and retrieving an article should also be taken into account. Those costs have been ignored in most of the cost estimates published.

All the major players in the scientific publishing process are currently faced with a quickly changing environment:

- Readers of scientific literature are facing the situation where most of the information they need to retrieve is only a few mouse-clicks away and available free. This has the side effect that the threshold to retrieve information that either involves extra effort (going to a university library), a long waiting time (ordering a paper copy through a traditional interlibrary service), or an extra cost (paying 20 US$ for downloading a digital copy of an article) is getting increasingly higher and that retrieving publications involving such efforts is done less and less. Students writing M.Sc. thesis in information-technology-related subjects almost exclusively use material available free on the Internet as references to their work. And Butterworth reports that Ph.D. students in his particle physics group "never, but never, look at a printed journal" (Butterworth 1998). For reasons like this, as well as constantly rising subscription prices, university departments and libraries are reducing the number of journals to which they subscribe. And even the scientists who have the luxury of digital access to most of their favourite journals have the extra chore of keeping track of numerous passwords and log-in routines.

- Authors are faced with the dilemma of choosing in which type of journal to publish their results. Should they prioritise journals with a fast turnaround rate from submission to publication, which can be achieved by electronic-only journals with no need to wait for scheduled issues? Especially in fast-moving fields such as any information-technology-related research, a waiting time of one year or more seems very prohibitive. Or should they prioritise a potentially large readership, which the major journals and freely available electronic journals can offer? Other important criteria include whether a journal is indexed by some well-known indexing services or whether the journal has a high academic status (important to help the author earn merits and in some cases even funds for his department).

- Publishers have to make choices about whether to publish full text on the Web in addition to the printed version, and if so, whether they should offer a screen-readable (HTML) version or a printable (PDF) version — or even both. They also need to decide how to make use of the possibilities offered by the medium in terms of hyperlinking, indexing, multi-media, e-mail alerting services, etc. The crucial question, however, is the pricing of the product. Is the price of the electronic version about the same as for the paper version? How should the site be protected from non-paying customers — by individual password protection or site licensing to universities? Another issue is whether to allow only yearly subscribers access to the full text or to also set up a mechanism to sell individual articles digitally, possibly using a specialised third party service.

- Libraries will see their role as intermediaries change radically in the future. The need to have paper copies on display and archived will diminish radically. Maybe a more important function will be to negotiate for the electronic site licenses of the university. Currently the question of long-term digital archiving is receiving increasing attention (Hodge and Carroll 1999). Studies show that already in the traditional paper-based process, the total archiving costs for journals (the sum for all libraries worldwide archiving a particular journal) are greater than the original publication costs of the publisher (Odlyzko 1998). With electronic archiving these costs can be reduced radically since the need for duplicating the archiving effort locally disappears. One obvious future development would be that publishers increasingly start to assume the archival responsibility for their journals. This, however, entails a lot of unsolved issues and risks. What happens when a university stops subscribing to an electronic journal? Does it still have the right of access to those volumes for which they already paid? (In the paper

process they *would* retain physical possession of the issues.) What happens when a publisher goes out of business or disappears in a merger? What guarantees are there that the old digital issues will be stored at all?

The survey: method and response

The survey takes a look at the current situation in the two areas of Construction Information Technology and Construction Management and Economics. The choice of construction information technology was obvious, given the background of the authors as researchers and publishers of an electronic journal in the domain (*Electronic Journal of Information Technology in Construction*). Construction management and economics has been included due to the fact that there is considerable overlap and synergy between these two research communities, on the personal level, in conference attendance, and in the choice of journals in which to publish.

The survey was conducted in February 2000. The purpose was to empirically survey the current situation, as well as to identify recent trends in usage and opinions. The research technique used has been a general Web questionnaire to authors and readers combined with case studies of individual journals and conferences.

The publication and reading of scientific journal articles cannot be studied in isolation since it belongs to the larger context of how researchers acquire the data and knowledge that is used as input into their own research. The survey was thus designed to include a comprehensive overview of the scientific information exchange, reading, and publishing habits among the researchers. For this reason other means of scientific communication — ranging from attending conference presentations to informal e-mail exchange and visits to construction sites — have also been included in the survey.

The survey tries to look at the scientific-publishing process in one particular scientific community as a whole rather than concentrating on the case study of one single journal or looking at scientific publishing in general. The survey did not target the information exchange between research and practice. This is a complex question that would well merit a separate research project. This concern was present in some comments made by the respondents. For example:

I have been generally satisfied with the readership and academic acceptance of my publications, but I am bitterly disappointed in how little my work has affected practice. It is also a little disappointing that your survey covers authorship and readership, but not one question deals with practical impact of authors' papers. Are we really just publishing for publishing's (and promotion's) sake?

The questions (see Björk and Turk, 2000) were grouped into the following sets:

- Questions about personal information to establish age, job, function, and research topic of the respondent.
- Questions on how people find, access, and read information sources.
- Questions on which journals people follow, what conferences they attend, and how often.
- Questions related to publishing their work, how much they publish, and what influences the decision where to publish.
- Questions related to electronic publishing.

