

# Building Open Source

In 2004, the most recent year for which data are available, the National Center for Education Statistics reports academic and public library spending topped \$14 billion, yet makes no mention of spending on research and development of library systems or services.<sup>1</sup>

Using open-source software can reduce the costs of experimentation—the GNU General Public License guarantees our right to experiment with the software—and can make possible the kinds of innovations that patrons have come to expect from electronic services. Indeed, we’ve enjoyed a steady stream of prototypes demonstrating new ways of using or manipulating library data from the growing cohort programming-savvy librarians in our midst.

But going from prototype to product, even an open-source product, is no small feat. California Digital Library’s Roy Tennant agrees: “We don’t have a lot of experience managing open source projects in libraries.”<sup>2</sup> Still, citing the ever-growing portion of library budgets allocated to vendors selling proprietary products, Tennant says he has high hopes for the future of open-source development. He points to Evergreen, an open-source ILS being developed by the Georgia Pines consortium, as an example of the possible future of OSS: “If they can pull this off and people can see that the entire state is running this, then we have a new ball game.”<sup>3</sup>

David Hughes, CEO of VisualArt Systems, however, worries about what happens when grants overtake organic evolution of a project.

A good or not so good idea gets run up a pole, usually by managers or grant writers or heads of foundations, the idea gets tossed over the fence to engineers who promise the moon, stuff gets funded often looking to appease the needs of the grantor or grantee and somewhere in the mix,

the greater good gets lost. And the life of such grant projects takes on less the need to attain functionality, but goals and objectives that end up as timebase reports. I’ve seen more of these than I care to shake a stick at, and most end up as non-sustaining projects.<sup>4</sup>

Open-source software, he says, works well when the people leading the project are passionate about what they do, and passionate about using the software they’re developing.

## In Detail: OSU’s LibraryFind

Federated searching, or metasearch, is a problem unique to libraries.

Many other industries put their catalogs online and track inventory in the way our ILSs do, but metasearch, the tools that allow a patron to search a number of databases simultaneously, is a rare bird. Indeed, University of Windsor systems librarian Art Rhyno notes, “Metasearch is peculiar to libraries.”<sup>5</sup>

So Oregon State University’s release of LibraryFind, “an open source metasearch application developed by librarians for libraries,” is particularly notable.<sup>6</sup> To better understand how open-source development in libraries works, I spoke with OSU’s Jeremy Frumkin, lead developer for LibraryFind (see figure 7), about the project.<sup>7</sup>

### Q: What is LibraryFind?

**Frumkin:** With LibraryFind, we want to redefine the information discovery experience. We want to simplify the user’s experience without losing the advantag-

es we provide to the user through our rich data, and we want to enable other portions of the user's workflow, whatever their particular workflow may be. We also want the tool to provide [better statistics on] the use of [library] collections and resources—we spend a great deal of our budgets on our article databases, and we need to have the ability to better assess the value we are getting in return for our investments.

### LibraryFind

<http://libraryfind.org>

### OSU's LibraryFind Search

<http://search2.library.oregonstate.edu>

We decided to pursue LibraryFind primarily because of our dissatisfaction at the time with our vended federated search product. Our users were not using the vended product, even though it was on our front page, and even our librarians had difficulty using it. At the same time, we understood that we needed to provide a more unified discovery service for our users.

**Q: Metasearch is pretty much unique to libraries. Did you feel as though you were creating something new? Was development perhaps more difficult because of this?**

**Frumkin:** We do feel that we are creating something new. Federated search itself is not new, by any means. The interesting thing is that until just recently, there really had been few if any improvements in federated search technology within the library community.

I don't think development has been made more difficult due to any deficiencies in any other tools; just the opposite, in fact. It's those features and components that we want that we don't see in other products that drive our development areas. If there were an ideal metasearch product out there, we would most likely not be developing LibraryFind.

**Q: When did LibraryFind development start?**

**Frumkin:** We first started on LibraryFind back in mid-2005, and we were live with a version that was written in php in February of 2006.

Terry Reese [had] built a prototype/proof-of-concept which resembled Amazon's A9 interface (at the time), and this prototype helped us move forward to formalizing the project.

We went live with the Ruby on Rails version of LibraryFind here at OSU in November of 2006, and we released the first public version of the code in



**Figure 7:** Search results on OSU's LibraryFind.

January 2007. The next version of the software (0.8) has a target date of April 2007.

