

# **The Scholar in the Forest: Scholarly Communication in the Internet Age**

David K. Jordan<sup>1</sup>  
University of California, San Diego

## **Abstract.**

The Internet, and attendant innovations in information technology, are changing all aspects of scholarly communication at a pace so rapid that few scholars to keep up with the changes. Striving to make out the forest while in its midst, a case can be made that these changes are bringing positive developments in global access and sharing of information, particularly in third-world countries and "distance learning" contexts. But there may be losses in our systems of scholarly information quality control and expectations about how information maintenance is financed or scholarship rewarded. And the effects upon language choice are still unknown. A cultural anthropologist and university administrator, David Jordan provides a consumer's eye view of possible trends

**Opening.** I want to begin by thanking the sponsors for inviting me to address this assembly. I am feel very honored by the opportunity, but also very humbled. You are all specialists in informational technology, and I am not. What can I possibly have to say that you do not already know? What problem can I point to that you have not already solved? What issue can I raise that is not already obvious to you? What perspective can I provide that you have not already taken account of? Most generically, what am I to say to people who are all my teachers?

---

<sup>1</sup> -In thinking through many of these issues I have profited a great deal from conversations with Brian Schottlaender and Bruce Miller of the UCSD Library, with Nancy Friedlander of UCSD's Eleanor Roosevelt College, with Sharon E. Jones in the UCSD Academic Senate office, and with many other friends and colleagues. Their stimulation is acknowledged, although of course they are not responsible for my opinions.

Because we know that even small changes can lead to large ones, it is a temptation for all of us to try to spot trends, to identify what the computer people call the “Next Big Thing.” In this talk, I would like to suggest some of the trends that I think I see in scholarly communication in the dawning age of the Internet, as they seem to affect me and my immediate faculty colleagues. Many of my colleagues, like me, are ignorant, unable to exploit the emerging technology very effectively. Many of us are gullible, believing outlandish claims about how the technology will solve problems that it in fact will not solve. And many of us are fearful of what we do not understand. I am a cultural anthropologist by training, and perhaps in some sense I am presenting a partially ethnographic perspective about how the professor in my university seem to approach scholarly communication and the Internet.

I will begin by talking about scholarly publishing, since that is where we professors do our most self-consciously formal communicating. This will lead me to the problem of quality control and peer review, which affect whether what we communicate is worth listening to. Then I shall talk about format standards and archiving, by which our work lives forever or vanishes immediately. At the end I shall turn briefly to teaching, and, if there is time, I shall close with some brief remarks about language.

I entitled this talk “The Scholar in the Forest.” There are several references packed into the title. You will probably find it inviting to interpret it as meaning that I can’t see the forest because of my preoccupation with the trees. But the title is actually intended to refer to the well-known logical conundrum that has provoked discussion in philosophy classes since something like it was propounded by Bishop George Berkeley in 1710: “If a tree falls in a forest, and no-one is there to hear it, does it make any noise?”<sup>2</sup> In its original context, the issue was the nature of reality, and whether there was any reality outside of our perception of it.

---

<sup>2</sup> -In its original context, the issue was the nature of reality, and whether there was any reality outside of our perception of it. Bishop Berkeley concluded that there was, since God was always around to perceive events, even if we weren’t, so all events had a perceiver and were therefore real.

The question of whether things exist if nobody notices them has become so famous that it provides the basis for a range of popular allusions and jokes. A recent American joke plays on the tendency of wives to criticize their husbands. “If a man talks and his wife is not around to hear him, is he still wrong?” Another takes the university as its context: “If a professor gives a lecture and all the students are asleep, did he say anything?”

So I begin by asking, if a scholar publishes an article and nobody reads it, is it scholarly communication?

### I. Finding Readers: Who’s Looking?

Unlike the falling tree, which presumably doesn’t care whether its disturbance of the ethers is interpreted as “noise” or not, most scholars want to be heard. They display their data and their interpretations to the world and crave a response. And yet sometimes there is little or no response, or the response suggests that the message, though heard, was not really understood.

Even since some ingenious and sadistic librarian invented the citation index scholars have been comparing the rates at which their colleagues have been cited, and some universities have even built formal consultation of citation indexes into their promotion systems. Several years ago a colleague asked me if I would please cite his work routinely in anything I published so as to help him get promotions, and he offered to do the same for me in return. Another made it a habit always to cite himself so as to increase the number of references he received.

**Citation: The Process.** Librarians and others maintain constant records of the rates of citation of scientific articles. Ordinarily, those of us who write such articles pay little attention to these rates. Many years back, however, an enterprising investigator generated and published rates of “non-citation,” asking what proportion of articles in various disciplines went utterly uncited. I regret that I am unable to locate the reference to this research, but my memory of it is clear enough for present purposes:

My own field of anthropology fared comparative badly, with well over 90% of all professional articles never cited by anybody. It was little comfort that a few fields did even worse. (In literary criticism, the figure was actually 99%, as I remember.) We anthropologists all suddenly had the feeling that our only readers were the “peer reviewers” who decided whether our work should appear in distinguished but apparently unread journals, and perhaps the department chairs who read our articles in order to explain to university reviewers why we should be promoted.

