February 1997

Things That Keep Me Awake At Night

Karen Hunter
Elsevier Science

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

Part of the [Library and Information Science Commons](#)

Recommended Citation

Hunter, Karen (1997) "Things That Keep Me Awake At Night," *Against the Grain*: Vol. 9: Iss. 1, Article 15.
DOI: <https://doi.org/10.7771/2380-176X.1988>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Things That Keep Me Awake At Night

by Karen Hunter (Senior Vice President, Elsevier Science) <k.hunter@elsevier.com>

Based on a presentation made at the 1995 16th Charleston Conference, November 7, 1996.

Introduction

In order to put into context my list of things that keep me awake at night, I need first to describe briefly some of the things Elsevier Science has been doing about electronic journals.

As most of you know, from 1991 through 1995 we partnered with nine universities (more correctly, 17 universities, if you count each of the University of California campuses) in an experiment known as TULIP (for The University Licensing Program). In this experiment we provided initially 43, later 80+, journals in materials science and materials engineering in bitmapped electronic form. The electronic files, which included searchable structured headers and unedited fulltext ASCII generated from an OCR scan, were provided to each campus and each university provided its own implementation over the campus network. We all learned an enormous amount from this effort, which is documented in the TULIP Final Report. (The report is available via our homepage — <<http://www.elsevier.com>> and click on General Information to get to TULIP.) Before the TULIP experiment ended, we decided there was sufficient interest from other institutions to offer all 1100 of the Elsevier Science journals commercially in this same format, also for local implementation. This program is known as Elsevier Electronic Subscriptions (EES) and has been available since January 1995.

We have recently announced ScienceDirect, which is a new online Web host which will be available from July 1997. We will start this with about 300 of our life sciences journals and a 1700 title subset of our EMBASE bibliographic database. Over the course of 1997 the file will expand to include all Elsevier titles and, we hope, journals from other publishers. The functionalities of ScienceDirect are quite robust, including A & I and fulltext searching, the ability to create personal profiles and alerting services, journal issue browsing, links to editorial information and, in some areas, additional discipline-specific databases and related information.

Again, in order to understand the concerns I will go into in a minute, I would

like to mention four reasons for our undertaking ScienceDirect.

1. As other publishers do, we currently have several individual journals up on the Web. Each of these was lovingly developed in a hand-crafted way by its in-house publisher and technical staff. This is the way to begin, but one cannot do 1100 journals efficiently in this way. We need a uniform, scalable platform for electronic development. ScienceDirect will provide that platform.

2. We also believe that bringing all of the journals together in a database will create significantly more value for the user and for the library than having individual titles scattered at individual Web sites.

3. Building this service will permit

"The assumption is the library's going to provide all of these neat new services in the future."

much more robust searching and SDI access, particularly when linked to A & I services.

4. Finally, the goal is to make it easier for scientists and students to do what they have to do — to improve their efficiency — and again we think a more consolidated service will do this.

And one final paragraph of introduction: what are we planning for the future? We are not just moving traditional paper journals from paper to electronic form. We have other ambitions, including improving the speed of publication (and separating the timing of paper and electronic publication by releasing the electronic first). We are already adding what we call "non-printable items" — data sets, links

to other Web sites, video, audio, 3-D graphics and animation, computational facilities, etc. We also plan to develop special services with clusters of journals and other community information for specific disciplines.

Concerns that Keep Me Awake At Night

1. Building the Infrastructure

You don't make these changes simply by waving your hands, particularly not when you are publishing 1.8 million pages per year and have to keep the present processes running smoothly. We have been retooling the organization for five years. A critical part is the change from traditional typesetting aimed at a print output (and one-time use of the input stream) to a production process which results in media-neutral databases which can be distributed — and reused — in a number of ways. We are at the end of a five-year process to convert to such a system. One step which was done this year was to build our Electronic Warehouse — our database storage and management system for the SGML and PDF files and for all illustrations. Given the resolution we are using for archival storage of illustrations, the last estimate I saw was that the annual storage requirements for this database would require 750 GB per year. A significant investment, to say the least.