**Q: There are number of vendors offering metasearch products. Why build your own?**

**Frumkin:** We decided to develop LibraryFind for a few reasons. First, we wanted a tool we could customize. At the time we started developing LibraryFind, there was only one vended product that allowed for some of the customizations we were looking to. However, even using that tool would have required a fairly extensive development effort.

Second, we had the in-house expertise to pursue this type of effort, both from the open source side and from the federated search technology side.

Third, we wanted to provide the library community with a tool/service/technology that was open to understanding by the people and institutions who deploy it, especially since such a tool is a critical piece of infrastructure.

By pursuing this development, we could understand how the relevancy worked, we could fine-tune not only the user interface (or create multiple ones) but also fine-tune the indexing and relevancy of results, and we could ensure that the tool could hook into other services that we might provide, be those other services from vendors or other services that we develop.

**Q: How does the value of “free” software compare to proprietary products?**

**Frumkin:** Looking at proprietary vs. “free” software is really a false comparison. Software is software. You

need to look at the value you are going to get out of any potential solution, you need to look at the variety of support that's available, and most importantly, you need to look at the available exit strategies that you can pursue because there will be a time when you will want to use a different piece of software. We chose to develop LibraryFind not because it was going to save us money, but because we felt the return on our investment versus the cost of building the software was greater than our return on investment vs. the cost of purchasing the available vended products. This isn't always the case—for instance, we license a commercial wiki product that has worked wonderfully for our staff.

**Q: So why release LibraryFind as open source?**

**Frumkin:** Because we viewed LF not only as a tool that would be useful to use, but a tool that the library community could adopt, we felt that releasing LF as an open source product would allow other libraries to adopt it more easily, and it would open up the opportunity to build a community around LF (DSpace is an example of this approach). Hopefully, over time, we will build not only a community of users, but also a community of code contributors as well.

**Q: How was this possible?**

**Frumkin:** As for how we pushed this forward at OSU, it may be sort of unique. My position is funded through an endowment, and my role in the library is to push forward innovative services and technologies—basically, I get to free-float and pursue new, interesting, and useful technologies for the library. I'm also in library administration, so I have some sway in shaping the library's strategic direction. I report directly to our University Librarian. Finally, we just did a mini-reorg, and right now the head of library technology (currently a vacant position we are filling) will be reporting to me.

The other thing, in this particular situation, is that they had already deployed our vended federated search product here at OSU prior to my arrival. I believe if our experience with that product had not been as bad as it was, we may not have pursued LibraryFind, at least not as a metasearch product. But since the experience was bad, I think that actually helped—as I stated, when we proposed building LibraryFind, there were no other good alternatives, except maybe Ex Libris' X-Server, and working with X-Server wasn't much removed from building our own product. So, I had to sell the idea to the rest of library administration, which here consists of our department heads, AULs, and the UL. Being as my

position is to actually pursue this sort of cutting-edge activity, it wasn't too difficult to get buy-in, especially after we were able to obtain the LSTA grant from our State Library to help fund the work.

**Q: What was the eureka moment?**

**Frumkin:** I can't say there was a eureka moment. We had no doubt of the feasibility of building LibraryFind—and when we were able to go live with the initial version here at OSU before our grant funding even kicked in, we knew we could make LibraryFind into a new type of discovery tool.

**Q: OSU has some experience with OSS development. Did that help?**

**Frumkin:** We actually have an open source lab here at the university, which does contract development work on open source projects and also recovers costs through hosting services. We didn't work with them on LF, since they mainly work with projects outside of the university, but because of their existence, there is a good understanding of open source at OSU, and our Office of Technology Transfer doesn't really have any concerns about our releasing software under an open source license. They would be more involved if we were to commercialize the product.

**Q: A project like LibraryFind is a big undertaking. How many people are involved in making it happen?**

**Frumkin:** Our extremely talented development team consists of Terry Reese, Tami Herlocker, and Dan Chudnov. The success of the project is truly a reflection of these three people, and they should receive the bulk of the credit for making LibraryFind a reality. In addition to the development team, we had a graduate student, Seikyung Jung, who led a formal usability test of the LibraryFind user interface. She worked with a number of our public services staff to plan and implement the usability testing. We also involved the broader library staff in the testing of the software, as well as for providing informal feedback to what worked well and what didn't. Finally, a huge chunk of credit needs to be given to our university librarian, Karyle Butcher, whose leadership at the OSU Libraries provides us the ability to move forward and to be forward-thinking on projects such as LibraryFind. The support of Karyle and the rest of the library's administration provides a solid foundation which enables LibraryFind to be successful.

