Learned Publishing (2005)**18**, 131–142

HighWire Press: ten years of publisher-driven innovation

John Sack
HighWire Press

ABSTRACT: An account is given of the beginnings and development of HighWire Press at Stanford University and the philosophy that sustains it.

Introduction

The HighWire story is about several not-for-profit publishers and scientific societies who took the risk to do something very new and very different. They took this risk because there was a new technology – from the perspective of 1995! – that presented real opportunities to change the method of dissemination of research data and results. Their insight was that the World Wide Web would solve some of the problems that were hampering research communication: lack of speed, rigid format and rising cost. There was also the sobering insight that the web would present new challenges for the way in which the tasks involved in publishing would be paid for, but the innovators saw that as a challenge to be met, not a fundamental flaw in the technology or concept. And, curiously, the US Postal Service had as much to do with the beginning of HighWire and the *Journal of Biological Chemistry* (*JBC*) *Online* as did the web technology; the print *JBC* was becoming too expensive to mail because it was overweight. Early 1995 brought together the necessity for innovation, as well as the opportunity of the technology and the insight of editors and others about how it was to be applied. Just as 'means, motive and opportunity' means a conviction in the TV show *Law & Order*, 'necessity, opportunity and insight' gave us the conviction to launch HighWire Press at Stanford.

The innovators who had the insight were researchers, journal editors, postdoctorate students, librarians, programmers and publishers. But it was not these people as individuals who made it happen; it was these people focusing their separate talents and perspectives on the same goal: to change how science was communicated. 2005 is the tenth birthday for all of us who worked together to begin the changes – the second- and third-order implications of which are

John Sack

still being discovered in concepts like mass digitization, publication on acceptance, institutional repositories and open access.

Early days

The story of the birth of HighWire is concurrent with the story of the birth of *JBC Online*. It is a story not only about technology and insight, but also and primarily about people: there was among the founding team a 'demo or die' commitment that matched their cleverness, insight and skill. We made 'demo or die' our mantra and even called ourselves 'DDT' for 'demo or die team', and to this day staff meetings at HighWire are called 'DDT'.

'Demo or Die' was originally a reference to our commitment to demonstrate an online version of *JBC* at the Society's annual meeting in San Francisco in May 1995. We knew the opportunity to demonstrate a breakthrough to several thousand scientists came only once a year, and if we did not have the demo, the project would, essentially, die. It was late January 1995 and we had only four months to deliver the product – not just a mockup, but a functional product, and no one really knew what the product should look like or how it should function. That was all part of the design debate that had to take place before we could begin programming.

we did not realize how much technology we had to design and implement

Even now, after ten years of maturing our processes and technologies at Highwire, it takes HighWire longer to launch a new journal site online than it took us to build *JBC Online*! Fortunately, at the time we did not realize how much technology we had to design and implement. Ignorance was bliss, at least for the first few weeks of the project.

Development

In this paper, I describe some of the early insights and anecdotes that were part of the project, and some of the follow-on discoveries and differentiators we developed as we went from *JBC Online* to develop online journal sites for *Science* magazine and its careers and news sites: *Proceedings of the National Academy of Sciences*, *Journal of Neuroscience*, *Journal of Clinical Investigation, Blood, The EMBO Journal*, *Journal of Nutrition*, *Journal of Histochemistry and Cytochemistry*, *Molecular Biology of the Cell*, and the journals of Rockefeller University Press, Cold Spring Harbor Laboratory Press, and the American Society for Pharmacology and Experimental Therapeutics (ASPET). All this was within the first two and a half years.

Perhaps it is humbling to begin the story with the first major mistaken assumption we had (since Bob Simoni has noted some of the mistaken assumptions that the ASBMB had when it launched *JBC Online*[1]). We believed that the follow-on to a successful launch of *JBC Online* would be the development of perhaps two additional online journal sites per year. While we had expected that HighWire's growth would come via word-of-mouth discussions of journal editors with their colleagues on other editorial boards – and that decisions ongoing online would be driven by the scientists that ran journals and societies in these early days – we had not accounted for the fact that these editors talked to each other very frequently, were very decisive and were also competing with each other to be first to implement services they knew were revolutionary for their authors and readers. I was more used to the time scale in which decisions get discussed and debated within a university like Stanford. HighWire decisions were made with what we later came to know as the speed of Silicon Valley entrepreneurs in the Internet boom.

