Cathedral, Bazaar, or Some Third Thing: Can Open Source Work for Libraries?

James Elmborg, School of Library & Information Science, University of Iowa

I recently attended the West Virginia Library Association conference. It was a fine event staged in a beautiful mountaintop setting. State and regional conferences always appeal to me more than big national ones. There are fewer attendees, which means you have better extended conversations, and you can navigate the program more easily. At this conference, I did a presentation about reference as spatial practice, which is something I have been thinking about for a while. Clearly, the library is in a transformative stage. What we think of as "library space" is changing rapidly. Libraries have been experimenting with space, and academic librarians have been trying to preserve the concept of "the library" while experimenting with new ideas about space. This is an important challenge with real consequences for the practice of librarianship.

The banquet speaker at the conference was Tim Spalding, founder of LibraryThing. LibraryThing is a tool for cataloging personal libraries; it works by allowing users to create tags to describe the books they own. In effect, LibraryThing allows people to have individualized library catalogs for their home collections and to share those catalogs with others. Spalding was an energetic speaker, moving from screen shots of LibraryThing in action to slides that compared LibraryThing's functions with how traditional library catalogs work. In addition to entertaining the conference attendees, he also obviously wanted to provoke them to think about the implications of Web 2.0 technologies for how libraries do business. He used several examples of books

that were tagged in LibraryThing and compared them to how books were classified by Library of Congress Subject Headings. The most striking example was the LibraryThing tag "chicklit." The book most often tagged as "chicklit" in LibraryThing is *Bridget Jones's Diary* by Helen Fielding. By searching for "chicklit" in LibraryThing, you can see that 5,373 users have used this tag 63,007 times. *Bridget Jones's Diary* has been tagged as "chicklit" by 875 users. [1]

Further down the list of books tagged as "chicklit" we find *Pride and Prejudice* by Jane Austen, tagged as such by 58 users. Spalding noted the difference between the tag, "chicklit," and the LC Subject Heading for Austen's book. Spalding chided librarians for their adherence to relatively useless controlled language to describe books like *Pride and Prejudice*, claiming that "chicklit" is a much more meaningful descriptor than anything we might find in LCSH. Indeed, when I replicate Spalding's search in our university catalog, I find *Pride and Prejudice* classified under such vague and general descriptions as: Social classes—England—Fiction, Courtship—Fiction, Domestic Fiction, and Love Stories. We might debate the differences between LibraryThing and a university catalog, especially about whether the tag "chicklit" is appropriate for an academic library subject or whether it will be viable even a few years from now. However, in doing so I think we would be missing the larger point. By allowing users to create tags for their own books, LibraryThing has become a dynamic and growing book catalog, one that users enjoy using, that they find effective at helping them find books to read, and, perhaps most importantly, that connects them to other readers of similar books. Spalding ended with some pointed advice for librarians. First, the good news (according to Spalding) is that readers still enjoy reading books, and they enjoy participating in a

community that helps them connect to other people who read books. The bad news (for librarians) is that these readers are not deeply invested in libraries or the future of libraries.

In speaking to this audience of librarians, Spalding could have given a warm and fuzzy speech about libraries and how he grew up loving them. We've all heard non-librarians give this feel-good talk at keynotes and banquets. I thought he did a more useful thing by addressing issues that were tough for librarians to hear discussed. He clearly knows a great deal about libraries and how they work, and he ended by taking on some sacred library cows. He showed us a link to a library book in the library catalog. He asked why library books never show up in Google search results. The answer, he noted, is that our library catalogs do not generate static URL's, so they can't be crawled or listed as hits by Google. He rightly noted that this is the first rule of Web 1.0, one that library catalogs still haven't mastered—you have to show up in Google. Finally, he observed that we (in libraries) are being led down a bad road by the vendors who sell us Integrated Library Systems, and by OCLC who continues to manage controlled cataloging in a centralized process that generates vague and meaningless categories like "Domestic Fiction." He encouraged libraries and librarians to make the leap to open source software that will allow them to take back control and direction and to have input into how the catalogs work.