At the same time we were building the production side, we have also developed our international sales and support structure, known as Regional Sales Offices. Paper journals are a wonderful product requiring relatively little sales effort or technical support. Buyer and seller understand the paper journal, so a new journal

will be sold by promotional mailings and sample copies. The journal issues themselves go patiently through the postal system. Everyone knows how to "access" them. And there is no negotiation on the price — it is basically a take-it-or-leave-it offering.

Electronic journals, as we have all learned, are quite different. Sales may have to be negotiated, licenses signed. What you get is not so clear cut — is the content the same as the paper, does it include software, what do you have in Year 2, etc.? The sale of electronic journals needs more explanation, technical support and sales support. And the caliber of person you need to hire to do this is higher and

continued on page 41

Things That Keep Me Awake from page 40

more expensive than in the paper journal environment.

Which leads to the question: what is the role of the agent in all of this? Hard to say. Right now we think that as we make the transition, we need to stay in close touch with our larger accounts in particular. But we cannot afford to negotiate with all customers, nor do all want to negotiate with us. That means that at some time we may welcome agents as part of this process.

2. Virtual Kiosks vs. Digital Libraries

I said that one of the reasons we are building ScienceDirect is a concern over the inherent inefficiency of individual journals residing at individual Web sites. What could evolve is a change not from the well-organized libraries of the present to the hoped-for digital library of the future, but to a world of thousands of virtual and essentially invisible information kiosks — URLs which may or may not be permanent and which require a new access procedure for each separate title. We do not believe this is an improvement over the present system. But others are creating such an environment, with each title a world unto itself. And I also wonder if we aren't creating a generation that believes indexing and searching means using Yahoo or Lycos. The information community has created well-refined tools which risk being lost in the world of the Web.

This means that I think there is a role for the aggregator — the party (or parties) which pull together in one organized space the publications of many publishers. Agents would like to play that role, as would traditional online hosts such as OCLC. ScienceDirect is also able to function in that way.

3. Creating Links

If I think there is a role for aggregators, by definition I am saying that bodies of information need to be related and linked to one another. We want to create secure gateways and seamless links, something more easily said than done. At the lowest common denominator, if all you want to do is move from one server with publicly available information to another, it is easy. You just put a hot link there, you click and you're gone. But if the information is security protected and you need a subscription or an access code, what you'd like to be able to do is to pass customers through from one system to another without having to re-enter a password or a whole new authentication procedure. That's not so simple, but we're determined to do it over

the long term.

We'd also like to connect the fulltext information on our server to locally-mounted abstracting and indexing services. What you gain is the ability to leverage your investment in your local searching of an ISI or Medline or INSPEC file. If you have a subscription to the fulltext, then from the local search you can retrieve the fulltext PDF files held remotely. That's not too hard to do. But you lose a lot in making journals into solely a document retrieval service. What studies are available of the use of electronic journals suggest the browsing of the issue still has significant value. The concept of the journal is still valid as more than a random collection of articles.

Another way of making links is between our journals and other relevant material. The easy example is the linking of references (cited articles) to the abstracts of those articles. Ultimately, as the files build, you'll be able in many cases to get to the actual fulltext of the referenced article. And, if the system is properly built, in the future you will also be able to take an article published a year earlier and link to where it has been cited. One of the key roles a publisher should play in the future is creating those links — adding value by

"We've particularly become much more conscious of student use than we ever have been before."

integrating information, letting people maneuver through the space and get a full range of information. As we talk with scientists, this is what they want. It is doable, it is possible — it's just going to take a lot of work.

Finally, we'd like to be able to let the information that's part of our service be easily linked and integrated with the reader's own personal files. This means downloading in formats which work easily with commonly-used file and reference management systems.

4. User vs. Buyer

When I talk about users, it raises another question which occasionally troubles me. Specifically, how in sync are the interests of the end users compared with the library's interests? We and many other publishers are certainly designing services which we hope are "market-oriented." We try to design with the end-user in mind. We do our focus groups with scientists and students. We've particularly become much more conscious of student use than we ever have been before. The fact that graduate students are the prime users of the general scientific literature on campus is something that I won't say escaped our notice before but, as they weren't our authors, we tended to focus more on the

faculty. So now we are trying to design systems that are appealing and useful to both students and faculty.