Perhaps two anecdotes illustrate this editor-driven aspect best.

1. After the successful launch of *JBC Online*, we contacted *Science*. A new editor-in-chief, Floyd Bloom, based at Scripps in La Jolla, California, was interested in the possibilities and we met with Floyd and his staff in Washington in the summer, and in August we sent them a proposal for a six-month development project. I then went on vacation to stay with friends in San Diego, California. While on vacation I got a call from Ellis Rubinstein, then the editor at *Science*. Ellis said they wanted to go ahead, and I was glad to hear that of course. But Ellis said 'six months won't work; what can you do for us in six weeks?' Apparently, rumour had it that *Nature* would launch its website in

eight weeks. I got in the car, drove to Floyd's lab, and we spent an afternoon parsing the print journal into a website. Eight weeks later, we launched *Science Online*. It was a fortuitous coincidence that the location of my vacation and the location of Floyd's lab were only a few miles apart.

The second mistake we made came home to us at this point, the week after we launched *Science*. We had assumed that, having put the largest weekly journal in the scholarly-journal world – *JBC* – online, that something as 'small' as *Science* would be much simpler. What we did not understand – which every printer and editor would have known – was that the variety of content in a news magazine such as *Science* outweighed in its complexity the sheer volume of all the articles in *JBC*. Fortunately, the managing editor, Monica Bradford, and the art director, Amy Henry, were used to 'rolling with the punches' and we were able to make decisions instantly to keep things moving forward with each weekly issue.

It was very important to our rapid development and deployment approach that journals and publishers would accept what we called the 'iterative' model of design and development, showing successively refined implementations, gathering comment and then cycling. This worked very well with both *JBC* and *Science*. Editors and management were very comfortable working by telephone and email; we simply had to add working by web browser to the mix. Even time zones did not seem to matter much, since both the deputy editor of *JBC*, Stanford professor Bob Simoni, and the editor-in-chief of *Science*, Floyd Bloom, were on the West Coast. But it was very important to our speed that people we were working for and with felt sufficient trust that they did not need to have frequent face-to-face meetings. Undoubtedly, this was only possible because HighWire was part of Stanford University, which already had a reputation for knowing what it was doing with science and with the technology of networking and the Internet (Google was getting started across the quad from HighWire at the same time, but of course wasn't known until the Google website went public in November 1995).

2. Soon after the launch of *Science*, I visited the laboratory of Sol Snyder at Johns Hopkins University. Snyder was the founder and chair of the publications committee for the *Journal of Neuroscience*. I talked with him and a colleague for an hour, talking about who HighWire was and what we do, and then left to return to Washington by train. I called the Society for Neuroscience's executive officer, Nancy Beang, on the phone from the train and asked, 'What are the next steps?', thinking that a series of discussions and proposals would follow. Nancy said, 'Sol's decided. Now we need to get it done.'

one thing everybody agreed on was that the print would go away

But not every meeting with editors and society-committee chairs was quite as sure in its outcome. When it came time for the ASBMB's finance committee to meet to determine the 'business model' for *JBC Online* – *JBC* had been freely available all through 1995 and 1996 – many models and costs were considered. The problem for the scientists making the decision was that there was no real data to discuss; that, however, did not dampen either the discussion, or the strength with which opinions were held. Later, one commentator on the meeting said, 'Put a necktie on a scientist and he thinks he's a businessman.' Through the consensus-building leadership of *JBC*'s editor-in-chief, Herb Tabor, *JBC* decided to go with an IP-based model, in which entire campuses could have concurrent, unlimited access. As Bob Simoni has written,[1] the chosen model was to 'unbundle' the print from the online because one thing everybody agreed on was that the print would go away and so the entire operation had to be sustained at least in principle by the revenue from the online journal. Further, institutional access was only for on campus use, not off campus; off-campus use was to be the driver for individual society memberships. And a 'campus' was to be considered a contiguous geographic location. Some of these decisions turned out to be popular and the way the market went, and others were not.