In total, there were twenty-seven major questions, including 169 sub-questions that demanded an answer, usually in the form of selecting one option from a choice of five (e.g. "1 = strongly disagree" to "5 = strongly agree" or "1 = not important" to "5 = very important"). Some of the questions also required respondents to input a number (e.g. "How many papers do you read per year?"). The form was designed in Microsoft Word and then translated into a Web-based form (Figure 2) that fed the answers into a file that could be

opened by popular spreadsheet programs.

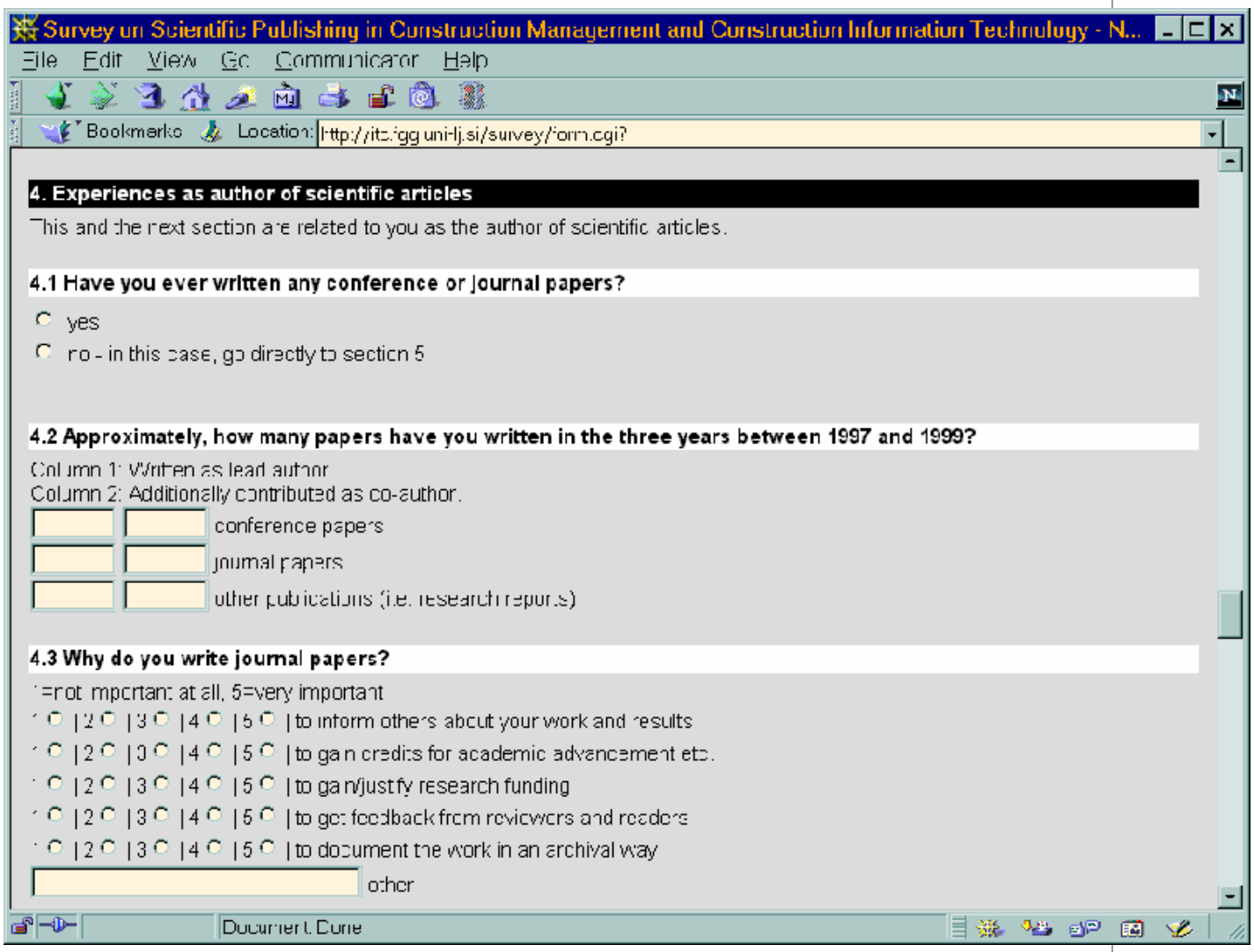


Figure 2: Questionnaire form on the Web (the full form is at http://itc.fgg.uni-lj.si/survey/form.cgi?.)

The form was announced on the Web on February 14, 2000, and the presented results are based on the data that accumulated during one month, until March 14. The survey was announced electronically to mailing lists subscribed to by academics and researchers in the field as well as any person who has at any time in the last seven years shared the same cyberspaces as the second author of this paper. Ziga Turk's mailbox was scanned for e-mail addresses, most of which were professional contacts. We estimate that the e-mail announcements reached between 3,000 and 3,500 persons. About 800 curious individuals actually followed the link to the form on the Web; 247 tried to fill in the form. A total of 236 of those respondents answered at least 45% of the questions. The results presented here are based on these 236 responses; 65% are from persons who stated that their "main interest area" was information technology in architecture, engineering, or construction and 20% who stated it was construction management and economics.

An average respondent answered 82% of all questions. Figure 3 shows some basic demographics figures.

[missing figure]

Figure 3: Some demographics of the survey: this figure is dumped from the Web where users can explore the survey results interactively (Björk and Turk, 2000).