**Q: OSU is big, with an obviously strong library system. Is LibraryFind an option for other libraries?**

**Frumkin:** Definitely. As part of our grant funding, we have partnered with a public library to test their ability to use LF. We are also looking to work with a number of libraries, both academic and public, so that we can better understand the challenges that they will face in adopting LF. The beauty of this, though, is that because LF is open source, we can meet these challenges and continuously improve LF so that it is adoptable by a wide range and various-sized libraries and institutions.

**Q: What benefits does LibraryFind offer to those libraries?**

**Frumkin:** I see the advantages being LF provides an easy-to-adopt, high quality metasearch tool that could be used 'out-of-the-box', or that could be customized at whatever level the library might want.

Instead of putting the cost of ownership into the price of the software, it can be transferred into the actual deployment at the particular library (i.e. Instead of paying \$25,000 for the software itself, that money could be used to support higher-level activities). It also provides a fully-functional OpenURL resolver.

**Q: Do you expect them to contribute code? What's the advantage to the project?**

**Frumkin:** My expectation is that most adopters of LF will not be contributors, though we would be ecstatic if we did become successful to the point of having a good core of contributing participants.

The advantage to the LF project is that each new user/adopter is an additional community member who can contribute feedback through the requesting of new features or reporting bugs, and who can potentially be a source of support for other community members. There are probably many other benefits on both sides, but those are the first that come to mind.

**Q: What about the unthinkable . . . what if you leave OSU or the project? What if OSU's interest and support in the project wanes? Then what?**

**Frumkin:** Well, you are really getting to sustainability, which is always an issue. Here's the thing:

With LibraryFind, there is no lock-in; Once you have the code, you have the code. If the project ended tomorrow, the code would still work. If OSU stopped working on LF, someone else could decide to fly with it. And because the code wouldn't disappear

if we stopped developing it, there would be plenty of time for an institution to look at its options.

If we are lucky, we'll develop a large enough community around LF so that the community will be a major sustaining force. Imagine if tomorrow MIT said "Hmm, that institutional repository idea—it wasn't as cool as we thought it would be. We're done with DSpace." If that happened (which is highly unlikely, but a good example), I seriously doubt many places would be overly worried about their particular DSpace implementation, at least for the short term. As I talked about earlier, one of the nice things about LibraryFind, and open source software in general, is the exit strategy.

**Q: What about the future of open source in libraries?**

**Frumkin:** Those of us who have been involved in open source in libraries for awhile now (10+ years, yikes!) are starting to move up the ranks, and hopefully we'll start to see more folks move into positions of influence within libraries. Whether we're talking about open source, open access, open API's, or open platforms, we still need to be able to make strategic decisions about our technology choices, and sometimes those strategic decisions will lead us to purchasing vended software and services, and sometimes it will make sense to collaborate and create open source software.

If we can get past the FUD [fear, uncertainty, and doubt] on both sides, which isn't easy to do, then we can move away from the religious wars and move towards what works best for libraries, both now and in the future.

**Notes**

1. B. Holton, K. Vaden, and P. O'Shea, "Academic Libraries: 2004," National Center for Education Statistics, Nov. 2006, available online at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007301> (accessed Mar. 19, 2007); Adrienne Chute et al., "E.D. TAB: Public Libraries in the United States: Fiscal Year 2004," National Center for Education Statistics, Aug. 2006, available online at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006349> (accessed Mar. 19, 2007).
2. Roy Tennant, interview by the author, Aug. 3, 2006.
3. Ibid.
4. David Hughes, e-mail to the author, Aug. 16, 2006.
5. Art Rhyno, e-mail interview by the author, Mar. 7, 2007.
6. "LibraryFind Is . . .," on the LibraryFind Web site, <http://libraryfind.org/about.html> (accessed Mar. 19, 2007).
7. Jeremy Frumkin, interview by the author via e-mail, Mar. 6-9, 2007.