As a longtime user of open source software, I join Spalding in wondering why libraries continue to depend on vendors who sell them unsatisfying, expensive software, and why LCSH remains the gold standard for cataloging. In my conversations with librarians, I rarely hear anyone talk about how much they love their catalog. More

common is the helpless shrug that says, "What choice do we have?" I serve on a

statewide leadership team for school libraries in Iowa. A few months ago, I presented

information to the group about open source options for library catalogs. I noted that a

few state library systems had adopted open source catalogs, as had some school systems

and small colleges. I suggested that it was worth considering open-source catalogs as

viable options for schools, especially as free alternatives to products with ongoing vendor

fees. The anxiety in the room was palpable. These librarians, selected as leaders in the

school library community, could not accept that open source solutions could work in

enterprise roles. In this conversation, I flashed back to an essay I assign in the

Foundations course I teach, "The Cathedral and the Bazaar" by Eric Raymond. This

essay had prompted many of the students in my summer class to examine the viability of

open source catalogs. These students concluded that open source catalogs (especially

Koha and Evergreen) were legitimate candidates for use in libraries. They could save

small libraries money in difficult economic times, especially once a community of local

users has developed to run the software without fee-based assistance. Most importantly,

open-source catalogs would allow for much more rapid innovation in features than

commercial vendors provided. However, as one student concluded, the culture of

librarianship is not ready to accept the philosophical change that open source would

entail. [2]

 Raymond's "The Cathedral and the Bazaar" helps us explain why this

philosophical shift is so difficult to make. While initially discussing the difference

between open source software and closed-source software, Raymond moves almost

immediately to the claim that open-source is not a software model so much as a

management model. In this analysis, traditional software (i.e. commercial software) is managed on the "cathedral" model. In cathedral development, a few specially selected people develop software through a process that is managed toward the end goal. The software's vision is designed at the upper levels of management and then written at the lower levels to perform according to the specifications of the vision. Most importantly, the code is "closed," which means that only a small group of programmers, hired and trained in and by the cathedral, produce the software and understand how it works. In Raymond's words, cathedral software is "carefully crafted by individual wizards or small bands of mages working in splendid isolation." [3] The internal workings of the software, the "code" itself, are not available outside the cathedral. When something does not work as it should, the solution to the problem must come from the cathedral.

The open-source management model is premised entirely differently. Raymond describes it as "a great babbling bazaar of differing agendas and approaches… out of which a coherent and stable system could seemingly emerge only by a succession of miracles." [4] Indeed, open-source software had much to prove to consumers of commercial software when it began to emerge into the public sphere in the early 2000's. It still seems risky to base crucial enterprise services (like library catalogs) on software that has no technical support team and no warranties behind it. But the amazing thing about the open-source/bazaar model is that *it works*, and in some cases, it works really well. It works because, unlike commercial software, which has long development cycles with supposedly perfect software releases, open-source software is under constant development. This is important because, according to Raymond, "given enough eyeballs, all bugs are shallow." [5] In other words, opening up the code to anyone who wants to

improve it encourages a participatory model that lends itself to constant development. If everyone is a bug fixer, then it is much easier to find and fix bugs if that process is ongoing. And the crucial factor that makes this model work is that in the successful open source projects, the volunteers who fix these bugs are *passionate* about fixing bugs, and about their participation in the community. They take on difficult tasks with energy and creativity, often for no reward except what Raymond calls "egoboo." "Egoboo" is short for ego-boosting, the social capital one accrues in showing off one's programming skill to an appreciative audience. [6]

Raymond describes five basic functions that software managers perform in the commercial/cathedral model. They are:

- To define goals and keep everybody pointed in the same direction
- To monitor and make sure crucial details don't get skipped
- To motivate people to do boring but necessary drudgework
- To organize the deployment of people for best productivity
- To marshal resources needed to sustain the project [7]

The assumptions behind this management approach help us understand the cathedral model. The vision comes from above. The manager keeps the team moving to execute the vision. The programmers tend to skip the details and get bored by their jobs, so the manager must monitor and motivate them. The manager must retain and organize resources (both human and non-human), including fending off competition for resources with other units within the company. This is the world of Dilbert, the corporate software company.

In contrast to the closed, commercial model, managers in open-source projects coordinate the activities of unpaid volunteers. They define challenges and recruit programmers to show off their skills in service of the greater good. This approach turns the task of the manager on its head. If the project manager of a cathedral project is responsible for defining a general goal, the job of the open-source manager is to "sell" the project to the potential pool of programmers and users who might value it. Once that pool of users is "sold," then all the managerial tasks take care of themselves. The community of users has accepted and then helps develop the direction of the project. The users and programmers monitor themselves to identify crucial details. The value of the project motivates everyone to do the work—even the drudgework. The community organizes itself, with users and programmers gravitating toward the places they are useful. These resources will stay in place as long as the project is valued by its community of users and programmers. This sounds utopian, and of course not every project ends up being successful. When it works, though, the bazaar model has created some of the world's best software (including Firefox and its plugins, the Apache web server, most of the content management systems like Wordpress and Drupal, and ultimately Linux itself). Given the success of such a variety of open-source projects, it is increasingly important to at least consider open source when making software decisions, if only to provide healthy competition and incentive to the corporate software giants who have been serving us so poorly. The open source model is transforming software and communities as it moves out of technical circles and becomes recognized as a new form of social organization with remarkable (and not coincidental) similarities to Web 2.0.

