

PUBLISHERS AND LIBRARIANS: COMPETITORS OR COLLABORATORS?

Karen Hunter
Senior Vice President
Elsevier
k.hunter@elsevier.com

Digital Libraries Symposium
Seattle, January 20, 2007

Both publishers and librarians are in the midst of redefining the roles we play in scholarly communication. As we go through a difficult period of repositioning, the relationships between us – already under strain – will change. There will not be a single new model that fits all libraries or all publishers. Nor will everything of the present or the past disappear by any means. But we should anticipate that much more will change in the next 10 years, even given the already dramatic changes of the past decade.

Looking back over those last 10 years, it is fair to say that together we have accomplished an enormous amount. Ten years ago web browsers were not quite four years old and electronic editions of journals were just emerging from pilot programs. Most scholarly journals had print runs fewer than 1500 copies – frequently less than 500 copies – and users had to physically come to the library to use them. That was perhaps good for libraries, drawing people in, but the inefficiencies compared with the present system could never otherwise be defended. If I compare the 1980s and 1990s with today, the change is astonishing. For Elsevier alone, on our ScienceDirect platform, we have over 10 million regular users, who have access to over 8 million articles (back to vol. 1, no. 1) and we recently celebrated our 1 billionth download. Indeed, in 2006 alone there were 310 million download, or nearly 530 per minute on a 24-hour, 7-day-a-week basis. And that's just Elsevier.

Can't we now declare success?

- Access is nearly universal, including free access in the lesser-developed world.
- Usage is continuing to increase at extraordinary levels.
- The sheer amount of information readily available has increased enormously.
- The speed of publication is much faster.
- Functionality has improved incredibly.
- The efficiency of information access and retrieval for the researcher is much greater.
- And there is rapidly falling cost per use and increased value.

Researchers, librarians and publishers should be celebrating what has been accomplished.

Instead, we have a heritage of 20 years of frustration. The fundamental source of that frustration is that research output has grown – and continues to grow -- at a steady rate of 3% per year. Publishers have experienced some ability to cut costs and incorporate improved efficiencies but these savings are counterbalanced by the need to invest in new systems and technology, new back office and sales infrastructures and new product development. The result, as we all know well, has been publishers' prices rising faster than library budgets, creating a generation of finger-pointing and defensive responses, with frustration on both sides. As one library director friend of mine said several years ago, "I never expected to spend my career downsizing my collection."

So here we stand – uncertain partners or at least players who in fact share a common mission: we both exist to support the research and education process, to improve the effectiveness and efficiency of our shared customer base (faculty and students). We both care about making the world better and want to find appropriate roles in a changing world.

Consider the traditional formal research information process.

- It starts with the research itself.
- The results of the research are written up and reviewed.
- If acceptable, these reports of the research are published.
- The publications are distributed via marketing and selling.
- They are purchased or otherwise acquired.
- They are made immediately accessible, but are also stored and archived for future generations.
- Tools are provided for location, retrieval and use.
- And, in the end, the research information is incorporated into new research and teaching.

In this traditional process in print, the roles were pretty clear and generally not overlapping.

1. Research – researcher
2. Writing and review – from researcher as author to other researchers acting as editors & reviewers, with the context and infrastructure provided by the publisher
3. Production/publication – publisher
4. Distribution – publisher
5. Purchase – library (if you go back far enough, researchers themselves)
6. Immediate access -- library
7. Storage – library
8. Archive – library
9. Locate/retrieve/use – library & researcher, with secondary services provided by publishers and enabled by third parties such as Dialog, both via the library
10. Incorporation into new research & teaching – basically by researchers, with publishers creating textbooks

Everyone knew what they were supposed to do and the system worked pretty well. But consider what has happened to roles in the current electronic environment:

1. As far as roles played, research itself is largely the same – done by the researcher -- although the growth in interdisciplinary and distributed research teams means a change in the social dynamics and tools used by researchers
2. Writing and review – still from researcher as author to other researchers acting as editors & reviewers, but generally with a more sophisticated infrastructure provided by the publisher – modest change
3. Production/publication – for the formal literature – the authoritative copy -- this is still a publisher role, but also researchers themselves are producing electronic versions, even if only as preprints
4. Distribution – Publisher post to their websites, but we also see researchers posting results directly and libraries distributing via institutional repositories
5. Purchase – libraries continue to purchase, but now there are also “open” variants of financing models and a likelihood – or perhaps it is my wishful thinking -- of ultimately more end-user purchasing and departmental and research team acquisition
6. Immediate access – The publisher is rapid posting to its online platform, often immediately after acceptance, but the researcher is also providing access, often enabled by library
7. Storage – in print, libraries; electronically, for the version of record, storage is now generally now done by the publisher
8. Archive – again, in print a library role; electronically this is a shared responsibility of the publisher and library and third parties (Portico); this has surely been one of the greatest areas of focus for academic librarians: how to ensure preservation that the preservation of electronic materials occurs, but many scholarly publishers have also early-on acknowledged the responsibilities they have here as well and are working with libraries to ensure proper archiving.
9. Locate/retrieve/use – Publishers, libraries, researchers and new third parties all play a role, as platforms become more sophisticated, new tools emerge and specifically search engines play a greater role than the third party online hosts of the past
10. Incorporation into new research & teaching – researchers, libraries, publishers and new third parties such as Blackboard. It appears that academic libraries may particularly be redefining their mission to include the support of teaching as a core element.

Is it any surprise that we are confused or uncertain about changing and multiplying roles? Are we links in a chain, competitors or collaborators – or all of these? In the future it is possible we will be competing in new ways, as we try to provide workflow tools to researchers that extend beyond any area in which either the library or the publisher has worked in the past. Is it possible for us to redefine at least some of our roles in a collaborative – or at least not hostile – way?

Perhaps it is useful to stand back and look at the ways in which libraries and publishers differ – and then what we have in common. (And let me thank my colleague Tony McSeán, our Director of Library Relations, for highlighting these differences and similarities to me.)

There are at least four ways in which we differ.

1. Our planning horizon – Typically today, publishers think of a two-year planning horizon, perhaps three. When I started with Elsevier in 1976, I was hired to do the first round of what was then called “long-term planning”. We did five year plans. And that was not unreasonable at the time. By the 1990s we were making major decisions that absolutely required a long-term perspective to justify the level of investment. Consider that we have invested over \$400 million in the last decade in ScienceDirect and Scopus alone and their underlying production and archiving systems. Today, however, all bets are off and – for better or worse – we work basically on a two-year planning horizon and review to “strategic planning” rather than “long-term planning”. Yes, some investments are framed in a longer time period, but the basic premise is that there is relatively little that one can count on beyond two years.

By contrast, my impression is that academic libraries’ planning horizon is much longer. I see libraries thinking still in five or ten year increments. This is perhaps a function of the university environment itself. I have had the privilege of being a part of two recent “academic library of the future” meetings and my sense is that the quality of long-term thinking that is going on is significant. And perhaps taking the longer-term perspective is exactly what is needed. But there may also need to be a sense of urgency.

2. Publishers and libraries may also differ as to their view of budgeting and costs. For publishers, costs are not always an absolute but something one considers relative to revenues. If a new product will generate sufficient revenues, the budget can generally (not always) increase to make the necessary investment. Our budgets have much more ability to be redefined annually – with winners and losers in the process. Few libraries have the same financial flexibility, unless there has been unusually successful fund-raising.

3. In fact, flexibility in general may be a difference between publishers and libraries. We can change location, outsource staffing and redefine our business. Libraries have far less flexibility in these matters.

4. Finally, we likely have a different ability or attitude toward risk-taking. Publishers have more access to serious resources and an ability to invest heavily on risky ventures. Ironically, having said that, I would also say that we are often among the most conservative and cautious of business investors.

Despite these differences, publishers and libraries also have much in common. First, we are faced with a similar set of challenges:

1. We are losing – or have lost – our captive audience. Just as it is no longer necessary for faculty and students to come to the library (or to acknowledge that much of what they access on their desktops is only there because of the library), neither is it certain that authors and readers will continue to rely upon publishers as intermediaries.

2. We face the potential disintegration of the structure of what we do. At a recent ACRL workshop one question that struck me was: What are the problems that the university administration looks to the library to solve? Storing physical materials and providing access to them is surely no longer a satisfactory answer. Publishers have to address the same question: What are the problems that researchers and educators look to the publisher to solve? If we are not providing answers to contemporary problems and needs, we face the possibility of being bypassed.

3. We both have a need for new and different expertise. For example, within Elsevier we have information specialists – whether search, text-mining, user-centered design – that a publisher never would have had in the past and large libraries have counterparts.

4. And we both need to redefine our businesses. To return to our shared mission: to support the research and education process, to improve the effectiveness and efficiency of our shared customer base. We must do that in new ways.

In addition to challenges, we also have strengths in common.