The most surprising response came from librarians. We had consulted with librarians before the subscription model was selected, and they had uniformly indicated that they wanted 'unbundled' management of online from print. But perhaps the pricing con-

sequences of this model were not clear to librarians at the time, since many complained, as Bob Simoni has noted, that they felt they were paying twice for the content, if they subscribed to both media. It also became clear from the early success of the Big Deal that librarians were not consistent across the profession in their belief that unbundled pricing was ultimately the right approach. A further problem we discovered with unbundled pricing was that only the most diligent of librarians – usually those in discipline-focused branch libraries – would change to or add an online subscription. The rest – whether because of time, attention or cost – were not yet prepared to watch for adding new-media versions of content they were already acquiring in print. Thus, the uptake on an unbundled online subscription was painfully slow, even for 'must have' journals like *JBC*. Even the communication channels that were needed to get the right people aware of new subscription options were not clear. The journal mailing staff would print e-journal subscription notices on the mailing insert that contained the print-journal shipping information, but did not realize that this information was usually discarded in the library receiving room before the print journal was sent on its way to the current periodicals shelving.

we knew from observing students and researchers that they wanted to read articles, not journals

Technical and business decisions

HighWire and its publishing partners made some early technical and business decisions that differentiated them from others who were working on similar projects at the time. HighWire was, of course, not the only technical or business enterprise working on this problem. However, we were one of the few university-based (and thus not-for-profit) groups engaged in bringing journals online. Other universities with related work going on include Johns Hopkins, University of Chicago and Cornell University. HighWire's non-profit status allowed us to take a mission focus rather than a financial focus from the very start. However, we had the mandate to become financially self-sufficient within eighteen months – a mandate we met.

Technical decisions

Our major early technical decisions included the following, none of which were foregone conclusions in 1995.

Be web-based

It was not obvious in 1995 that the World Wide Web would be a reliable medium to transmit highly complex content in scientific journals. Mosaic was a new term to most people. One had to say what a 'browser' was in a conversation with non-technical colleagues. OCLC at the time was, for example, working on a proprietary viewer called Guidon, which was the way it would present page images from journals.

Be article-oriented, not page oriented

We decided that the unit of interest to a reader was an article, not a page. So we would present articles to users, not individual pages. We felt the page metaphor was based in the medium of print, not the intellectual unit of the article. (Of course, the article itself would begin to mutate soon enough.) This meant that we would take a bit longer to display an article than a page-based system would take to display a page, because there was much more information to download; and early browsers did indeed have trouble with the number of bytes of content in a typical *JBC* article. But once the user had the article on his or her workstation, the interaction was a natural one for a PC user: the document was local and the PC power could be used to manipulate it.

Be article-oriented, not journal oriented

We knew from observing students and researchers that they wanted to read articles, not journals. There were very few journals that any one individual would scan from cover to cover – perhaps three journals per person – but they frequently were reading articles from a dozen or more journals. We decided there were two consequences from this observation: (i) access controls had to operate at the article level, not at the issue or journal level, so that some individuals might have access to one article, but not to all articles in an issue or journal; and (ii) we

built the system with the assumption that readers might find their way to articles via all sorts of paths, since we had already seen in 1994 the rise of copying of URLs into emails, and the inclusion of URLs that would go directly to an article into web pages and text files. We also saw a corollary that said we needed to provide navigation from any one page to all the major functions on the site.

Use HTML, not images

We decided to present articles as HTML text, rather than as images (such as JPEGs) of page layout. We knew this was true to the best design for usability and malleability of the web, but we also knew that this would give us the greatest challenges in presenting complex scientific material such as special characters (e.g. Greek), equations and tables. So for these three in particular, we embedded images in the HTML.

Offer PDF for printing

We saw how ugly HTML text was to print out, and how many more pages of paper it took than a beautifully composed printed page. We decided that it was not a case of either online or print, but that readers would want both. From earlier projects we knew of the Acrobat PDF format and felt it would be successful because of Adobe's backing. We encouraged publishers to have their printers create PDFs which we would provide to support user's requirements for printed output. Some other online-journal projects either were PDF-only (essentially putting the print and only the print online) or avoided the printed-page simulacrum to forestall any cannibalization of print.

