ATG Interview with Victoria Reich

Margaret Landesman

University of Utah, margaret.landesman@library.utal.edu

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

Part of the Library and Information Science Commons

Recommended Citation
Landesman, Margaret (2002) "ATG Interview with Victoria Reich," Against the Grain: Vol. 14: Iss. 6, Article 16.
DOI: https://doi.org/10.7771/2380-176X.3948

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.
group on its own is enough. We need a sort of “biodiversity” model, with several groups taking different approaches.

**ATG**: On a lighter note, I wondered where HighWire Press is located?

**JS**: HighWire is in the Stanford Industrial Park, about half a mile from the Stanford Library and the center of campus. It used to be that my office was in the library and the rest of the staff was in trailers right next to it, but the university didn’t like having to continually add trailers as we grew. The University had a building available in the area right off campus (across the street from Hewlett-Packard’s main building), and HighWire leased it. It’s a good space and centrally located, with easy access to the campus. But the best thing is that we got the University to agree that we wouldn’t have to have parking permits. As permits cost $200-400 dollars a year, this was a very popular decision among HighWire staff.

**ATG**: The photos on HighWire’s Staff Directory seem to show lots of dogs and babies on the payroll.

**JS**: Well, we have about a dozen dogs in during the day. It’s definitely a dog friendly and child friendly place. It’s good when kids are there — there’s usually only one at a time. Many of us don’t have kids of our own, so having kids around is an invigorating thing. It reminds you of the rest of the world. And dogs are fascinating because they are dogs — running around doing doggy sorts of things. Kids and dogs depress you pretty rapidly.

**ATG**: Perhaps we could end with some ideas about how you see the future of HighWire — what do you see happening next?

**JS**: What we see happening is that in a few years, people will be doing most of their literature discovery work in just a few places. There has been quite a consolidation of general-purpose Internet-search sites and we see this trend continuing on into the science-search domain. People form habits as to where they work. If HighWire isn’t one of those places, our publishers’ content could be ignored when researchers are doing “discover” (vs. “known item”) types of searches.

Researchers will use multiple approaches to literature discovery — and that is a good thing. But they will use a small number of different approaches — choosing just a few from among the most powerful and useful tools and content. Our goal is to make HighWire’s portal at http://highwire.stanford.edu one of the choices on their short list, to be sure our publishers’ content is highly visible.

---

**ATG Interview with Victoria Reich**

**Director (and founder) of the LOCKSS program <http://lockss.stanford.edu>**

**Addressing the Challenges of e-Journals, February 28, 2002**

by Margaret Landesman (Head, Collection Development, Marriott Library, 295 S. 1500 E. University of Utah, Salt Lake City, UT 84112; Phone: 801-581-7741) <mlandesm@library.utah.edu>

As Vicky and I talked, we discussed her various activities in more or less chronological order. Later on, Vicky said of the work we’d been discussing, “but that’s not my true love of the moment.”

There’s not much to say to that, except to ask:

**ATG**: And what is your true love of the moment?

**VR**: That would be LOCKSS, Lots of Copies Keep Stuff Safe.

**ATG**: Why this fondness for LOCKSS?

**VR**: As with HighWire back in 1995, we have a new idea about how to solve a difficult problem, this time e-journal archiving and preservation. This is a critical issue for both scholarly publishers and libraries. So far, the LOCKSS technology is promising. We are in the early days, developing the idea and helping it come to life. I’ve been given an opportunity to work with many international librarians and publishers. Meeting new colleagues and helping to solve an important problem is exciting and fun.

**ATG**: How did the idea of LOCKSS come about?

**VR**: About three years ago, while hiking with a friend, I was voicing frustration with the lack of progress around digital archiving and preservation. HighWire had built an excellent electronic journal online environment. We had made it easier for readers to use and find the literature. However, even three years ago, the authoritative version of most titles was not the paper version, it was the electronic version. The paper edition of many titles provides a subset of the whole journal, which is digital.

This is causing a serious problem. Libraries aren’t giving up paper copies because they worry about archiving — as they should. But keeping the paper is no longer archiving, because paper isn’t the definitive edition. So, librarians are spending money on the print, which is no longer the authoritative version. Even when they purchase the electronic, in addition to or instead of the print, they can’t take custody of the content. In the current electronic journal world, librarians, through little fault of their own, have abridged building local collections. They are unable to take custody of material, to guarantee access to important material for both their current and future communities.