1. We both are passionate about quality.
2. We both are adaptable and open to innovation
3. We both are collaborative – and are used to working together
4. We both have a network of contacts within the research community – common customers and a goal of not only supporting them but of delighting them with the services we can and do provide

So, how do we work from our collective strengths in a collaborative rather than competitive way? When we look at our academic customer base, what changes do we see in an era of rapidly changing technology?

1. Researchers have formed e-communities and some work within the world of e-science. There is a renewed emphasis on data and distributed computing.
2. There is instant communication within the research group. For students, e-mail is seen as old-fashioned.
3. There is a premium on quality – “great stuff”. This is expected to be available quickly, easily, everywhere, anytime, at no perceived charge
4. The community is more literate, self-sufficient, with a self-service culture
5. There is a resistance to paying for enhancements – what I call a PC mentality, i.e. that every year there should be more features and functionality at the same or lower price -- leading to the question of how to monetize new services
6. There is both an appreciation and a frustration with the abundance of information – much of it unauthenticated. How to find the best, most relevant

information? How do you authenticate the author of a work found on the Web, particularly as work is increasingly in interdisciplinary fields, moving the researcher outside of their known circle of colleagues?

7. There is a desire to get access to everything but also to be able to cut, paste, combine and recombine information. How do we make this easy -- and legal? How do we make it affordable?

If these are some of the attributes of the environment we face, where might we look for collaborative areas? I'm going to mention seven areas -- in no special order -- where we could at least explore working together and I know there are many more.

1. Data storage -- The management of large data sets is now getting much more serious attention internationally. Generally this has been left to authors and there is no permanent preservation. Publishers have little experience here as to data per se, but we do understand the creation and preservation of information objects. Institutional repositories have been suggested as a logical home for data sets. Perhaps there is room for co-operative work here. One nagging question: who will fund this? And will researchers support these efforts? We did a survey not long ago that reported, no surprise, that twice as many researchers want access to other researchers' data than were willing to make their own data available. It has also been suggested that this is perhaps the area that funding bodies should be focusing on as to preservation and access rather than focusing on the journal articles. Preserve the data that results from the research you have funded, so it can be explored and evaluated further.

2. Improved search and filtering tools -- We all are more than familiar with the overload of information that is returned by search engines. Some do a better job than others -- I'm not hesitant to mention the awards that our search engine Scirus has won -- but there is still room for better tools to help filter search results to get the really relevant information.

3. A related area: text-mining -- There has been a growing interest in text-mining of the full text of articles. Text-mining can be much more nuanced and productive than straight search. Are there ways we can work together to improve the possibilities and ease of text-mining? Ultimately we may need a fundamental rethinking of the organization of information and who better to work on that than librarians and publishers.

4. Rights clearance -- We need to develop systems to allow for the handling of information at a much more granular level. Four years ago, when he was a speaker at this symposium, Clifford Lynch of CNI said, "We really have very little experience in developing economic and legal and business agreements that allow computation, extraction, derivation, reuse, correlation, reintegration into other databases." He was absolutely right, of course, and I think it is a shame that so little progress has been made in four years. How do we move forward, including rights clearance at a micro level?

5. New economic models – There are few other things I could mention that would likely draw as great an agreement as the need for one or more new economic models. I noted with interest this week the announcement from the University of California that they want to negotiate with journal publishers on the basis of “value-based pricing”. It would be inappropriate for me to comment in any depth on their proposal, as I really haven’t had time to study it nor presumably have other publishers or librarians but I admit to being initially skeptical. But setting this specific UC proposition aside, the underlying issue remains: we need new pricing models.

6. A sixth possible area of collaboration: is there a way we can broaden the user community? We have had librarians express an interest, for example, in providing access to alumni. Not an easy question, but one worth considering collaboratively.

7. Finally, we can work collaboratively to do research on user behavior and user needs.

In conclusion, the future is too important for all researchers and students – our shared customers – to leave to chance. Librarians and publishers – and researchers – are taking new roles in the scholarly communication process. Librarians have been the traditional advocates of the user. Publishers have been the advocates of the author. Those roles continue – but they also are broadening, as publishers become much more closely intertwined with users and academic librarians engage their faculty in their role as authors and editors. We’ve become competitors in many areas – or, if not competitors, actors in overlapping roles.

But within this confusion there are new needs and opportunities to support the research and teaching process and some of these may be best met with collaborative efforts. We need new initiatives that can bring all parties to the table, with an open mind for balanced analysis. Among the areas we can look at: data preservation, new search and filtering tools, text-mining, rights clearance at more granular levels, new economic models, the extension of information distribution to broader communities and the study of user behavior and user needs. Let’s see if we can’t make something happen together. Thank you.