No-one blamed the high price of journals and books or the difficulty of retrieving materials from libraries. Instead we blamed ourselves for writing material that no-one wanted to read. We felt unloved. We felt useless. We felt ashamed.

The feeling lasted about a week, as I recall, and then, of course, we forgot about the whole business and went back to our typewriters.

### Publishing on the Internet

The world wide web places us in a new historical period. Three fundamental facts change everything. One is that on-line publishing is extremely cheap. The other is that on-line reading is extremely cheap. And third is that both publisher and reader can be located anywhere on the planet.

Never in human history have we had the ability to make so much information available to so many people, at a cost that human society could so easily afford. This is the moment scholars have always said they were always waiting for. We are living in what we have always imagined would be Paradise.

Unless, of course, we screw it up. Decisions being made right now may delay or restrict many of the most benevolent effects of it. Let me lay out my vision and my fears.

**Hopes & Fears: Hope.** As an author, my goal is to be read. Very few scholars make any money directly from publication. Scholars are paid by

their universities or their research institutes, or they work independently and support themselves on grants or with inheritances or with “day jobs” of other kinds. Scholars don’t publish to get rich, although being rich is of course a nice idea; they publish to be read.

Thus the widest possible availability of knowledge with as few barriers as possible is a value to which virtually all of my colleagues subscribe.

One of the audiences that is potentially included by the Internet for the first time is the scholarly community of third-world nations. A whole roomful of computers connected to the Internet is a great deal cheaper than even a modest library, and as an author I find it exhilarating to think that an unknown colleague in Haiti or New Guinea or Bolivia may read an article or book on my web site and find it useful, or may publish on his own web site and have his work enter the international scholarly world almost without regard to the financial resources of his institute.

Let me provide a couple of personal examples of what world-wide access provides.

Example 1. I teach a freshman course on paleontology and archaeology, and I am now using “late breaking news” sources from around the world that were never easily available to me before.

Example 2. I put a book and a couple of my articles on my web site and a few months later received a request for additional comments by an anthropology student in Hong Kong studying “the work of the famous David Jordan.” I was flattered, of course. (Academics are easily flattered.) But I was also a little puzzled, until I noticed that all of my writings that she referred to were those available on my web site. In other words, the “famous” David Jordan was famous only to the extent that his writing was available on the World Wide Web.

Example 3. On my website there is a small textbook and reference materials on Classical Nahuatl, the language of the Aztecs. So far it is the only web-based course or reference grammar of Nahuatl, so most search engines locate it easily. I have received Email from several

countries from people who have made use of these materials, people who never would have found these sources if I had published them in a necessarily small paper edition available only through the limited, specialized publishing outlets of Mexico or the United States

**Hopes & Fears: Fear.** In conflict with this goal of universal access to information is the goal of making money by charging for access to information. This is the goal of publishers and of commercial research firms. They argue, correctly, that the cost of editing and distributing scholarly material must be covered, and if they are supported by venture capital, then profit must be generated as well. Free access, it can be argued, does not generate a stream of revenue to pay the bills. Further, some would argue that the prospect of someone getting access to David Jordan's writing without paying for it is inherently unfair, that it takes something away from David Jordan (or his publisher), even that it is theft.

I believe that my vision will be realized and my fear will prove groundless. That is, I think I see evolutionary trends that favor universal access to information, and that suggest charging for access to information has no future. Let me suggest what these trends may be.

### Publishing As Making Public

It is in recognition of the collaborative nature of science that learned societies have journals, and that universities and academies of science around the world underwrite the cost of publication of scholarly research. In principle scholarly publishers, like public libraries, seek to make human knowledge available as widely as possible by removing financial barriers to propagating knowledge.

In the world of paper publishing, there have of course always been practical limits. Just as no library could afford to own everything published, no publisher could ignore the fact that there are costs, and there had to be limits on how many pages could be printed, how many photographs or graphs or maps could be included, how long copies could

remain in the warehouse, how much could be spent on advertising to find an appropriate audience, and so on.

The dawn of on-line publishing has drastically reduced most of these costs. I will list some of these, but the point I will try to make is that as the costs drop, the need for professional publishers also drops, and self-publishing becomes an inevitable result.

**1. Production Costs.** Compared with the costs of paper books and journals, server space is extremely cheap and additional readers bring no added costs. The cost of publishing a journal that may be read only by members of the Paleontological Society of Tanzania and the cost of producing a journal that may be read by anybody in the world is exactly the same. And such enhancements as color photographs and in-text graphs have no effect on cost.

**2. Marketing Costs.** Finding an audience on-line is potentially simply a matter of being included in a bibliography. That is why I get queries about my on-line textbook of Nahuatl; it is only one mouse-click away from a search engine. There are complications, to be sure, but the point is that the difference is both real and substantial.