And there is important literature at risk here. For example, take Pediatrics, the journal of the American Association of Pediatrics. Pediatrics has unique content in its paper edition and, almost since day one, substantial unique content in the electronic. But libraries aren’t capturing the electronic part.

**ATG**: Ok, my friend responded, tell me about how libraries work. How do they do it with paper copies?

**VR**: With paper, there are often many copies of any particular journal scattered around the world. Libraries know where some of the copies are, though it’s difficult to find all of the copies. We talked about people who go to libraries with razer blades, and the ways in which some paper copies might get destroyed. But, because the system is “redundant,” though we can’t find all of the copies, we can easily find some of the copies and deliver them to authorized readers.

Could there be something like that for the electronic world; would it be possible to build a system that allows libraries to take ownership of the journals and to have a similar level of security in archiving?

**ATG**: So how did you go about building LOCKSS?

**VR**: With the support of [Stanford University Librarian] Michael Keller, we received funding from Michael LeShk, then of the National Science Foundation, to build a first version of the system. AAAS ScienceOnline gave us permission to use their content for testing. We distributed ScienceOnline content to six partner libraries (Berkeley, Columbia, Harvard, LANL, Stanford and Tennessee). Each library brought up a persistent LOCKSS cache and cached the ScienceOnline test content. We deliberately damaged the content files. Sometimes content files just became damaged due to disks failures or major fires in New Mexico, or human error. As planned, the machines talked to each other and repaired the damage. The machines collected new issues as they were published. Over the course of that year, the scheme appeared viable and the specific tests worked.

We then requested funds to continue testing from the Andrew W. Mellon Foundation. Fortunately Mellon staff agreed that LOCKSS seemed like a promising idea and with their support we’ve been running a beta test. The beta test has over 50 participating libraries, about half of which are in the U.S. We’re using PNAS Online, ScienceOnline, bmj.com, and the jbc Online as test content. Many publishers have been extremely generous and supportive; we continued on page 65
have about 40+ endorsing the beta test. We hope to soon begin to work with Blackwell Science content. It's important to test the system with content published via a variety of online publishing platforms.

LOCKSS reinstates a library ownership model for digital materials, without threatening the publisher's interests. One important step will be to configure these local LOCKSS caches so they provide access for the individual library's clients when the publisher site isn't available. LOCKSS is a low cost way to allow libraries to make local electronic collection development decisions and to have access to content after subscriptions are canceled or terminated. LOCKSS makes sure the content maintains its integrity, but only the local community can see.

LOCKSS is now a joint program of the Stanford Libraries and Sun Microsystems, which has given us a lot of support. David Rosenthal, senior engineer at Sun Labs, has been the chief theorist and design architect from the beginning; it really would never have gone anywhere without his incredible vision and effort. The program is, intentionally, not a part of or controlled by HighWire; we want this solution to be available to ejournal archiving without restriction. On the other hand, my working at HighWire and having long-standing relations with some of the publishers has made it a lot easier for me to understand publishers' concerns in general and to bring them into the design and implementation process.

ATG: What are the requirements for LOCKSS to create archival access?

VR: At the moment, LOCKSS works best for materials that are delivered through the http protocols, in other words, that are Web based. The content should be immutable. LOCKSS doesn't work for the CNN Website, because every time you click on the URL what's displayed to the reader has changed. The content has to have a beginning and an end, or you have to be able to define for the LOCKSS Web crawler an "archive unit."

In addition, the publishers MUST give explicit permission to the libraries that their content may be cached. We recommend language for publishers to add to their online subscription agreements. By and large, this language says, librarians may use LOCKSS to collect, to preserve, and to provide access to their community provided the terms of access from the LOCKSS cache are identical to the original terms of the subscription agreement. Publishers are fairly happy with the idea of adding that kind of permission to subscription licenses. A local LOCKSS cache provides the local community with access to the cached content whenever the publisher's online system is unavailable. Access is seamless; there are no trigger events.