Make links

Links-R-Us. Instinctively, those of us on the team who were working with the web knew that the 'magic' of the web was not just its ability to deliver pages on a desktop, but to connect pages you knew enough to find to those you did not know enough to find. We wanted to exploit this linking power to the fullest. We also realized that every article should link to all articles related to it, not just through cited references but through common authors, common subjects, etc. The mantra for this system design was 'every article is a portal to all the other articles related to it'. Just as the system had navigation buttons in each article, it also had navigation elements gathered in a sidebar (which we called, inelegantly, the 'content box') that took readers to all related articles and services. Unfortunately, at the time, there were not many other data sources on the web to link to, except for linking to other articles we were placing online, and to a system called Entrez. What was not obvious at the time for some developers of similar systems was that the power of the links to send traffic to your site was greater than the power of the links to send traffic off your site.

we also realized that every article should link to all articles related to it

Link to NCBI Entrez

Entrez, which later evolved and enlarged into PubMed, was a key resource for us to tie to. Any cited reference in a biomedical journal should 'resolve' to a reference with an abstract in PubMed. We knew that researchers would be very pleased to be able to read the abstracts online. And, working with David Lipman and Jim Ostell and others at NCBI (the developers at NIH/NLM of PubMed) we developed automatic link resolvers, automatic deposits so that publication metadata would be electronically included in PubMed in about 24 hours from publication, and (key to our publishers, and to traffic building models) links back from Entrez/PubMed to the journal articles.

Emphasize user-centred design, and usability–validation

Some of the members of our original team had academic training and practical experience in what Apple Computer had called 'user-centred design'. In this the designer begins with the tasks the user has to accomplish rather than from the objects the system has to offer. For us this meant we did not attempt to imitate a print journal online (most other systems at the time were designed around the print metaphor, i.e. the designers goal and the publisher's comfort level were to put the print online), but instead we

observed actual users to see what they did with journals. And it was those tasks and actions that we sought to put online. As we developed the designs, we showed our early implementations to postdocs in biology labs, and found they were quick to respond and very opinionated. One postdoc, Todd McGee, offered us so many opinions and suggestions that we hired him from Bob Simoni's lab; ten years later, Todd is one of our system architects and is still quick to respond and very opinionated.

The proof that we had a user-centred system design was visible in the first public demonstration of *JBC Online* in May 1995. We had set up several workstations in a convention centre booth, and invited scientists to give the system a try. We offered no documentation or handouts, believing the system should be self-evident if its design was good. We had loaded the five most recent weekly issues of *JBC* into the system; because of the usual mail lag, these were issues of articles that almost no one at the conference (except *JBC*'s editors, perhaps) had seen. What we observed was that the scientists simply used the system naturally and did not notice or stumble over the 'interface'. What they did notice was the papers they had never seen were ready weeks before they might otherwise see them in print. The system was transparent to the user trying to accomplish the task. We were delighted. (One thing we did not anticipate for the convention floor was the great demand our demo would create for printed output of the papers that people were reading online. We had only one printer in the booth.)

the search functions of the system were actually more important to users than the ability to browse the current issue

Recognize that authors and readers and editors are the same community

This was a subtle insight, which lead us and *JBC* to the development of the 'Papers in Press' model which publishes final peer-reviewed manuscripts online within an hour of their acceptance by the journal's editors in the automated peer-review system. Because the readers were also authors, they were familiar with the style and status of such manuscripts. For a clinically focused journal, the same would not be true; relatively few clinicians are writing papers for journals. From a more basic business perspective, this principle meant that HighWire was working with its potential future customers whenever it was working with an editor or author; that is, the editor of one journal might be the chair of the publications committee at another society and would of course be a reader of the online journals we produced. Because the basic-science community was linked this way, HighWire was able to grow in its early years through word-of-mouth referrals.

Emphasize searching, not only browsing

Several members of our team had significant expertise in information retrieval. In part for this reason, but primarily from user observation, we realized that the search functions of the system – the ability to identify articles based on author names or words in the title, abstract or anywhere in the text – were actually more important to users than the ability to browse the current issue. From this insight – which was not what journal editors and publishers had expected – we knew to develop the site around a content database and a text-retrieval search engine. Most other systems at the time were developed from the browsing metaphor: select an issue to read, then scan the table of contents, then select an article. While we knew that the browsing function had to be supported – browsing current issues was the most important function of online journals when there were only a few issues online – we built issues up from the raw material of individual articles. In fact, in the early HighWire system, many functions were implemented by database searches, thus the table of contents of any issue could dynamically be assembled by making a query.