**3. Technical Preparation Costs.** With computers, even before the moment came for on-line publishing, more and more of the editorial task formerly involved with paper publishing had been turned over to authors. Today even major university presses expect to receive well edited manuscripts in electronic format suitable for computer-mediated typesetting with little additional editorial attention. Indeed it can be a point of pride for publishers to demand computer-readable copy from authors.<sup>3</sup>

An important reason for all this is that software packages make it feasible for individual scholars or their immediate assistants to generate

---

<sup>3</sup> -I was amused several years ago when an American university press required a book MS to be submitted in electronic form, and then simply printed it out and proceeded to work with the paper copy, which was recopied by hand when they did the typesetting. The need for me to submit the original MS on a disk was to show me that the press was progressive, not because they knew what to do with the electronic version.

legible graphs, draw maps, process digital photographs, make proof corrections, and create attractive layouts that put to shame the efforts of many professional publishing houses of only a few years ago.

**Remaining Costs.** Two editorial functions have not dropped in cost. One is translating or polishing the prose of the writer. Another is critically reviewing the content so that the imprint of the publisher effectively represents a guarantee of quality. And this of course brings us to the issue of information quality.

## II. Quality Assurance in Scholarly Publishing

The quality of information that is made public is not a new concern. When printed books first appeared in China, there was, I have been told, opposition, and various people argued that the dawn of printing posed a potential threat to civilization as they knew it. The enormous expansion of printing in the Song dynasty brought repeated, only partially effective, government attempts to restrict the circulation of printed materials. There were, as I understand it, two important arguments that opponents of printing used to defend their position, both of them related to information quality control:

First, while opponents of printing acknowledged that it was desirable to make edifying works available to a wider public, it was argued that undesirable books (particularly books undermining state power) could also be distributed in printed editions, posing a danger to public morality. An edict issued in 1009 required that all books be inspected by the government before they could be printed or sold, and detailed restrictions were issued in 1090 that included, among other things, a blanket prohibition on all “frivolous and licentious publications.”<sup>4</sup>

Beyond objectionable content, another argument against printed books apparently maintained that accidentally printing multiple identical copies

---

<sup>4</sup> -Denis Twitchett 1983 *Printing and Publishing in Medieval China*. New York: Frederic C. Beil. P. 61. From the Song to the Ming dynasties, the government imprimatur sometimes included a statement of author's rights and constitute the world's first government certified copyright. The earliest known example, from Sichuan, dates between 1190 and 1194, Twitchett tells us (p. 62).



of a work containing an error would produce multiple identical copies of the error as well. Worse yet, since a copy once distributed could not easily be located and recalled for correction, such errors had the possibility of propagating themselves for generation after generation and across the whole of the civilized world.

Thus critics of printing saw it as cheapening the quality of the written text and potentially damaging society through the repetition of error. They were correct on both counts. Printing really has been used to promote anti-social causes, and it really has duplicated both large and small errors on a grand scale. But the cost-benefit ratio is so skewed in the direction of benefits, that no-one today would seriously consider banning printing, and indeed even the most minor restrictions on printing are met with suspicion in most countries.

Similar criticisms are made about the Internet. Like printing, electronic media can be used for anti-social purposes. Modern governments worry about the ability of criminals to make use of means of electronic communication to increase the efficiency of their crimes, from sales of fraudulent products to credit card abuse, from money laundering and smuggling to the coordination of terrorist attacks.

The second objection to printing in China, namely that it would promote the duplication of errors, also applies to electronic media, and this Song-dynasty objection is raised by many of my university colleagues against using the Internet as a source of information (particularly information involved with the production of student term papers).

“The Internet is full of junk,” they say. “Anyone who uses it knows that.” I cannot speak of Internet publishing without colleagues pointing this out, as though this ended the discussion. The internet as a place where archives of previously published material are made available does not disturb their imagination, but the internet as a place where new material is published seems dangerous to them. “There must be peer review,” is the rallying cry of the opponents of on-line publication.

There are of course peer-reviewed web sites, but most web sites are not peer reviewed, and it is more interesting to turn our attention to peer review itself and what it has done for us. I will argue that its future utility to us is likely to be much reduced. Since peer review is a matter very close to academic hearts, I expect those of you who are still awake when I finish will want to attack this view with special vigor.

## Peer Review

**Economic Need for Peer Review.** In a world of limited publication space, restricted by the economic considerations involved with typesetting, manufacturing, where-housing, and marketing, it was critical to publish only what had the highest probability of using these resources well. In the case of commercial presses, the goal was to sell lots of copies. In the case of university presses and learned societies, the goal was (theoretically) to serve the scholarly public. Review by the author's peers was one way to make the decision about whether the quality of the research was adequate. Review by a marketing board was one way to make the decision about whether the work could be afforded.

In other words, traditional peer review, like traditional market review, has been a mechanism for the rational allocation of scarce publication space, of printed "real estate." With the advent of the internet, the bottom has dropped out of this "real estate" market. Space for articles no longer matters. (Neither do color pictures or placing footnotes on the page cited or using better or worse paper or bindings, and all those other economic considerations.) Peer review is unnecessary because the job it did no longer needs to be done. There is infinite land. We don't need to fight over it any more.