The publishers also need to give explicit permission to the LOCKSS crawler. It's illegal to systematically collect someone's content! The LOCKSS crawlers work very, very slowly; we're currently testing to make sure the network will not stress the publisher's servers. The publishers grant permission once for each "archive unit." In most cases, the "archive unit" is a volume and permission is granted through a html page, a TOC/Volume, that the publisher needs to add.

ATG: How can libraries and publishers join LOCKSS?

VR: We're starting the second phase of testing in Fall 2002. Any library or publisher who would like to participate is welcome to contact me. The LOCKSS software is open source and more details about our work are available at http://lockss.stanford.edu.

ATG: But we really didn't start the interview talking about LOCKSS; we talked about HighWire Press. How and when did it start?

VR: Bob Simon, the godfather of HighWire, was and is a senior editor of the Journal of Biological Chemistry — he's also a grandfather, an amateur wine-maker, a well-known biochemist, and a leader on the Stanford campus. In 1995, Michael Keller gave a presentation on serial prices at a Faculty Senate meeting. Bob caught him after the meeting. They are both the kind of leader who decides to do something and it really goes forward. They decided to put up JBC online and they did.

Michael pulled together a core group — we all had other jobs — to bring up JBC. The pub-
The unveiling of the JBC Online was at the ASBMB convention in San Francisco, May 21, 1996. Working to build the HighWire Press platform was a wonderful adventure. All of us were doing it out of sense of commitment to improve the state of scholarly publication and because it was so much fun! Mike hired John Sack, a longtime Stanford technical development manager, to lead the team, and he has not only held the team together since 1995, he has created a unique forum for scholarly society publishers through the semi-annual HighWire Publishers’ conferences.

The JBC in hardcopy is published weekly. It’s about an inch thick, about the size of a telephone book. There is a lot of data, but the format is consistent and straightforward. It was the perfect first journal to put online.

**ATG: Where did HighWire go after JBC?**

**VR:** It was not planned, but the second journal HighWire put online is also a weekly, *Science Online*. *Science Online* has a very complicated structure and includes letters to the editor, news, research reports, job ads, etc. So within the first two ventures, *HighWire* had pretty much staked out two opposite ends of the publishing structure.

For the first few years, if you plot the “genealogy” of the journals *HighWire* put online, you could track how word had passed from editor to editor to friend to colleague. Business came to us, by word of mouth.

**ATG: How did you come to be part of this?**

**VR:** We had just moved into my dream house and had started a new job working for Sarah Thomas at the National Agriculture Library. My late husband was a computer science professor at the University of Maryland and had been spending summers at Xerox PARC, the premiere research lab of its day, in Palo Alto. After a two summers, Xerox PARC offered him a very attractive position, and he wanted to accept the offer. “OK,” I said. “We both move with jobs, I’ll go.”

Unbeknownst to me, he started reading *American Libraries*, wrote up a résumé and a cover letter in my name and applied for a job as head of Serials and Acquisitions at Stanford. I was shocked and surprised when Stanford University Libraries called me for an interview. I was offered the position; I kept my promise, the family moved. I was head of Serials and Acquisitions for five years. Being at Stanford has been wonderful.

**ATG: What did you do with HighWire?**

**VR:** In the early days, everyone was involved with almost everything. While there was no choice (we were a small team), it was lovely. I couldn’t write software and choose not to learn UNIX; but I did everything else. After the first 18 months, I began to focus on issues like business models, intellectual property, licenses, and eventually archiving and preservation.

HighWire’s development and direction is driven by the needs of scientists. Stanford University Libraries has a two year grant from the Mellon Foundation to study the impact of electronic journals on science and how scientists use e-journals. I’m part of that project team. The study has generated an impressive amount of data, analysis, and reports. To assess the needs of scientific users, the eJUST project team conducted three surveys with approximately 10,000 respondents, a number of ethnographic-style interviews, and data mining of selected journal logs. All results are publicly available: [http://EJUST.stanford.edu](http://EJUST.stanford.edu).

Over the last seven years, there has been an incredible amount of change in the serials world, some of it good, some of it ugly. I’m lucky to have been directly involved in some of the most positive developments in e-journal publishing. With both HighWire and LOCKSS, we set out to make the world a better place, and I think we’re succeeding.