At the same time, we don't fool ourselves. These systems need to be sold to the library because, if you ask scientists how much money they're willing to spend on information or how much they actually spend now, it is not much. They total up their memberships in the societies and the couple of other journals they buy and come to a number and that number is consistent across disciplines (and not very high). Now, you don't know what will happen if you build a better mouse trap, but there is no inclination on the part of scientists to spend more. The assumption is the library's going to provide all of these neat new services in the future. The library's always gotten it for me, so they'll get it in the future. That means we are more stretched than we were in the past in trying to see whether we're meeting the needs of both our buyer and our user.

I would add a point here that, as I mentioned earlier, we're also looking at offering services which are much more focused on specific user communities. We could offer clusters of services, for example in neuroscience, and include not just our long list of journals, but also brain maps and links to genetics databases. When we talk with the specialist scientists about these ideas, they love it — there is no question about it. But is the library going to buy this sort of connected community for the neuroscience faculty? That does give me some sleepless nights.

5. Role of Consortia

Consortia — increasingly we are talking with consortia. I know they're here in the present and I suspect there's a consortium in my future as well, but I also sometimes wonder if it's just the flavor of the month. Consortia are a hot topic now, but will they hold up over the long term? It is not clear to me and it can have a dramatic effect on how we organize our business now and in the future.

One of the things that comes with consortia is how you define a user community. We spend a lot of time on this when we do licenses. We need to define who has the right to access the material we are licensing. Any of you who have been through licensing discussions, and most of you have, know that this is not necessarily easy. We're getting better at it and I think that as a community we're all getting closer to understanding how to create definitions. Often, however, new problems arise as soon as you try to figure out how to implement the definition — by ID and password? By ever-problematic IP address?

continued on page 42

Things That Keep Me Awake from page 41

addresses? Maybe this doesn't keep me awake, but it disturbs my snoring from time to time.

6. Interlibrary Loan

The relevance of defining the user community is that you are also simultaneously defining who is not part of that community. Which means, what about interlibrary loan? And that's where the license discussion can get into trouble. In TULIP we had one of the universities that went through the whole experiment (and completed it splendidly) but with whom we never had a signed license because we could not agree on the interlibrary loan clause. They were unwilling to sign a license that said they couldn't use the electronic files for ILL and I was unwilling to sign a license that said they could. That won't happen in a commercial environment. We have to come to some agreement.

I've been part of the ILL discussions within the context of the recent Conference on Fair Use (also known as CONFU). While some of the issues discussed within CONFU, such as multimedia course work, have resulted in guidelines which both libraries and publishers can accept, ILL has been a non-starter. We have tried and tried and gotten nowhere, to the point where discussions were simply stopped. And that's an extreme frustration. It is also a personal frustration for me because I've tried to put forward new guidelines that haven't gone anywhere.

What is the issue here? If I look at Elsevier Science titles, there is not an article from our recent journal volumes that cannot be rapidly obtained through a legitimate document delivery service, in 24 hours if you want it. The people of UnCover will get it for you tomorrow. And, therefore, if you choose to go the route of interlibrary loan, what you're saying to me is "I don't want to pay the publisher for it." While I understand that, I can't condone it if I'm moving into an electronic age.

I'm being asked not just to allow a photocopy to be made (which is what the original CONTU guidelines are about — photocopying). I am ultimately being asked to permit "interlibrary loan" to become a hands-off, remote access system from a user's desktop, where the user comes into the electronic files of a supplying library with no people intervention by libraries on either side and pulls off the desired article. Journal publishers just can't live with the notion of one big lending library in the sky (or on the Net). It's a

major issue for publishers and it remains on my personal list of priorities. Publishers may be willing to do some compromising (individually or collectively) on printing ILL hardcopies at the supplying site from electronic files, but libraries must as well be willing to compromise on electronic delivery from electronic files if we are to reach agreement on new guidelines.

One final comment on this topic. When we did CONTU 20 years ago, it really was to create a safe harbor for the borrowing library. It was to say, here are the number of copies you can make without having to worry about getting permission. What we as publishers are hearing now is not from the borrower but from the lending university (and it is generally a university) saying that it would be ever so much more efficient if I would just let everyone come in and pull off articles. Yeah, I know it would be, but we're not in that business yet. This keeps me awake and I want to solve it.