**Other Uses for Peer Review.** Peer review has of course had some other functions. One was to associate a publisher, whether of books or of journals, with a certain standard of quality. Princeton University Press is a prestigious place to publish. Jordan's Garage is not. The reader believes that he can trust what comes from Princeton. He correctly doubts the he can trust what comes from Jordan's Garage. The

disparaging term “vanity press” recognizes the difference that peer review makes in our evaluation of content. In its absence every reader must make up his own mind about quality, which not all readers are prepared to do. The problem is real, and we will need to address it.

Peer review has other importance, unrelated to protecting readers, of course. In universities it is critical in the evaluation of candidates for academic hiring and promotion, and we should expect it to remain important there. However we may find that the nature of such peer review changes. At my university, the Committee on Academic Personnel has actually stated in writing that “the acceptance of an article by a journal known for its rigorous review and high rate of rejection is considered a guarantee of the quality of the work.” (Internal UCSD memo, “Where CAP Stood 1997-97.) In a world in which there is no economic need to reject scholarly work, evaluating a professor’s worth by his ability to publish in journals with high rejection rates is becomes laughable as a “guarantee of the quality of the work”! Peer review will remain important, but it will need to done in the context of its goal. Solicited peer review specifically focused on personnel actions may actually increase in importance.

**The Unreliability of Peer Review.** Unfortunately the process of peer review does not always usefully recognize quality. In some fields standards of excellence are well canonized. I have the impression that my colleagues in chemistry, for example, are in good agreement about whether a given scholarly article is a solid and useful contribution or not. In some other fields, standards scarcely exist. In cultural anthropology many of the recommendations made by anonymous peers can be too easily dismissed as largely political, with evolutionary psychologists unlikely to find merit in the work of postmodernists, for example, and vice versa. In such a situation, the quality of work which is accepted for publication is not guaranteed, and even “peer reviewed” journals tend to slide into becoming the party organs of one or another school of thought, while some more respected journals depend largely or entirely on the judgment of a single editor.

**Trend.** In the absence of an economic need for anonymous peer review to allocate intellectual real estate, we should expect it to become far less important in on-line publication than it has been in the past. We should then see ripple effects in such areas as faculty promotions.

Let me now consider peer review from a somewhat different perspective. Let us ask how knowledge itself is advanced, and what kind of peer review contributes to knowledge and to what degree. I will try to argue that publishing everything is better than publishing only what has already been shown to pass muster. I begin with an excursus on open-source software, with the caution that this is one of those trees that can keep us from seeing the forest.

### III. Open-Source Knowledge

**Excursus on Open-Source Software.** Computers are nothing without software, and the development of software remains the major preoccupation of the computer industry in our era. Yet a curious trend is developing in this field, namely the tendency to give software away for free. Important examples are the Linux operating system, which has become prominent as the preferred operating system for web servers, the Perl programming language, used by many or most of the web servers that process user input on the server side, and the Netscape web browser, used by a substantial proportion of the world's web users. The modern world depends on these products to such an extent that the sudden absence of any one of them would produce a major global catastrophe, yet each is free for the taking and can be modified by users as they wish.<sup>5</sup>

The obvious puzzle is how a programmer or company can make any money giving a product away, and how a product developed in their

---

<sup>5</sup>-A key distinction is between free software (which is simply given away but sometimes in a form that cannot be modified), and open-source software (where the underlying program is made publicly available and can be modified by any user). We are speaking here of open-source software, for that is what has been making such rapid strides in creating new universal standards of software performance and market penetration. In an important way, however, the same argument can be made about internationally agreed-upon standards for data representation to computers, such as ASCII, Unicode, or TIFF. My university's "data warehouse" tells me that it prefers to store documents in TIFF format rather than Acrobat format because the former is open and the latter proprietary.

spare time by unpaid programmers can possibly compete with a product developed in a well-financed lab that expects to make a profit by selling it.

Perhaps the best known analyst of this phenomenon is Eric S. Raymond<sup>6</sup>, who argues that the reason for the success of open-source software is that its easy availability inspires broad use by a wide range of sophisticated people many of whom actually enjoy improving it and becoming known as improvers of it. The community of interested and competent users of important software so far exceeds the closed development team that a company can typically put together that open-source products have a level of focused brainpower behind them that enables them to out-compete commercial products.

Most professional programmers who work on open-source software are actually hired by companies and agencies to develop programs for internal use. Less than 5% of the software they write, Raymond estimates, is intended for sale; most of it is intended to solve problems internal to the companies that hire them. Thus in the vast majority of cases, release of software to an “outer world” represents no loss of income, and can actually be profitable in that external users may help to debug it and extend its functionality.