Raymond's identification of these two types of organizations, the cathedral and the bazaar, has useful implications for libraries. Library catalogs are developed and managed on the cathedral model, and indeed, libraries themselves have traditionally been managed as cathedrals. To re-purpose Raymond's language, catalogs (and libraries) are "carefully crafted by individual wizards or small bands of mages working in splendid isolation." Once the "product" is released, we expect it to be more or less "finished" so that it can be used it as librarians have designed it to be used. If there are "bugs" in the catalog (or in library services), they tend to be tenacious and long-lived, and often we are asked to just live with them. In the worst cases, we must actually develop instructional strategies to teach people how to overcome these bugs. Whenever library software is not intuitive, we have to teach students and faculty how to navigate the design flaws as part of the natural landscape, at least until the next release when the bug may or may not be fixed. When I worked in library instruction, we had a saying, "any software problem is an educational problem." This situation is perverse, of course, when your goal is to teach people higher order thinking skills and to foster a passion for genuine inquiry and research.

There have been efforts to allow users to tag library catalog items. The University of Michigan boasts a tagging system. My initial searches for tag clouds were disappointing. Not many items I could find were tagged at all, though admittedly, my efforts were neither comprehensive nor thorough. The University of Pennsylvania has tagging software, and several tag clouds can be browsed. They classify a tag as popular if it has been used more than 109 times, a far cry from the 61,968 times the tag "chicklit" has been used in LibraryThing. The power of tags lies in numbers. According to

Spalding, "to do anything useful with tags, you need numbers. With only a few tags, you can't conclude much. The tags could just be 'noise'." How many tags are required for usefulness? Again, according to Spalding "Critical mass is important, even if we can't pinpoint the line. Ten tags are never enough; a thousand almost always is." [8] As we examine the differences between how LibraryThing operates (as a bazaar) and how academic library catalogs operate (as cathedrals), I think we have reason to be concerned about why our catalogs lack more robust implementation of Web 2.0, and about why we continue to rely on corporate software to deliver it. Discussions in the library world have been bubbling along for years now about the impact of Web 2.0, but relatively little has changed about our catalogs.

Some serious questions loom, and there is urgency to answer them. How can libraries create the kind of loyalty and energy that users of LibraryThing have? How can library catalogs evolve into social networking sites? Should they? How can libraries be sites of community activity and purpose? Perhaps the most important question of all is: why is there such hesitancy on the part of libraries in general to entertain the viability of the bazaar development model? There is such inertia in library culture, so much invested in the cathedral and the library's place as a complete and produced space, a place defined and served up by librarians for the community's use with very little input from the community. The idea of "tag clouds" already seems quaint, a relic of 2005 thinking. We are so far behind, and it seems we get farther behind as time passes. In a very short while Google will have begun serving up the entire collections of the nations' research libraries, at which time the only viable use of library catalogs will be to locate holdings already identified through Google, over-produced shelf lists. As Spalding notes, the catalogs that

vendors produce for our use (and the way our institutions are managed) are still struggling to enact Web 1.0 technologies. Google Books will further widen that gap.

As I teach classes for librarians-in-training, we often discuss the college library of the future, the library these students will work in. I have never felt libraries would go away, but I do fear they might evolve into empty and meaningless "centers." The library is too symbolically important for any campus to abandon. However, it is very possible for that symbolism to become empty, for meaning to be evacuated. Indeed, on many campuses, the library itself is evolving into a generic social space. Ray Oldenburg, in his book *The Great Good Place*, argues that people need a "third place," one that is not home and not work. According to Oldenburg, the "third place" fulfills a unique social need, a need for camaraderie, conversation, and relaxation. He provides five essential things a "third place" must have. They "exist on neutral ground;" they flatten social hierarchy and create social equality; "conversation is the primary activity;" there is a crowd of "regulars;" the place itself has a "low profile;" and the mood is "playful." Above all, those who enjoy the "third place" see it as theirs, a sort of "home away from home." [9] It is tempting to see the future of libraries as "third places" on the Oldenburg model, and perhaps to see this as a way of creating the connection to community that virtual users have with LibraryThing. What concerns me about this vision is that, at some point, the library stops being a library and becomes what Henri Lefebvre calls "abstract space." [10] Meaningful uses for space are made generic, and the space becomes a kind of commodity, intentionally produced for consumers, who occupy it and make use of it without any real deeply held connection to it. In Oldenburg's model, the users of third

places "appropriate" them without regard for their original intent. Is that the fate of libraries?