7. Pricing

Pricing is an issue that keeps everyone awake. I used to say, as recently as a year ago, that we hadn't a clue on how to price. It's not that bad now. I think we've lived through a lot of discussions and we've

"It is not so simple ... to offer access to just a fragment of the file."

spent a lot of time looking at pricing. We've tried to understand how to do at least interim pricing of electronic journals, and I think we are reasonably comfortable about what to do for the next year or two. I don't think the community as a whole has a long-term model yet. We collectively may have some notions of long-term models, but it's too early to try to put them in place.

What I would like — in my dreams — is to find a way to effectively lower the price. I would love to experiment with the demand curve and see what would happen if, say, in April or May we announced a lower price which would take effect if there were enough subscribers — you know, \$500 if there were 1500 subscribers, instead of \$1500 with 500 subscribers. I'd love to try this, with some guarantees of pricing over two or three years, because it would take this long to really test if there was still a possibility of increasing subscriptions by lowering price.

Assuming that there is not such a direct experiment, what we can look for is ways of creating packages which offer more information and functionality for the same money (or a little more money). I think this may be possible in an electronic environment, although I'm not making promises. But it is on our minds — how to deliver more and how to provide some long-term price assurances.

8. Subscription or Transaction

All of the pricing discussion in part begs a fundamental question. Will we be perpetuating the subscription model or are we in a state where at some point we're going to have to bite the bullet and live with selling by the drink? Publishers like the subscription model, but we are also considering some alternatives. What about a subscription to a pre-paid group of transactions where those transactions go across a group of journals — or you can create a virtual journal in a particular subject field. Again, that's why I want the database of 1100 journals and not just individual journals sitting on the Web. I want to be able to say that if you really want to access all of our environmental material regardless of the journal that it's in, there may be ways we can work that out.

9. Parallel Publishing

One of the other things that worries me is parallel publishing — paper and electronic. All of us are suffering under the financial strain of doing two things at once. We'd like to say for a given journal it is either in paper or electronic, but the reality is we're living in a world of paper and electronic, and supporting two systems from your side and from our side is very expensive. It is not clear when that's going to change.

10. Short-term Archival Access

That leads to the short-term archival access struggle. Right now most of what Elsevier delivers electronically is delivered locally, so the archival question is moot: you have the files and own them. But as we move more into an online environment, we have to deal with the fact that people's collections will change and, when they cancel, what will be their access to the canceled material? Ideally we would like to say you continue to get access permanently to what you once subscribed to. But is it realistic that this be permanently available without additional charge, as we are continuing to have costs associated with your access? As years go by, you have School X and University Y who five years ago subscribed to this one journal for two volumes — it is not so simple (or clear to users) to offer access to just a fragment of the file. We're in discussion with customers now to try to craft an effective solution for this issue.

11. Long-term Archives

Which brings me to long-term archival questions. I was on a national task force last year on digital archiving. I think the Task Force was successful in identifying the issues, but not in identifying answers to those issues. This is a discussion which

continued on page 74

Barlow Interview from page 20

political and economic dominance, or, more accurately, continue its dominance.

JPB: Not necessarily, if you look carefully at what happened in the industrial period. It was the northern Europeans who had a syncretic time-based culture that was naturally conducive to allowing itself to be turned into machinery. If you are punctual by cultural nature you don't mind being an interchangeable part in a large organization. But most everybody else on the planet suffered because they had a completely different view of the nature of time. As we get deeper into the Information Age, time will be irrelevant.

ATG: Do you see a role for publishers in the Age of Information?

JPB: There is some role for publishers. I'm not saying they are going to go away. But what publishers continue to do is to put human thought in a physical format and ship it around in trucks to a place where it is put on a shelf where somebody buys it as if it were some kind of mercantile object. You don't need to do that anymore. It's now possible to take anything that a human

being can do with his or her mind and reproduce it infinitely and distribute it infinitely at zero cost. That is going to make for a big shift, so publishers will have to reexamine their role in the information age. I've talked to publishers about where their future lies. They aren't doomed. They just have to think a lot about what they do that is still going to be useful.