**Open-Source Physics.** The parallel with science in general is arresting. Physics does not progress rapidly when a physics lab keeps its research secret, like the “mad scientist” of a Victorian novel. Rather, physics makes rapid progress when large numbers of people think about the same problem, seeing different aspects of it, imagining different approaches to it, conceiving different applications related to it, and being esteemed by their peers for breakthrough insights. Successful science is the science that is vastly collaborative. This is what is recognized in the process of peer review, but it is also what is recognized in what we might

---

<sup>6</sup> -Eric S. Raymond 1999 *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. Sebastopol, CA: O'Reilly.

call “open-source scholarship,” in which broad accessibility of information invites broad experimentation with each idea proposed.

The analogy with Open-Source software is to be seen in the recent Open Archives Initiative, by which “Scholarly authors can make electronic documents available to a global audience by submitting them to e-print archives” (<http://www.openarchives.org>). The initial example is in physics (<http://arXiv.org>), which already has a “hit log” far exceeding the circulation of most scholarly journals. The organizers are at present concerned to create technical “interoperability standards” among the constituent archives, but importantly for present purposes, they represent a new kind of scholarly institution, one that combines publishing and archiving, that leaves content to the authors, and that represents, in a sense, an end-run around the economic interests of commercial and academic publishers.

Significantly, the initiative, at least as it presents itself at the moment, is not concerned with providing peer review or with limiting listings to peer reviewed materials, but rather with promoting unsolicited and extremely broad voluntary peer review by making listings available for the inspection and comment of “all comers,” just as occurs with open-source software.

Not all scholars are enthusiastic about such developments, and some journals (such as the *Journal of Organic Chemistry*) refuse to print articles which have already been made public through electronic pre-prints. One point of contention is the moment that information is construed to be “published.” Established journals seek to “break the news” through first publication. However a related point is the version of research that is to be understood as the statement of record to which subsequent research properly should refer.<sup>7</sup>

---

<sup>7</sup> -Jessica Gorman 2001 “The End of Good Science? Some Chemists Oppose Posting Research Online; Others Say, ‘Get With the Program!’” *Science News* 159: 76:78. One direction in which “digital libraries” are evolving is toward the establishment of an unchangeable version of submitted research —peer reviewed or not— so that subsequent references to it (including critiques of it) are not vitiated by its mutability.

Probably far more publication of preliminary scientific findings already occurs on the web sites of individuals researchers or their laboratories than in paper sources. One example will suffice. In December, 2000, the Associated Press reported the discovery of a field of hydrothermal vents and carbonate deposits nearly a thousand meters below the surface of the Atlantic. The discovery was made by a team surveying the area in a research submarine. The newspaper matter-of-factly reported:

The scientists have posted their findings on a Web site — <http://www.sio.ucsd.edu>— but have had little contact with on-shore colleagues. They are scheduled to return to Woods Hole, Mass., today aboard the research ship Atlantis. [San Diego Union Tribune, 001216-A10]

No editors. No peer review. No censorship. No intellectual-impact statements. Merely excited scientists eager to make their discovery known to the world as soon as possible. Complete with video clips. This is the initial, not the final, presentation. But it puts the most arresting information in the public record and immediately invites the world's interest and thought.

The trend toward simply putting one's research on a web site has been noted with consternation by the Committee on Academic Personnel of my university in the course of its deliberations about how to recognize faculty accomplishment. (These were the people who wanted to measure journal quality based on rejection rates.) They write:

CAP has noted an increasing tendency for faculty in some disciplines (for example, engineering) to dispense their research through the electronic media instead of published, peer-reviewed journals. CAP has no objection to the medium, but is concerned, rather, with the quality and rigor of review employed.”(Internal UCSD document “Where CAP Stood, 1996-97.”)

They continue:

“More complicated is the evaluation of publication by posting an article at one's website, or disseminating "preprints" by e-mail. These issues are under consideration by committees of the systemwide Academic Senate. (Internal UCSD document “Where CAP Stood, 1997-98.”)

Given unlimited publishing capacity, the cost of substandard scholarship to the publisher is negligible. The cost to the reader is directly proportional to the ability of the reader to discriminate good from bad work. In the case of open-source software, substandard efforts go unrewarded in that they are little used, while excellent efforts receive the recognition of advancing the product.

**Trends.** I believe that all of this suggests that we will soon see the end of peer review in most internet publication. As a researcher in a field where peer review has been a dubious benefit. I am tentatively pleased. Knowing how much my colleagues in the sciences believe that we all depend on it, I am, of course, somewhat troubled.

**Quality Control and the Public.** Not all scholarly communication is directed to other scholars. Not all psychology papers are read by psychologists. There is a whole world of communication directed to the lay public, ignorant, naïve, and potentially gullible. All of us are laity with respect to most fields of knowledge, so this issue should be of universal concern. The need for clear measures of quality control understandable to the laity may be far greater than it is for communications directed from scholars to other scholars.

A couple of years back I assigned a group of freshman students a book on famous frauds in archaeology. They read the book, and rejoiced that *they* were too clever ever to be taken in by such silly ideas as the author described. Then I assigned them the task of searching the Internet to find archaeology sites. They were to find both fraudulent and legitimate sites, and then write papers explaining how one could tell the difference.