Theoretically, the response rate compared to the original mailing was a little over 7%; however, given the number of the researchers in the fields and active participants at

conferences, we can estimate that at least in the construction information-technology domain, we have reached a very significant number of professors and university academics. If the responses are compared to the people who actually had a look at the Web questionnaire, then the response rate is closer to 29%. We had hoped that more students and younger researchers would complete the survey. However, the student population is not stable and seems not to be leaving a lasting e-mail trace on the Internet. In addition, dated e-mail addresses for students may not have been reaching their target as some students have moved off to other opportunities after graduation. The survey was anonymous, but there was an optional field where the respondents could provide their e-mail address (question 1.6). Only 20% preferred to hide it.

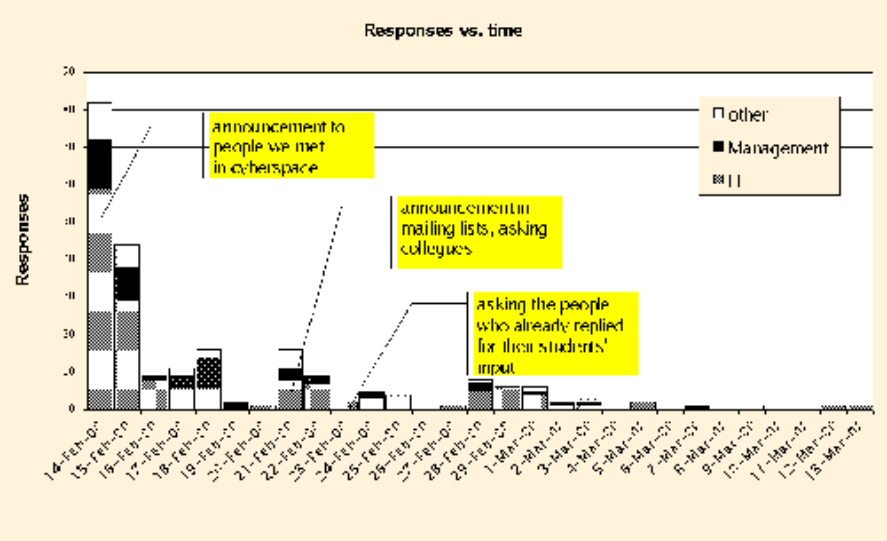
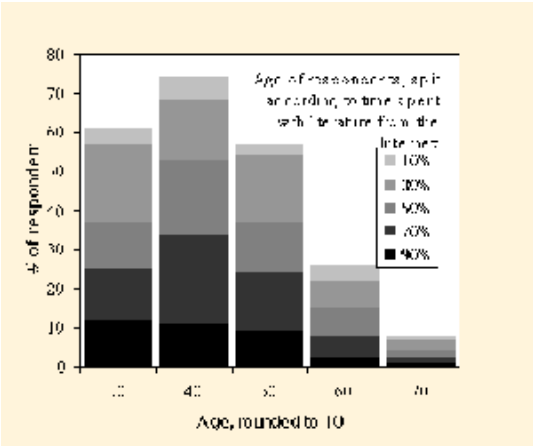


Figure 4. The responses to the form as a function of time.

We were pleasantly surprised by the volume and speed of the response (Figure 4). Although it took an average respondent twenty-six minutes to fill in the form (the system measured the time between displaying the form and the time at which it was submitted), more than half of all replied to the form in the first twenty-four hours and 75% during the first week. After one week other announcements were made; therefore, we can assume that people respond to e-mail messages either very quickly or not at all. It is also noticeable in the data that the construction-management community reacts to e-mail more slowly. The responses came from forty-seven different countries (17% UK, 11% USA, 9% Sweden, 7% Australia, 5% The Netherlands, 4% Germany and Canada, 3% Belgium, Finland, Slovenia and Denmark).

The average age of the respondents was forty-four years (question 1.4).



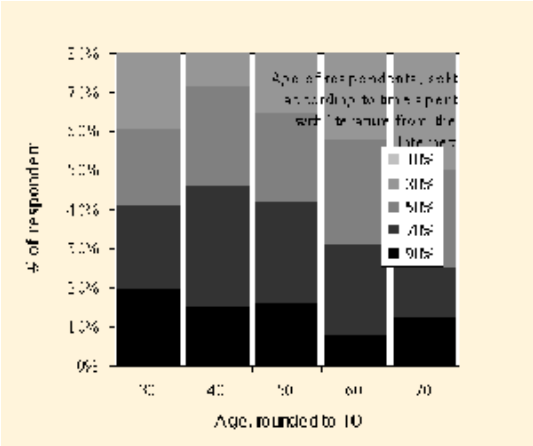


Figure 5: Age vs. the use of the Internet. Note that the heavy use of the Internet drops with age. However, the population from 35-45 uses the Internet slightly more than younger or older colleagues.

The discussion in the following sections is a summary of some of the more interesting results as shown in Björk and Turk, 2000. It is also a result of a correlation analysis among the 161 sub-questions. Figure 6 shows a correlation matrix where relatively high correlations (<+0.5) are painted dark green, (<+0.25) light green, and negative correlations (> -0.25) orange. Absolute correlations were not very high, because most are based on answers where values between one and five had to be selected. This analysis enabled us to find correlated questions quickly. Figure 6, for example, shows that the construction information-technology and construction-management communities are rather distinct — correlation exists for the visits of the conferences in each field separately.