Benjamin Lewin of Cell Press (HighWire hosted the Cell Press journals for several years) remarked to me that this insight about the primacy of search was plainest to see in a journal like *JBC*, which – at over 100 articles and 1,000 pages/week of print, about the size of a phone book each week – was 'a database in an archaic format'. By this Benjamin meant that *JBC* was a database to which about 100 articles were added each

week; then someone printed out each week's new entries and mailed them to subscribers, because that was the only way to distribute a database in the 1980s. No one was expected to read all 100 articles, but instead to skim the table of contents and to keep their on-shelf database current.

The validity of this search-oriented design is now evident. Across a large number of journals, typically 60% or more of the usage of an article comes from search engines on or off the site, while 40% or less comes from browsing through issue tables of contents. As we've worked with Google to index the full text of all HighWire-hosted journal sites, the search-driven usage of articles has increased in proportion.

Offer a database of publishable articles, not only a shelf of completed issues

One of the frequent complaints we heard from authors and editors was the time it took for articles to be published; we understood from their comments that articles were frequently complete in final form but needed to await the printing of an issue before they were published. We knew from this that article-level publishing would come at some point when procedures that had been based for years around the need to bundle articles into issues for printing and mailing could be replaced by article-level procedures that yielded a complete, publishable article to go online. The implementation of this concept did not come for several years after 1995, when the American Chemical Society released its 'ASAP' (As Soon As Publishable) function, a few months ahead of HighWire introducing something similar for *JBC*.

The insight that readers read articles, not issues, was plainly obvious to us as we did our usability studies on the Stanford campus. We noticed that graduate students and postdocs rarely, if ever, had print issues of a journal at hand; they were all reading photocopies (this was before articles could be downloaded in PDF form), sometimes second-generation photocopies. The article in the hands of the reader was completely divorced from the issue it had been published in. And, as was plain to see as well, the photocopied article was also visually divorced from the journal it had appeared in. There was little if any branding at the article level to show what journal an article had appeared in. You had to read the fine print in the page margins to identify the source journal.

Provide all the articles, not just recent articles

In almost all cases where you have information retrieval from a database, completeness or at least broad historical coverage is important to the users of the database. The database grows in value as it grows in size. As a journal site grew to have more than just the last few issues online, the use of the site would increase as it was used for research, not just current awareness. Early on, there was debate about whether online journal sites should contain only the most recent few years' content, and whether older content could be deleted from the online site as new content was added (the consideration of this at the time was driven from a cost standpoint but also to keep interest in print subscriptions); we knew this would be a mistake and made sure that the technology and costs would allow publishers to grow the content store (we saw that disk storage was declining in price faster than publishers were creating new content), and that search and display functions would allow readers to search and select new content only, if they wished. This was certainly the case with *JBC*, and, since it was one of the two largest journals published, our database grew very rapidly. We also worked with NLM/NCBI to extract metadata from *PubMed* and copy it into *JBC Online* site, so that readers could search *JBC Online* back to the 1960s. We implemented this same approach to populating the backfile of a journal site with nearly every journal site we developed that was indexed in *PubMed*.

readers read articles, not issues

Create the archive. Research is 'out of mind if not online'

We observed that students and postdocs – people whose research careers had developed with or after the web and search engines were part of the landscape – were not making use of literature that was not

online. Literature was 'invisible' if it was not online. We also heard from senior scientists that they found it unfortunate that their early-career's work was ignored by students because it was not available online. As technology allowed, online content was extended from the metadata to the full text: *JBC* was one of the first STM journals to have its entire backfile – 100 years of *JBC* – scanned and loaded online (*PNAS* and *Science* now have similarly large backfiles online as well). Now dozens of HighWire-hosted journals are putting full backfiles online.