The papers were of indifferent quality –these were freshmen, after all. But in the course of class discussion it became evident that they did not really have adequate experience with archaeology to avoid being taken in by foolish ideas. However they tended to believe material presented on sites where the URL ended in .edu, which were universities and must be telling the truth, they reasoned, while those sites ending in .com were commercial and were not to be trusted. Just as I tend to believe that Princeton University Press would not publish a truly stupid and irresponsible book, my students tended to believe that university personnel would not publish a stupid or irresponsible web site. We are both wrong, of course —well anyway they are— but I suggest that server-affiliation or URL has already become an indicator of reliability for some Internet users.

**Trends.** If the “.edu” suffix, or any other, is to be taken as representing a minimum standard of quality, one can imagine greater concern on the part of institutions with what kind of information is associated with them. As a member of an electronic information task force at my university I have already been party to discussions about what should and should not carry the official seal, or what pages should and should not carry the university’s name in their URLs. We may be only at the beginning of this process.

**Hopes & Fears.** Before moving on, let me review this rather meandering discussion. Returning to my vision and my fear, as I stated them at the beginning, my hope is an open Internet, in which publication thrives, and in which the commerce in ideas expands to people and nations far beyond what we have seen in the past. My fear is that this will somehow be constrained. And I believe that the natural evolution of the medium favors my hopes over my fears.

#### IV. Archiving & Technical Standards

I turn now to another hope and another fear. The hope is that the efficiency of digital storage will make it possible to keep scholarly materials in print and available forever, despite their rapid increase. The

fear is that our arrangements for technical standards and for archival responsibility will not be up to the challenge. In this case, my fear is rather greater than my hope.

The problem is summarized in another of those in-house reports from my university's Committee on Academic Personnel as it struggles with the task of evaluating academic files. They write:

CAP was concerned about the possibility that in some instances electronic publishing is associated with reduction of the time after publication that work is accessible to other scholars. This could have the effect of truncating the impact of scholarly work, and of diminishing the quality of ongoing creative work that would of necessity be done without access to the earlier contributions. (UCSD internal document, "CAP Annual Report, 1998-99.")

**Technical Standards.** First let me nod briefly to the issue of technical standards. The Internet works because there are standards. Transmission protocols like ftp and http, encoding standards like Guobiao and Unicode, display standards like HTML, and database query standards like SQL. The Open Archives Initiative is importantly about technical standards, which will be the principle issue of its meeting in Berlin next month (<http://www.openarchives.org/Berlin2001/OpenMeeting.html>).

Like many other people in this room, I have had to enter Chinese characters into my computerized fieldnotes over and over as the notes migrated from one computer system to another before there were standards for embedding Chinese into English text. They are now mostly in Unicode. In theory I have reached the end of the road. But an article in this month's issue of *Multilingual Computing & Technology* describes the new GB-18030 "mandatory standard" for character encoding in China, suggesting that the dust has not yet settled.<sup>8</sup>

---

<sup>8</sup> -Dirk Meyer 2001 "A New Chinese Character Set Standard." *Multilingual Computing & Technology* 12(1): 63-68.

The obvious point is that the cost of shifting standards presents a tremendous danger of old materials becoming unusable or having to be renewed at great cost. Just as wire recordings of the 1940s can no longer be easily read, the computer formats of today will eventually become obsolete.

From my perspective it is irresponsible to chain substantial scholarly publication efforts to a preservation format that does not have general endorsement. As we move toward more and more on-line publication, the issue of readability a century or two from now becomes more and more obvious. Who will use Microsoft Word fifty years from now? How legible with Adobe Acrobat files be in the year 2300? Whose responsibility will it be to move our archives to new formats?

Unfortunately, few of my colleagues have very much understanding of the formats in which their efforts are encoded, displayed, or stored. Other of my colleagues are eager to see new research become available immediately, and have little interest in long-term accessibility.

It falls to the information specialists to ensure robust international standards, ease of data transformation, backward compatibility, and the honest discussion of these considerations in software reviews.

**Publisher & Archive.** Now let me worry about archiving. On-line publishing blurs the distinction between a publisher and a library. The Bibliothèque Nationale de France, for example, has scanned and made available a large number of reports of the Bureau of American Ethnology (<http://gallica.bnf.fr>). Does this constitute republication or circulation or archiving?

The problem goes beyond semantics. The ease of publishing on-line obviously makes it inviting to do this oneself not only in the form of articles, but even in the form of on-line books. It is unclear who will end up assuming responsibility for the maintenance of on-line sources. Governments? Libraries? Publishers? Scholarly associations? New consortia that combine these in some way? Worse yet, since on-line

materials, unlike paper materials, can easily be modified, the impulse to keep them “up to date” may be a disaster from the perspective of our archival instincts, even if we do figure out who is to be the archivist.