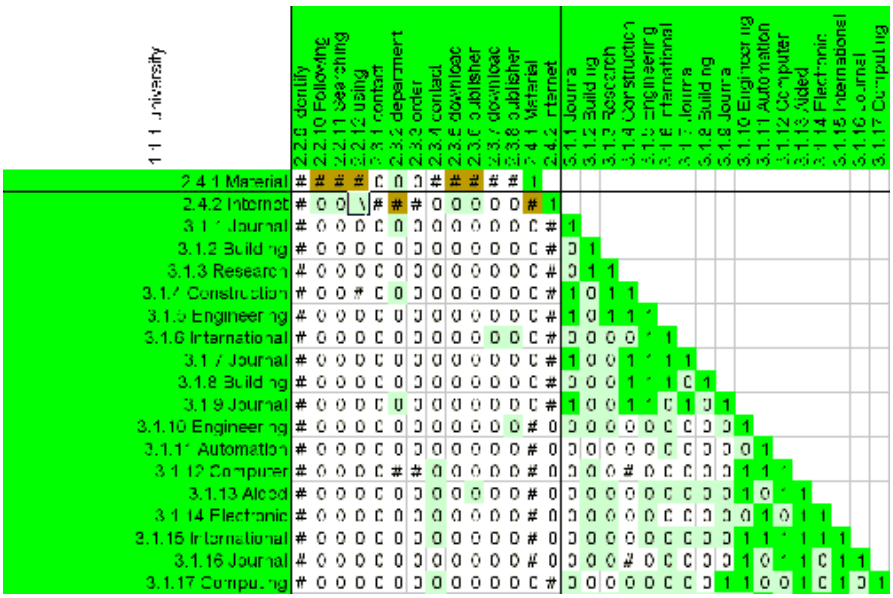


Figure 6: Part of the correlation matrix. Row and column headers refer to the question numbers. Dark green cells mean significant correlation.

Reading habits

The survey showed (question 2.1) that personal discussions are the most important source for learning "about what is going on and should be of relevance" (3.9 on 1 for low through 5 for high scale), followed by books (3.7), journals (3.6) and conference proceedings (3.4). The question did not ask about the medium (e.g. paper or the Internet). However, many respondents used the option "other" to say that they use the Web as such and do not care if the item on the Web is a journal article, report, or book. The relatively least utilised sources are mailing lists and workshop presentations.

To "identify or find particular interesting items worth reading" (question 2.2), the most

important are "references in other publications" (4.0), Web search using general (3.9) or topic-specific search engines (3.6), and hyperlinks (3.7). Among the least important are searches in the traditional bibliographic databases (2.7) or browsing in libraries (2.3) or the Web, not looking for something in particular (2.8). This set of answers clearly shows the strong preference for the just-in-time search in the readily available (references in other publications) or free (Web) resources. According to the survey, just-in-case browsing is somehow more important to the professor/teachers cohort than other groups. This might be explained by the fact that the professor/teachers cohort, through their network and connections, receive many more free paper copies of material to browse than junior researchers and Ph.D. students. That group's function also is to keep a broad perspective, which they can maintain through the just-in-case browsing.

"Definitively, researchers in our community do not want to pay for electronic documents, either through paid subscriptions or pay-per-view mechanisms"

The interesting publications, when not readily available, are most often retrieved (question 2.3) by downloading them free from the author's (3.9) or some organisation's Web site (3.5). Going to the library to fetch a publication or ordering it through a library services is nearly a whole level less important (3.1). There is a positive correlation between "downloading from an author's Web site," the use of "topic specific portals and search engines," "browsing the Web not looking for something in particular," "following hyperlinks," "trying out IT tools," and, of course, the requirement for "fast access to Web sites." Few individuals contact the author and ask for a copy; however, if they do, they do not care much whether it is a paper (2.5) or a digital copy (2.4). In fact, there is a strong correlation between the two, showing that some simply go and ask the author, no matter if a paper or digital copy is in question. It seems, however, that those who would ask for a digital copy find it on the Web anyhow. Definitively, researchers in our community do not want to pay for electronic documents, either through paid subscriptions or pay-per-view mechanisms (1.7 and 2.0 on a 1-5 scale).

One of the simplest but most important questions (question 2.4) in the survey was "Estimate the way you received or accessed all the material that you read, Internet vs. paper. Enter in percentages of time spent reading each category, adding up to 100." The average respondent does it almost 50-50% (Internet : Paper). There are, however, insignificant differences when it comes to age (Figure 5). For the construction-management respondents only, the ratio is 65:35 in favour of paper, while for the construction information technologists, the ratio is 45:55 favouring the Internet. Those who use the Internet a lot do not read or subscribe to journals, because they find "essentially the same information on the Web." They believe that "a paper on the Web will be read by more colleagues than one printed" and that "a paper on the Web is more likely to generate personal contact with readers." Apart from this, there are no other significant correlations between this information and the views on electronic publishing.