Write programs, not pages

Some systems for presenting online journal pages in the mid-1990s relied on people writing some number of pages of HTML code to display new-issue content to readers. We knew this 'boutique' web-page approach would not work for *JBC Online*: we couldn't type that fast! We knew everything had to be fully automated in order for each issue to be released online simultaneously with the print mailing. So, all of our pages were generated by programs. This required discipline from the compositors, who had to prepare the data files for the programs using identical rules with each issue. We were very fortunate to begin our work with a large compositor – Cadmus Journal Services – who understood the need for rigour and repeatability in file supply. Aaron Bigman and Fran Steck of Cadmus were excellent colleagues on this project, fully committed to the goals we shared among ASBMB, HighWire and Cadmus. Fran today works for HighWire, as a journal manager providing support to many of the journals whose Cadmus composition files she delivered to HighWire in the 1990s.

not all the early decisions that led to HighWire's success were technical decisions

Bring the site to the user, in order to draw the user to the site

In the early days of the web, many people had just developed regular habits of checking their email, but not yet regular habits of checking a website for a new issue once a week or once a month. We knew it would be essential to announce to readers that updates were available on a site, rather than expecting them to check the site periodically for updates. We developed two types of alerts: 'eTOCs' or electronic tables of contents; and 'research alerts', which allowed a reader to type in a word, phrase or name that the system would monitor for across all published articles and alert the reader on any match in any journal. The alerts would only give the user a citation; this would bring the user back to the website for abstracts or full text.

Provide better service through statistics

Quite a few years before COUNTER helpfully standardized a set of usage statistics to help librarians compare journal usage across sites, HighWire worked with librarians to identify metrics that would be of interest in collection development and management, as well as patron services. While all the COUNTER statistics are present in the reports that HighWire developed with librarians, the HighWire reports also help librarians and editors know something about the interests of the readers in general, and at a librarian's institution. Statistics are available on which articles are most frequently read, as well as which topic areas are most frequently viewed. Librarians can see whether usage is widely spread across an institution or a journal, or whether most usage is limited to a few individuals or to a few articles. (Of course, HighWire also offers all standard COUNTER journal reports as well.)

Business decisions

But not all the early decisions that led to HighWire's success were technical decisions. In fact, many were business decisions and were less obvious than the technical decisions.

Start with STM

Perhaps the most important business decision that HighWire made was to begin its work with a big challenge, one of the two largest STM journals, and the most frequently cited STM journal, *JBC Online*. It was important that we begin with a very large and highly cited journal, so that our work would be visible and important to a large number of

scientists. Indeed, almost anyone working in molecular biology would need to use *JBC Online* to keep up with their colleagues' (and competitors') work.

It was also important that we began with an STM journal because such journals were, in general, well funded. The high individual and institutional circulation of *JBC Online* meant not only visibility for our work, but also a strong sponsoring society which could, with Stanford, share in the significant funding of a risky project. Smaller journals, and especially those in the humanities or social sciences, would probably be unable to afford the costs of this new kind of venture.

Follow the mission, not the money

HighWire's mission from its inception to the present is to foster the success of society and not-for-profit journals as they make their way in a complex, risky and expensive technological and business environment. Our belief was that an excellent, shared, technical infrastructure would provide societies and publishers with lower risk and higher capacity to innovate, beyond what they could do on their own. Successful large commercial publishers had such a multi-journal infrastructure, of course; we sought to provide the same for societies.

Not for profit; not for loss; not for sale

We also knew Stanford's preferred style of innovation: seed money was available to help entrepreneurial ventures that served the academy, but it was only seed money, not a continuing subsidy. This led to another HighWire mantra, that we were 'not for profit, but not for loss either'. In the heyday of the Internet boom, when many startups were being bought by each other, we expanded the mantra to 'not for profit, not for loss, and not for sale'. We had to share with our customers the full cost of the work they wanted us to do.

Sell services, not software or subscriptions

We wanted to provide a service to publishers, but not to sell or license software, and not to license, own or produce any content. This service-provider model felt comfortable to most publishers because it was like their relationship to their printer. We did not want to be publishers, knowing that this could lead to conflicts of interest with our own customers, to be avoided at all cost.

Our publishers and societies made some of the important business decisions early on as well. Today, with perfect hindsight, it seems hard to imagine there were debates, but there were.

Be online on time

Every publisher quickly agreed that it was essential that the online journal be posted and available at the same time (or before) the print was mailed. Some other online journals felt that a delay in the online posting would help sustain the value of print from erosion through online use.

Offer abstracts for free

All but one publisher (no longer with HighWire) agreed that it was essential for abstracts to be available to non-subscribers. In part this decision was based on free availability of abstracts in Entrez/PubMed, and in part it was based on the belief that abstracts could be given away without harming the subscription base.