As a researcher working in a field where old books and articles are often even more useful than new ones are, the prospect of such materials abruptly disappearing is deeply disturbing. My enthusiasm for uniting the libraries of the world into one vast global resource is tempered by my recognition that this may be one of the most fragile resources the human race has ever created if we cannot somehow solve the problem of sustaining its stability.

Let me provide a very personal example. I am a rogue publisher. My old book *Gods, Ghosts, and Ancestors*, finally out of print in its second, Taiwan edition, still is in use as a textbook by some instructors. Since I myself own the copyright, I decided to create a third edition and put it on the Internet. Making a new edition allowed me to switch the Romanization from Gwoyeu Romatzyh spellings to Hanyu Pinyin spellings, and it allowed me to restore the additional photographs that had been omitted from the earlier editions to reduce publishing costs. Now when I receive reprint requests, I can refer the instructors to the on-line location and they can do their own reprint or refer students directly to the site.

I thought very little about what I had created when I posted my book until I discovered another “new edition” on line, this one a second edition of a bilingual dictionary prepared by her colleagues after the original author’s death. Unlike the earlier print edition, the on-line one incorporated extensive revisions she had written into the margins of the first edition. Unlike my new edition, which added very little new content beyond pictures, the dictionary was substantially revised. No plans were made for a paper edition, since the on-line version was already available.

The advantages are obvious: a new edition is inexpensive, easily updated, and freely available. Indeed, if I chose, I could go through my files and find more pictures and add them to the book tomorrow. Or if I rethought

an argument, there is nothing to prevent my revising the on-line book. There ceases to be any significant dividing line between a third edition and a fourth, or a fourth and a twenty-seventy. The book becomes a living text, evolving as the author's thinking evolves, interacting with its time, responsive to new intellectual currents. As an author, I love this.

The disadvantages are also obvious: The on-line version can vary from day to day and can vanish just as easily as it can appear, at the whim of the author or the expiration of the author's computer account. For the person who cites its arguments, the possibility of the argument later changing is disconcerting and, if one is legally inclined, begs the issue of whether or not one is misquoting another's work.

Information specialists are addressing this problem with various technical means of time-stamping and archiving web pages. It is possible to create a more or less constant on-line library in which an unchanging, unchangeable, canonical version of an on-line work can be safeguarded from further modification by a fickle author like me. For the legal-minded this is desirable, perhaps even essential. For the author it trades the immortality of static archiving for the flexibility of readapting and updating one's work. The trade-off is only sometimes worth it, and people like me are part of the problem.

**Closed Sources.** Some sources and even some indexes are hidden from anyone who doesn't pay a consortium membership, a user fee, or some other cost. The sources that are hidden are not conspicuously better than the sources that are freely available, in my experience. And certainly my students and colleagues have never expressed a preference for fee-based sources. Charging a fee to use something is probably a sure way to reduce use of it.

I think the trend does not favor closed Internet archives or pass-word protected Internet journals. If my university works to provide information to facilitate my research, it costs neither more nor less if such materials are made available to the general public, and it has the additional benefit of increasing the university's visibility and general

academic renown. If, on the other hand, my university is reluctant to “give away” the fruits of its labors without outside payment, its visibility is reduced. To the extent that all or most universities and research institutes develop closed and proprietary archives, journals, or indexes of research materials, the net benefit to all of us is clearly less than if all or most of them are accessible to the whole public. From the perspective of a scholar seeking to have his work found, it is an advantage to be included in a data base that anyone can access. Academic authors want to be read, not protected from readers who don’t pay publishers.

**Trend.** My prediction is that closed-access web resources will wither and die, probably quite soon, for want of customers willing to bother with them. If I am right, then we need to find a way to finance knowledge production and knowledge propagation that does not assume income from the knowledge consumers. This may be a major challenge.

## V. Teaching

Not all scholarly communication is inter-scholarly communication, of course. I have alluded to the lay public. But a hot issue on my campus is the potential role of the internet in teaching, particularly college teaching. It is perhaps indicative of the anxiety we all feel about this topic that a couple of years back my university established a task force on instructional technology that was charged not with evaluating instructional technology, but with recommending an administrative structure that would be able to deal with this ever-changing world in a way that would maximize benefit while minimizing the money spent on crackpot proposals of unrealistically technophile faculty.

Like any new medium, the true potential of the Internet in university teaching is not always obvious. Posting one’s reading list on a web site is not a very creative use of the medium, but probably the majority of American university class web sites consist of little more. A colleague was praised for his “creative” use of the web in teaching, but a quick visit to the web site showed that all its pages were simply converted from paper materials distributed in class, plus the possibility of taking the

course off-site by viewing the professor's lectures on streaming video. Lectures are hardly an innovation.

I myself have been making extensive use of on-line quizzes that provide the student a response to every question as it is answered. The instant-response quizzes are arguably an innovation, something not possible in a non-web format, and the students are enthusiastic and make heavy use of them without being required to do so (which is certainly not their reaction to paper quizzes!). But I have been unable to show that this results in any measurable increase in what they actually learn by the end of the term.