Surprisingly, professors and teachers use the Internet more than students. The heaviest users of the Internet are engineers and professors not working at a university (Figure 7). The Internet seems popular in environments where traditional paper publications are not available, are difficult or too much trouble to get, or where tangible results are expected more quickly (engineers, researchers, non-university environments). This chart also shows that the largest advantage of the Internet lies in the vertical communication of the scientific results between academia, research, and industry. The respondents from research and industry, however, often found scientific articles "too academic or too long."

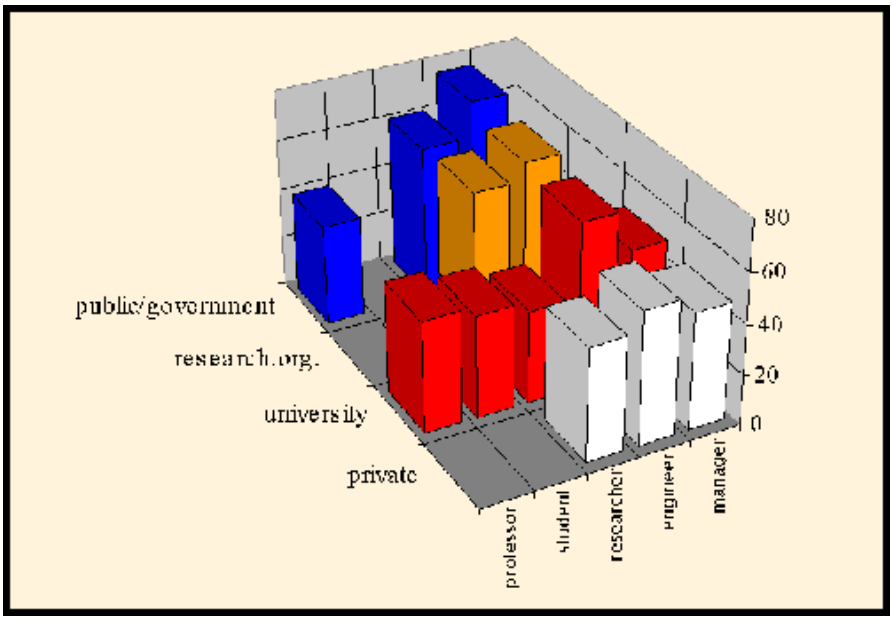


Figure 7: Time spent (in %) reading scientific information retrieved over the Internet rather than in paper form, by the job function and affiliation.

Readership of journals and visits to conferences

Section 3 of the survey dealt with the readership of journals and conference participation. Fifty-five percent of the respondents were not familiar with an average journal listed (question 3.1) and 70% were not familiar with an average conference (question 3.4) listed. On average (question 3.2), the survey respondents "browse through or read in detail" 107 papers per year and travel to 2.75 conferences. Some manage to read 500, 1,000 or even 2,000 papers a year! A group that reads more is construction management people (average 130), "working at a university" (average 128). Construction information-technology colleagues read less (102), particularly the students (60). In all, those declaring themselves as "researchers" read more (176), which is more than the professors (103), students, or engineers (78) and managers (8).

Professors seem to attend twice as many conferences (question 3.5) as students per year (3.4 vs. 1.7). Construction-management people travel to conferences more than their information-technology colleagues (3.2 vs. 2.5). For some strange reason, there is a relatively strong correlation between the number of conferences visited and time taken to fill in the form! This was the strongest correlation to the time taken to fill in the form. Heavy Internet users travel less than other cohorts.

People who spend more than 55% of the time with material retrieved over the Internet seem to read fewer articles (95). Why don't people read more (question 3.3)? Mostly the reason is: "simply don't have time to read more than I do currently" (3.9). People who did not find this reason important "browse the Web, not looking for something in particular" a lot and or learn a lot from books. Other important reasons for not reading more are subscription problems (3.6) and wrong topics (3.4). There is a positive correlation among those who found "content available elsewhere" and finding "the papers too long or academic."

Authoring of papers

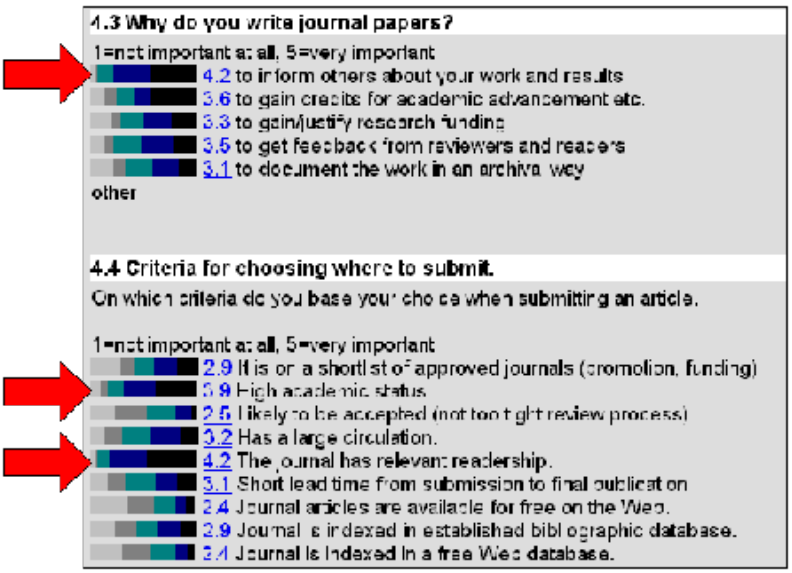


Figure 8: Why scientists write journal papers and how they choose where to submit them.