'not for profit, but not for loss either'

It is probably important to note here that HighWire does not control publisher business models. Each publisher makes its own decisions about the models to use. This was important in order to ensure experimentation with a variety of business models in the early days of Internet-based journal services, rather than too soon deciding which model was 'right' and implementing only that model. In a conceptual sense, a HighWire-hosted journal site is owned by the publisher, not by HighWire. HighWire is responsible for the technology, but not the business of the site.

Provide cited-references' full-text for free

One early innovation, still not matched in other systems that I am aware of, is the decision by publishers that the reader of a full-text article could obtain the full text of a second article directly cited in that first

article without requiring a subscription. These were called 'reciprocal toll-free inter-journal links'. Publishers realized that they were a limited risk. Such articles were only made free by a direct link, and those articles that were made free were typically a year or more older. Readers loved this capability, using the word 'seamless' to describe it. It is the HighWire innovation most often mentioned by readers.

Offer alerts free

After some initial debate about whether alerts were a value-added service to subscribers, or a form of advertising, all publishers eventually agreed that alerts were to be available free to anyone who wanted to sign up for them. In addition, publishers recognized the value of the 'research alerts' was significantly greater when it was applied to many journals, rather than to a single journal, and they agreed that they would share in cross-journal alerting.

Open the archive

After several years of content was online, Nick Cozzarelli (*PNAS*), Bob Simoni (*JBC*) and Michael Held (Rockefeller University Press) presented a concept of 'free back issues' to their colleague HighWire publishers. Their view was that librarians and researchers were subscribing because they needed access to absolutely current issues, and that there was significant educational benefit in issues that were months old. They proposed that back issues (6 or 12 months old) be made freely available to the public to support educational uses, and expected that this would have no significant effect on subscription count. Gradually more and more journals came to this same belief, and today the programme comprises the largest archive of free full-text research articles that we know of: over 825,000 articles from about 220 journals.

the largest archive of free full-text research articles that we know of

Realize that content is king

In the early years of online journals, there was a significant debate over whether services (such as alerting, enhanced types of searching and linking, off-campus access by username and password, etc.) would be important enough to ensure subscriber and society-member retention as more and more access to content was available via institution-wide IP-based subscriptions. In other words, could value-added services be differentiators, or was content king? Users quickly taught us that no single service was valuable enough to ensure individual subscriptions, and that it was access to content that was most important to readers. Publishers responded by adding content such as backfiles, online-only articles and online-only content supplements. Societies included online journal access as part of society membership (and eventually some societies provided print journals to members only at extra cost) and provided access to the full backfile to society members.

Work with librarians as reliable, trusted collaborators

Librarians were part of the team that founded and fostered HighWire, and HighWire itself resides organizationally with the Stanford University Libraries. Especially with society publications like *JBC*, librarians were part of the chain that delivered research to faculty. Librarians helped faculty overcome the early problems with the new technology of the web, and saw the great promise the web held for changing access to the literature. Publishers recognized the essential roles of librarians, and valued them as colleagues. In practical terms, this comfort with librarians as advisors and collaborators in the research-delivery enterprise meant that publishers sought advice from librarians on subscription models, and trusted librarians to abide by subscription licenses (or to clearly object and suggest improvements).

Optimize access and subscriptions

Several of the above business decisions were the most visible expressions of the principle the not-for-profit community shared of maximizing access to the literature, while somehow retaining a subscription base that would adequately fund editorial, technical and business operations. But there were a number of smaller decisions made by publishers that were true to this principle. For

example, publishers who placed supplemental data online but not in print relatively quickly moved to have such data files made open to the public without a subscription, so that readers of the print journal would not be denied access to supporting data. A more obvious example: after a brief debate about potential cannibalization of subscriptions, most HighWire-hosted publishers decided to offer pay-per-view at very low cost to ensure accessibility to non-subscribers. Eventually, the debate about the optimization would shift to incorporate alternative income sources – such as author and funder payments – beyond subscription income, as the open access debate took off.