Although there have been a few successes, and a growing industry of small educational technology firms eager to sell under-tested products to "simplify" our task, relatively few American professors seem to use web pages in a way that make demonstrable new progress in teaching and learning.

It is not particularly surprising that the professoriate is still getting its sea legs with a new medium for the delivery of education. More surprising is the rising fear among many of my colleagues that lecture notes on the web will be "stolen" or "published" without their consent and to someone else's profit.<sup>9</sup> Some fear that a future "Bill Gates University" will deliver all education without the active participation of classroom teachers, making university professors unnecessary. How can a student have a true "college experience," they ask, when taught by a "faceless computer."

**Hopes & Fears.** "Distance learning" has been with us since the invention of writing. With the introduction of the internet, it has taken a long step forward, and presumably even American will eventually catch up with many countries already extensively using the Internet in this

---

<sup>9</sup> -A new section has just been added to our "student conduct code" making it an offense for a student to sell class materials, including notes taken in class, for commercial purposes.

way.<sup>10</sup> My hope is that in fifty years the boring lecture will have gone the way of the single shared book on reserve in the library. My fear is that the level of innovation will be limited to streaming video showing professors writing equations on blackboards.

## VI. Language

Finally, I would like to add a word about language. There is no doubt that English today is indeed the premier language of scholarly communication, just as it is of political posters intended for CNN reports. However this is far more the case in some fields, such as computer science, than in others, such as Peruvian archaeology. English is also the language of the Internet, although some estimates see it overtaken by Chinese within the next five years in number of web users, web sites, and so on.

One of the very rare scholars who seriously considers the future of English, David Graddoll,<sup>11</sup> notes that native speakers of English are dropping as a proportion of the world population, and that non-native speakers of English are now roughly as numerous as native speakers, and are increasing rapidly. Given demographic and economic trends, we should shortly see an end to the monopolization of so much of the world's economic wealth by English speakers, reducing the motivation of non-native speakers to study and learn the language, perhaps slowing the rate of increase in the proportion of non-native speakers. Although Graddoll sees a continuing place for English as an inter-language among non-native speakers if present trends continue, he also sees it becoming less dominant than it now is, as Chinese, Spanish, Arabic, and

---

<sup>10</sup> -As American universities dither in ambivalence about the role of "distance learning" in higher education, the rest of the world plunges ahead and simply does it. A conference in at the City University of Hong Kong in 1998 [?] assembled representatives of an impressive array of active distance-learning-based open universities, from Britain and Turkey to China and Korea, that are offering education to a wider and wider public at a lower and lower per-student cost. It may not be the "great college experience" about which American academics become so easily nostalgic, but it enables the education of people who otherwise would have little access to higher education.

<sup>11</sup> -David Graddol 1997 *The Future of English*. London: The British Council. (See also Graddol's web site, <http://www.english.co.uk>.)



Indonesian move to positions of greater world prominence. These trends have nothing to do with computers or the Internet.

Through the development of Unicode, the Internet already welcomes a host of orthographies and is actively developing the capacity to welcome many more. Being able to publish cheaply and globally in any language may lead to a substantial increase in publication of scholarly materials, particularly preliminary or “advance” announcements of research results, in languages other than English.

About twenty years ago, a colleague and I surveyed the citation rates for articles in psychology and marine biology.<sup>12</sup> The dominance of English was not a surprising finding, but we found that it reached such an extent that, for example, “A few journals issued in Japan (such as *Plant and Cell Physiology* ...) publish articles in English (and occasionally also in other Western languages), but do not even accept articles written in Japanese” (p. 436). One Japanese informant explained that he published preliminary findings in Japanese because it was easier and faster, but considered it necessary, despite the delay and cost, to publish his “important” work in English. It appeared to us that writers in many other languages did something similar.

**Trend.** As the Internet makes cost insignificant in publishing untranslated work, we should expect to see more and more non-English “preprints” of new science appearing on the Internet, some of which may never move through the expensive process of full translation to enter the world of international Anglophone scholarship. On the other hand, the rapidly increasing success of machine-assisted translation<sup>13</sup> is likely to make it easily feasible to include imperfect translations into English or other languages at the time when the work is initially posted. The effect

---

<sup>12</sup> -Ralph A. Lewin & David K. Jordan 1981 “The Predominance of English and the Potential Use of Esperanto for Abstracts of Scientific Articles” in M. Kageyama et al. (eds.) *Science and Scientists: Essays by Biochemists, Biologists and Chemists*. Tokyo: Japan Scientific Societies Press.

<sup>13</sup> -Kurt Godden 2001 “The Coming Revolution in Translation Efficiency.” *Multilingual Computing & Technology* 12(2): 39-42.

on global scientific English of such a trend, should it materialize, is difficult to anticipate.

## Closing

In closing let me say that I continue to have an acute awareness of being lost in the woods, of not knowing the forest for the trees, and all those other intellectual risks implied by my title. Nevertheless I hope that what I have said has not been completely silly and may have inspired someone to see some issue in a way that may grow to a productive insight. My fear, of course, is that you have all fallen asleep.