Open Source Software in Libraries: Opportunities and Expenses

Open source software (OSS) is not a panacea; it will not cure all problems computer. On the other hand, it does provide the library profession with enumerable opportunities as long as we are willing to pay a few expenses. This essay elaborates on these ideas by: 1) outlining what open source software is, 2) describing how its principles are similar to the principles of librarianship, and 3) enumerating a number of open source software applications. By the end it is hoped you will be have a better understanding of what open source can and cannot do for libraries. You will be better able to discuss topics related to open source software with "techies". Finally, and probably most importantly, you will have learned the definition of "free" in the context of open source.

What is OSS?

Open source software is both a software development process as well a software distribution process. On one hand it is a way of creating and maintaining computer software that is only really feasible in a globally networked environment such as the Internet. The process usually starts out with a person who has some sort of computing problem and then goes about writing computer program to solve the problem. This is often referred to as "scratching an itch." The archetypical example is the development of Linux. A person -- Linus Trivols -- wanted to run a version of Unix on this Intel-based computer. In order to do this he needed to write the core of an operating system, and so he did. The next step in the development process is to share the often fledgling solution with sets of one's peers. Allow them to read the computer code, make changes or enhancements, and incorporate them back into the original software. The process is then reiterated. Rinse. Shampoo. Repeat.[1] As this

---

[1] Did you hear the one about the programmer who couldn't get out of the shower? He got stuck in an infinite loop while washing his hair. Rinse. Shampoo. Repeat.

process builds momentum a community is hopefully formed. Norms are established, and if the code "scratches many itches" it takes on a life of its own.

Open source software is also about software distribution, and the essential characteristic of this distribution is "free". In this case, the word free has two meanings. On one hand, free can be equated with gratis -- without a charge. No money is paid for the software. Money might be exchanged to put the software on a medium. Tape. CD. Installation. Money might be exchanged in order to provide technical support, but in all cases, the software is gratis. Second, and more importantly, free is equated with liberty, meaning, with few restrictions. In this sense of the word, the person obtaining the software is allowed to modify it for their own use. You are "free" to make changes.

OSS and librarianship

There are many similarities between open source software and librarianship. Both value the peer-review process. Open source software is "peer reviewed" given the shear numbers of people who look at code, point out its defects, and provide solutions. As they say in the open source world, "Given enough eyeballs, all bugs are shallow." Librarianship, especially academic librarianship, values peer review since it is a similar process to winnowing the wheat from the chaff.

Both open source software and librarianship put few restrictions on how their content is used. Open source software licensing models are very explicit about this. You can use the software in any way you see fit, as long as you understand that it comes with no warranty of any kind. Librarianship is similar. We do not ask people what they are going to use the information for. That is considered none for our business.

Both open source software and librarianship value the concept of "free". When was the last time you paid for open source software? When was the last time you paid to use a library? In both cases, it is not about money as much as it is about service.

Collaboration is another key value shared between the open source community and

librarianship. Both institutions pay high regard to group work over the work of the individual.

Finally, both communities deal primarily in data and information. Let's face it, software does nothing but take data in, perform some actions against it, and output the data in another form. Software is about data and information. Libraries are very similar except the majority of their data and information is usually manifested in books, journals, microforms, etc. Libraries provide tools and an environment for processing the data and information for the purposes of creating knowledge and wisdom.

A Few OSS projects

Let alone the open source software projects that literally run the Internet, there are a whole host of library-related projects. Probably the most well-known are Koha and Evergreen -- both integrated library systems (ILS). Then there is Zebra, an indexing application is used in a number of commercial ILS's. VUFind is a "next generation" library catalog application -- a discovery system as they are increasingly called. Similar applications include BlackLight, SOPAC, and Scriblio. DSpace, Fedora, and ePrints are three very popular institutional repository applications, all three are open source software.

Most open source software packages are generic enough to used in environments besides libraries. For example, libraries create and maintain many lists. Books. Journals. Internet resources. In a digital environment, lists are best created and maintained using relational database applications, and MySQL is probably the most popular open source software database application available. Where databases are great at data organization, indexers are great at search. Probably the most popular open source indexer is called Lucene, and it is the heart of many applications including a number of open source and commercial discovery systems. Increasingly the world of data and information is manifested as XML. Two open source software "libraries" known as LibXML2 and Xerces are the core of quite a number of XML processors. Then there is Open Office -- an almost complete open source software replacement for Microsoft Office.

Open source software is everywhere, and whether you know it or not, it drives much of our everyday work lives, especially when it comes to the Internet.

Opportunities and expenses

It is difficult to over-estimate, in my opinion, the number of opportunities open source software provides for librarianship. From my perspective, librarianship is not necessarily about books but rather what the books manifest. Librarianship is about the intertwined processes of collection, preservation, organization, and dissemination of data, information, and knowledge for the benefit of a library's specific audience. In the past books, journals, and a myriad of other physical items have been the object of these processes and content. Such an environment will continue to be important but also pale in comparison to an environment where content is born digital and distributed over the Internet.

The same processes librarianship applies to physical content can be applied to digital content. To do this most effectively, the profession needs to master the use of computer technology. No, not every librarian needs to be programmer, but the skills of the profession need to augmented. Open source software provide a means to have more control over one's computing environment. Open source software is more oftentimes standards compliant than "closed" source software. The skills used in deploying open source software are easily transferable to other computer problems. In this way, the application of open source software is a redirection of resources or an investment in personnel. These are expenses. Instead of spending money for computer systems that you have little control over, you can spend money on people who are in your local environment and have them work on any number of computer-related problems.

Open source software is a viable choice for librarians who want to provide library and information services in a digital environment.

Eric Lease Morgan
University Libraries of Notre Dame

November 22, 2008