So what makes a library a library? That seems to me a crucial question, and one whose answer is not clear at this historical moment. Perhaps the most important thing about an academic library is that it be academic. It can't be about the coffee shop and the computing facility or the social conversations that take place there. While these amenities seem comfy and homelike, I think it is strategically unwise to convert libraries entirely into social centers. LibraryThing is popular because it engages its users on an intellectual level. People who are passionate about their books invest great energy in tagging them. They do so for complex reasons: they love their books, they value belonging to a community of book lovers, they like building something as part of a community, and undoubtedly (and perhaps most importantly) they enjoy the intellectual activity of tagging and the "egoboo" of showing off their collections. Academic libraries, in general, seem hesitant to open up the cathedral to such activity, which I think is a mistake. This week, I was asked to do a workshop for faculty on our campus on Zotero, an open source bibliographic tool. Zotero does what many other tools do (Endnote, RefWorks, etc.). Indeed, our library subscribes to RefWorks and does workshops on it. When I asked why these faculty members wanted to use Zotero rather than RefWorks, they responded that they wanted the opportunity to build and share collections with other faculty around the country as a way of creating broader intellectual community.

Libraries claim to seek relevance and connection to their communities, yet overall, libraries have been extremely slow to embrace technologies like Zotero and

LibraryThing that are modeling the best ways to find and nurture those relationships.

Clearly I am making sweeping, general statements. Throughout the professional

landscape, we can see librarians working with Web 2.0 technologies and finding creative

ways to use them to connect with users. Librarians tweet and blog and tag. They have

avatars that build libraries in virtual worlds. LibraryThing has developed a modular

plugin that allows tags from its collections to display in results of searches in a library's

catalog. Over a thousand libraries (most of them public) have incorporated this plugin to

their catalogs. Still, there is an overall resistance to these powerful new directions—a

lack of trust in what will happen if we let go of our cherished positions as "individual

wizards or small bands of mages working in splendid isolation." And often that

resistance comes from key powerful players in the library, many of them the "wizards

and mages" with the most to lose in any transformation to a bazaar culture.

In Europe, most cathedrals long ago stopped being vibrant centers of culture and

became instead tourist destinations. The deep meaning attached to them by their

communities long ago has given way to a kind of majestic abstraction that feeds the

tourist's imagination about bygone times. At best, visitors connect to these cathedrals as

romantic attachments to an idealized past. Seen this way, the cathedral is both a

management model and a cautionary tale. The center of activity is now the bazaar, the

teaming activity of the online intellectual marketplace. As LibraryThing and many other

open-community projects demonstrate, the bazaar can be both a lively *and* intellectual

place. That seems like a model academic libraries would do well to embrace going

forward.

[1] chick lit | books tagged chick lit | LibraryThing. http://www.librarything.com/tag/chick+lit

(accessed December 15, 2009).

[2] Erickson, J. *Open Source Library Systems: Positioning Your Library for the Future of Change*. Unpublished Course Paper. Aug. 1, 2009.

[3] Raymond, p. 21

[4] Raymond, 21-22

[5] Raymond, p. 30

[6] Raymond, p 53

[7] Raymond, p. 58

[8] Spalding, T. "When Tags Work and When They Don't: Amazon and LibraryThing." Thingology (LibraryThing's ideas blog). http://www.librarything.com/thingology/2007/02/when-tags-works-and-when-they-dont.php (accessed December 15, 2009).

[9] Oldenburg, R., & Ph.D, R. O. *The Great Good Place: Cafes, Coffee Shops, Bookstores, Bars, Hair Salons, and Other Hangouts at the Heart of a Community* (New York: Da Capo Press, 1999), 22-42.

[10] Lefebvre, H. (1991). *The Production of Space* (Oxford: Basil Blackwell, 1991), 49-53.