Of the respondents, 85% have at some time during their career written a conference or journal paper (question 4.1). In the last three years they authored or co-authored 7.7 conference papers and 4.8 journal papers (question 4.2). This fits with the number of conferences they had attended. The survey also indicated there is a strong correlation between people writing any kind of publications. Those who publish more, publish more in all of the formats.

The important reasons for publishing in journals (question 4.3, Figure 8) are "to inform others about your work and results" (4.2) and to "gain credits for academic advancement" (3.6). When deciding where to submit an article (question 4.4), the most important reasons are relevant readership (4.2) and high academic status (3.9). Being on a "short list of approved journals (promotion, funding)" is much less important (2.9). Ironically, respondents don't care much if journal articles are available free on the Web (very important for them, when asked as readers, but not when asked as authors!). Given the preferences for finding and retrieving the information, this should be essential if their aim is to actually inform others about their work — unless, of course, the "others" are not their peer researchers. One could speculate that publishing in a reputable journal, despite the answers given in the questionnaire, is more important than a large readership. Some comments given in the Web form are also indicative:

> Until the motivation for publishing is true to the intent of archiving and disseminating knowledge, rather that just an academic requirement for promotion and tenure, or a $$$ opportunity from a publisher's perspective, we will still have a lot of meaningless published work around...

And also:

> The form failed to address the most significant issue to many of us. We need a media, which serves the need to get information — not to publish it for academic merits. There are too many publications containing similar papers from the same authors. Too many papers mostly refer to other papers without having any original content. The value of journals is becoming questionable.

Forty-two percent agree or strongly agree that the review process is not blind, 47% that reviewers like papers that go along with traditional (their) approaches in the field and want their work cited vs. only 22% who feel that reviewers like radically novel ideas or approaches (question 4.5). There is a lukewarm feeling towards the quality of the reviews: 51% agree that "reviewers' comments help to improve the paper a lot," while 43% also find that "reviewers' comments help improve the research."

The average time from initial submission to printing, estimated by authors for their own articles of the past three years, was 9.1 months (question 4.6). A quick calculation, done by

the first author of this paper, from submission and publication figures posted on the Web site of the ASCE Journal of Computing in Civil Engineering gave an average of 9.6 months (for the 19 articles of the 1999 volume, excluding special issues, for which the month of submission was stated). Similar figures could probably be obtained for the majority of journals included in the survey, but that involves quite a bit of work and would in many cases require access to the paper issues. The average for articles in the *Electronic Journal of Information Technology in Construction* is 4.5 months, which is to be expected since papers can be published immediately once they have passed the review and editing stage.

The reasons for writing conference papers (question 4.7) are similar to those for writing journal papers, but the most important one is "to get the immediate feedback, discussion and contacts based on the presentation" (average score 4.3). To 76% "spending some time near a beach" is not important. If the respondents have been truthful in answering this question, then conference venues should move to campus environments or large international airports.

In the questionnaire, respondents were asked to rate "relatively how much credit do they get for the publishing in the different types of media" (question 4.8). They were asked to rate the credit on a scale from 1 to 10. On average, there is a 8 points vs. 3 points ratio for publishing in the best journal vs. at the most obscure conference. Electronic journals received about 5 points, which is less than conference proceedings published as a bound book with page numbers.

Views on electronic publishing

Students and professors found it particularly important that electronic publications are peer reviewed (especially in the construction management field). On average, 69% find peer review important or very important (question 5.1); 75% find it important "that the papers should be available quickly on the Web, not bothering delays of formal publishing procedures." The questionnaire failed to point out that peer review is part of such a formal publishing procedure. All except the professors find the "quick availability" more important than peer review. Other important features of electronic journals are e-mail alerting services and multimedia attachments.

When it comes to the format of electronic publications (question 5.2), fast access, full text available in any HTML browser, and the option to print on paper are important or crucial to at least 75% of the respondents. The attributes not deemed important are professional graphics design, many hyperlinks, availability in printed format, and availability on a CD-ROM.

A strong majority (67%) believes that papers on the Web are more likely to generate personal contacts and that they will be read by more people (59%) than printed articles. The opinions are split on the controversial statement that "authors have a moral right to post copies of their own publications on their department's Web site, despite possible breaches of copyright" (39% agree, 26% disagree, others have no strong opinions) (question 5.3).

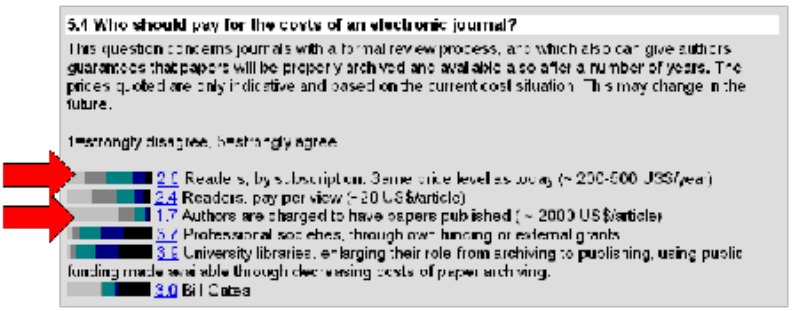


Figure 9: Who should pay for the cost of an electronic journal.