ATG: So how would you characterize the role of the librarian in all this? Should we be anxious or is it an exciting time to be a librarian.

JPB: Oh, definitely, yes...an exciting time because people are going to need librarians more than ever. In the future, people are going to have to use libraries more and more for economic reasons, not just for personal enrichment. Libraries are going to be much more central because people are going to be making money on the basis of what librarians can do for them. You are going to be essential to them.

Ron Chepesiuk is Special Collections Librarian at Winthrop University in Rock Hill, SC. He is also the author of six books and 1,200 plus articles that have appeared in *The New York Times*, *The Progressive*, *Modern Maturity*, *USA Today*, and more than 200 other publications.

Things That Keep Me Awake from page 42

will go on for some time, as there does not seem to be an answer in sight. Who is going to create and maintain the digital archives of the future? Are publishers going to do that? Would you trust us to do it? Elsevier recognizes that it's vital that an archive be maintained and will ensure that there is an appropriate archival home for our electronic material. Will we be the one to do that archive? I honestly don't know yet. Libraries...governments...it's not clear yet. It is expensive to archive and to continually refresh files as technologies change. How is it going to be paid for?

Once you even decide all of that, there are still issues relating to access to the archive — how can it be used? Publishers and librarians had some wonderful theological discussions last year as to whether things on paper could be scanned and put into a digital archive. And the publishers didn't have — surprising as it may seem — too much problem with the notion of scanning for preservation — so long as no one had access to it after it was scanned!

Not surprisingly, that notion didn't fly with librarians.

The other issue here is to ask what the definitive edition is, anyway. I said earlier that we're creating these non-printable items. There will be an increasing divergence between the paper and electronic versions of a journal. We have a new journal called *New Astronomy*. It's got video and it's wonderful. It comes out first electronically, then there will be a paper version, but they're not going to be the same, and that's true of other journals from other publishers.

12. Where Is the Money Going to Come From?

We can ask now the last question: where is the money going to come from? We know that budgets are more than strained. We also know what it costs on our side and it costs a lot to get this stuff up and ready for electronic access. There is so much more technically possible now than we have the money to do, so we're having to make choices all the time. And I mean the whole community. Which means we need to set priorities, and the process

of setting priorities is not clear to me. I don't think the dialog that needs to occur between the libraries, the agents and the publishers (and the students, faculty and administration) about where we should spend an increasingly finite set of dollars is taking place. How do we collectively decide on what is the highest priority? If there is a way we can do that, it desperately needs to go on, because each of us now is operating independently.

IN CONCLUSION....

So in the end what I worry about is: are we too early because we're building a very complicated system? Or are we too late because too much has already happened in other ways and maybe we're overbuilding? Are we simply misdirected? Or we doing things just about right and on the correct path? All of the above...none of the above...and then I get to the point I ask, what was the question again?

Which leads me to ask you to count with me. One sheep, two sheep, a third sheep — oops, here comes a goat, two cats, three donkeys and an elephant. Rats, I'm still wide awake.

Back Talk from page 77

Jack replied, "He hadn't been making all of his house payments on time. I just got the bank to threaten foreclosure and he was more than willing to take Angela's offer."

4. **The Plight of Libraries Programs.** "I thought I would save the best to the last, I just hired an old friend from the Firm. Got himself a nice job at Princeton to help libraries. He has given a bunch of papers and writes articles about the money problems facing libraries. He is going to get the college presidents' association to agree to a uniform percent of their total budgets so that libraries can get their periodicals budgets back where they need to be." And he added with

a smile, "back to where we think they should be."

Alfred was impressed, but asked "But what about Chuck? It won't take him long to get on to you and we can't afford to take any heat on any of this." Jack got more serious and said simply, "Chuck is out of hand, but you don't need to worry about him." Alfred motioned with his hand and the expression on his face that he had heard that before: "That is what you said when he was supposed to take early retirement." Jack replied, "The twit turned it down. He doesn't seem to have a life beyond the library." Alfred looked him in the eyes and said "Well, the problem is yours, take care of it before our next meeting."

(More in next year's ATG annual issue)