Treat indexes as sources of readers, not as competitors

This is a business principle that is consistent with the technical insight that readers search more than they browse. There was, for a time, a lively debate among our early publishing partners whether it was better to send metadata to secondary-indexing services of all sorts; whether one should be selective about which secondaries had your publication metadata and abstracts; or whether one's own journal site should be the exclusive location for searching for content in that journal site. Some publishers believed it was necessary to send data to PubMed, but not to others. Eventually, however, publishers almost universally decided that the more indexes your metadata had, the more likelihood your literature would be found. Publishers mostly agreed that leading readers to a journal article was more important than having high amounts of search traffic on their journal site. This principle eventually led us to develop simple methods for secondaries to pick up metadata, and also to collaborate with Google to develop the techniques for indexing the full text of the research literature. It has also led some publishers to supply full text to research groups who are developing 'data mining' tools to be used on the literature; these indexes and tools would refer readers to full-text papers on the journal's online site.

Encourage others to link not host

This is still a widely debated principle: whether it is better to offer others a link to journal content or to provide them with an electronic copy for them to host on their own site. This principle led, for example, to the development of electronic reprints, which provides authors with a 'toll-free' link that they can send to colleagues, rather than a PDF that is emailed; the link takes the reader directly to a paper on a journal site, without any need of a subscription. The benefit to the author and reader is simplicity; the benefit to the publisher is traffic to the journal site and brand recognition. All parties benefit because the online journal site provides many links to and from the research paper. This principle also lead about half the HighWire-hosted publishers who participate in PubMed Central to opt for *PMC*'s 'PubLink' model in which the literature is deposited for indexing and archiving, but display for readers is directed from PMC to the journal site.

debate the results, not the possibilities

'Time-box' development

Publishers were interested in rapid delivery of new function. It was more important to most publishers that a new feature be delivered quickly, rather than that it include all the bells and whistles that anyone at HighWire or the publisher's office could think of. Publishers generally also appreciated that we could sometimes suggest an implementation that provided 80% of the function at 20% of the implementation time or cost. Rapid delivery would allow rapid user feedback, and a second iteration could then be designed and deployed with strong user input.

Devise an experiment; debate the results, not the possibilities

In Silicon Valley in the mid-1990s, there was a saying (popularized, if not invented, by management-by-chaos guru Tom Peters), 'Ready. Fire. Aim.' One of the great benefits of the web that publishers quickly leveraged was the ability to develop new features and services without the huge cost of printing and distribution. This allowed them to rapidly

mount editorial, technical or business experiments, collect data on reader response, and learn from the data; there was real hope that this rapid-cycle experimentation would short-circuit some of the more 'religious' debates going on in publishing. There is a bias towards experimentation among the HighWire-hosted publishers, and a respect for data over rhetoric. Today, the best example of this principle in action is found with over a dozen HighWire-hosted publications that are experimenting with open access and 'open choice' publishing models, and sharing their data with other HighWire-hosted publishers.

Foster collegiality

The key engine of innovation at HighWire is not the 120 staff who work at Stanford University, but the 130 different publishers and societies who collaborate with HighWire and with each other to identify opportunities, meet challenges and share a wealth of information about how their experiments in online publishing are going. HighWire hosts two meetings at year – one in the spring at Stanford and one in the autumn in Washington, DC – at which all HighWire publishers and societies gather to share this knowledge. By far the majority of presentations at the meeting are publishers speaking to publishers rather than HighWire staff speaking to customers. The publishers are able to 'stand on each others' shoulders' if you will. The diversity of publishing and editorial voices provides a wide scope for innovation on many fronts. In the early days of any disruptive technology – the Internet's effect on scholarly communication is nothing if not that – it is important to avoid standardizing too early while innovation has still to run its course. HighWire is fortunate to have smart, creative and collaborative colleagues for our customers. Perhaps the only common quip about the meeting is that there is too much useful information and too many new ideas packed into the time. Attendees are exhausted, but never bored.

In conclusion

Even ten years on, we don't think online publishing should be boring. And we look forward to another decade of innovation with the HighWire publishers.

Reference

1. Simoni, B. Serving science while paying the bills: the history of the *Journal of Biological Chemistry Online*. *Learned Publishing*, 2005:18(2), 127–30 (www.ingentaselect.com/rpsv/catchword/alpsp/09531513/v18n2/s7/p127).

John Sack
Director
HighWire Press
Stanford University

HighWire Press is a division of the Stanford University Library. It won the ALPSP Award for Services to Not-for-Profit Publishing in 2003 for 'providing a magnificent service to not-for-profit and society publishers'.