There is much support for free electronic journals where the costs would be covered by, for instance, professional societies or university libraries, rather than electronic journals accessible only by traditional paid subscriptions or site licensing (question 5.4, Figure 9). The options "Professional societies, through own funding or external grants" (3.7) and "University libraries, enlarging their role from archiving to publishing, using public funding made available through decreasing costs of paper archiving" (3.9) received very high support, whereas "Readers, by subscription" (2.6) and "Readers, pay per view" (2.4) were much less popular. The option of

author charges was ruled out (1.7).

Discussion

The authors are well aware of the limitations of the particular questionnaire method used. In an ideal world, we should first have identified the relevant target population of researchers rather exactly and then either sent a questionnaire to all of them or to a statistical sample. Also, phone interviews with rather few questions would probably have provided results that are more representative for the whole population, since it is more difficult to avoid answering a phone interview than an e-mail message or a letter. Techniques such as the ones outlined above were in fact used by the research team of the first author in the information-technology-barometer study (Howard et al 1998), where the usage of information technology in the Swedish construction industry was studied. The costs of that survey were, however, in the order of 50,000 Euros and there was, despite all precautions, a strong suspicion that those members in the studied population who had a positive attitude to information technology responded more eagerly to the study.

For cost reasons, doing this survey in such a way was out of the question. Approaching the potential respondents by e-mail was very inexpensive, and several existing e-mail lists were used, the union of which covered the intended research community reasonably well. Using a Web interface for obtaining the answers, rather than a printed form, dramatically reduced the manual work needed to handle the raw data.

Due to this process, the results should be interpreted with great caution. There is a high likelihood that researchers who use the Internet a lot and who have a positive attitude toward electronic publishing have answered more eagerly than others. Some of the results can, nevertheless, be compared with earlier published results, in particular those reported by Tenopir and King (1998) in an overview article collecting data from several earlier studies and sources.

According to Tenopir and King, the average number of scientific articles (authored or co-authored) per university scientist has increased from about one in 1977 to 2.1 in 1995. This can be compared to the figure of 1.6 of this survey. According to Tenopir and King, scientists at the University of Tennessee read 188 scholarly articles per year (in 1990-93) which is a higher figure than the 107 of this survey.

A very interesting figure quoted by Tenopir and King is the average number — between 500 and 1,500 — of readings beyond the abstract of each scholarly article for nine sampled fields of science. Unfortunately, this survey provides no means for getting comparable figures for our domains. For paper-based journals, getting data about the number of subscribers to journals would help, but doesn't tell much about the readership of individual articles, which can vary a lot. (We did in fact send a questionnaire to fifteen editors of the journals included in the survey. Not a single one answered! Because of the often-low subscription bases in our field, editors tend to regard this sort of data as trade secrets.) Also there may be several readers — or, in some cases, null readers — for library copies of journals. One of the positive things about information on the Web is that it is easy to check the number of accesses to individual pages. Also, there should be an almost one-to-one correspondence between the number of accesses and readings (after excluding Web robots). As an example, the average number of accesses to the full text versions of an article published in the first four volumes of the *Electronic Journal of Information Technology in Construction* was 770 for the HTML versions and 582 for the PDF versions. We haven't analysed what proportion of accesses is from web robots, but assuming that such an analysis was possible and robot accesses were excluded, the remaining figures should correspond well with the true readership of the articles. This is especially true in the PDF versions since people would normally download these only if they intended to print out personal copies of articles.

Finally, Tenopir and King provide data on the information-seeking pattern of scientists surveyed a number of times between 1977 and 1998. Most of the figures describe the situation before the proliferation of the World Wide Web. According to Tenopir and King, the most common way for scientists to find articles to read is by browsing paper material that comes their way via subscriptions or is located in the departmental library etc. (50-72% depending on the survey). The second most common method is being alerted by colleagues (10-18%). Citations in other works that trigger the interest accounted for only 6-13%, depending on the survey. On-line searching in bibliographic databases accounted for between

1-14% and was clearly growing in the 1990s. These figures are not directly comparable to the figures from our survey because we asked people to grade the relative importance of these different routes rather than percentages. Nevertheless this is an area where Web technology is starting to have a very important impact. It is, for instance, becoming much easier to quickly follow up an interesting reference if a URL is provided. Even if the reference is in a paper version and lacks a URL, one can always try keying in the title of a paper into a search engine!

We hope that this survey has provided a reasonably reliable snapshot of the situation in the very beginning of the 21st century, looked at from the viewpoint of the readers and authors of papers in two narrow scholarly domains. We would be happy to make our questionnaire software available to any colleagues who would like repeat the survey in other fields of science and collaborate on comparing the results.

It would also be useful to follow up by a more thorough study of the "supply side" of the equation by looking at how well the existing journals function in the domain — for instance by benchmarking them against journals and pre-print archives in other scholarly domains. Currently journals are often rated based on prestige or sometimes based on the citation they generate, but we feel that benchmarking methods should be developed that also take into account total readership (as opposed to subscription), failure rate of submissions, and speed of publication from initial submission to final publication.

Also, we believe that the possibilities for free Internet publishing should be studied using more formalised business-process re-engineering methods, which would look at the total life-cycle cost from drafting to retrieval of publications and would provide business cases for changing the process. Such business cases could, for instance, be used to influence public-sector decision makers, who influence the situation more than the mere enthusiasm of founders of free Internet journals and pre-print repositories.

The two most influential players are in fact the United States government and the European Union, including the Commission and the member states, since they together control a large part of the global funding for R&D and also have the legislative means to change the rules of the game. The European Union has issued a directive that important sports events such as the Olympics or World Cup football should be broadcase on public television channels and not limited to special subscription sports channels. The statement said that the free access is of vital public interest. Would there not be an even better case for a directive concerning free availability on the Web of published scientific results from any publicly funded research in a Union member country? This would affect a third of all research publications produced globally.

Acknowledgements

The results of our survey where originally reported in an article in the *Electronic Journal of Information Technology* (Björk and Turk, 2000). Since the audience of this journal is restricted almost exclusively to scientists in construction information technology, and since we believe that the results can be of some interest to scientists from other areas who care about the scientific communication process, we have chosen to submit a slightly reworked paper to *The Journal of Electronic Publishing.* The central empirical results are the same, but some material, which only would interest scientists from our domain, has been omitted. Also, our thinking and our conclusions have slightly changed during the last half year and this is reflected in the current paper.

Bo-Christer Björk is Professor of Information Systems Science at the Swedish school of Economics and Business Administration in Helsinki, Finland. He has a background in Construction IT research and spent seven years as professor in that subject in Stockholm, Sweden. He developed an interest in electronic publishing in the mid 90s and is the main editor of the Electronic Journal of Information Technology in Construction http://itcon.org. In his spare time he is a cinema buff, with a home library of more than 1000 feature films on video. You may reach him by e-mail at bo-christer.bjork@shh.fi.

Ziga Turk is Associate Professor of Construction Information Technology and Documentation at the Faculty of Civil and Geodetic Engineering in Ljubljana, Slovenia. His research topic is construction information technology, particularly computer integrated construction, engineering information systems and communication technology such as the Internet. Dr. Turk is involved in editorial or scientific boards of two journals and in the steering committees of many relevant international associations. In his free time he used to make Web sites. His most popular invention is the www.shareware.com. His homepage and full bibliography is available on the web at http://itc.fgg.uni-lj.si/zturk/. You may reach him by e-mail at ziga.turk@uni-lj.si.

References

Anon 1999, *Survey of the Use of Network Based Products and Services,* International Council for Scientific and Technical Information. Unpublished report.

B. C. Björk, and Z. Turk. 2000. A Survey of the Impact of the Internet on Scientific Publishing in Construction IT and Construction Management. *Electronic Journal of Information Technology in Construction,* ITcon, Vol. 5(2000), No.5, http://itcon.org/2000/5/

I. Butterworth (ed.). 1998. *The Impact of Electronic Publishing on the Academic Community.* Proceedings of an International Workshop organized by the Academia Europaea and the Wenner-Gren Foundation, Stockholm, April 1997, Portland Press, UK. http://www.portlandpress.co.uk/books/isbn/1855781220.htm

G. Hodge and B. Carrol. 1999. *Digital Electronic Archiving: The state of the art and the state of the practice.* International Council for Scientific and Technical Information. [formerly http://www.icsti.nrc.ca/icsti/]

R. Howard, A. Kiviniemi, O.Samuelson. 1998. Surveys of IT in the Construction Industry and Experience of the IT Barometer in Scandinavia, *Electronic Journal of Information Technology in Construction,* Vol 3/ 4. http://itcon.org/1998/4/

D. King and C. Tenopir. 1998. *Economic Cost Models of Scientific Scholarly Journals.* Paper presented to the ICSU Press Workshop, Keble College, Oxford, UK, 31 March to 2 April 1998. http://www.bodley.ox.ac.uk/icsu/kingppr.htm

A.Odlyzko. 1998. The Economics of Electronic Journals. *The Journal of Electronic Publishing,* Vol 4/1.

C. Tenopir. and D. King. 1998. Designing Electronic Journals With 30 Years of Lessons from Print. *The Journal of Electronic Publishing,* December 1998.

C. Tenopir and D. King. 2000. Towards Electronic Journals: Realities for Scientists, Librarians, and Publishers. *Special Libraries Association,* Washington DC.

T. J. Walker. 1998. Free Internet Access to Traditional Journals. *American Scientist,* Vol 86, No. 5, Sept-Oct 1998, pp. 463-471.
http://www.sigmaxi.org/amsci/articles/98articles/walker.html

A. Wells. 1999. *Exploring the development of the independent, electronic, scholarly journal* M.Sc. in Information Management, Electronic Dissertations Library, Department of Information Studies, University of Sheffield, UK. http://panizzi.shef.ac.uk/elecdiss/edl0001/index.html

Product of the Scholarly Publishing Office
of the University of Michigan Library
and DLXS
for more information please contact
spo-help@umich.edu