

“Open Source” Software

By Richard W. Boss

The term “open source” refers to software that is free and that includes the original source code used to create it so that users can modify it. It also includes the right of redistribution; therefore, there may be products that are based on other open source products.

While open source software may be free, a developer or distributor may charge for services, including custom programming, installation, training, technical support, and hosting services for those who do not want to implement and maintain the software in-house. Some distributors of open source software are for-profit organizations and may charge prices for their services comparable to those charged by proprietary software vendors.

Not all “free” software is “open source.” For example, Google offers a number of applications software products without charge, but it is in complete control of development. Source code is not included and there is no right of distribution. Google covers the cost of development and support by selling on-line advertising.

The perceived **advantages** of open source software are:

- *Ability to tailor to fit local needs*

The availability of the source code means that a user can modify and enhance the software to more closely fit its own needs. Unlike with proprietary products, the development priorities are set by the user, not a vendor. The user is also able to set its own priorities for bug fixes.

- *No restriction on use*

Unlike proprietary software, there are no contractual restrictions on how the software is used. While some developers use the GNU General Public License that assures users that they have the right to distribution and those to whom they distribute also have the right to modify and distribute, other developers merely declare that their software is in the public domain. A subsequent user may, therefore, decide to protect the enhancements that it makes by copyrighting them. The GNU General Public License is, therefore, preferable.

- *Low cost*

There is no charge for the software itself, therefore, the capital outlay required by proprietary software is avoided. The major costs are ongoing development and maintenance. If the number of users is large, and they share their efforts, each user's cost is reduced. However, if the number is small or a user does a lot of tailoring to fit unique local needs that are not shared by other users, the cost can escalate.

The potential **disadvantages** of open source software are:

- *Unanticipated cost*

Over a period of five years, the cost of a software package, whether open source or proprietary, is typically ten to twenty percent of the total cost of a project. With few exceptions, open source software tends to be less complete than proprietary software; therefore, a library may find that it needs to do a great deal more work than anticipated to adapt the software to local needs. If local development is undertaken, new open source releases may not be compatible with what has been done locally.

Considerably more staff expertise and time may be required to use it. It is, therefore, essential to compare costs for multiple options, rather than assuming that the open source solution will be less costly.

- *Lack of coordination*

The decentralized development of open source software means that progress can be chaotic and there may be delays in addressing bugs and in completing planned enhancements. This may increase the burden on a library that decides to proceed on its own.

Not only may there be lack of coordination within an open source development project, there may be little attention paid to integration with other applications.

- *Inadequate training and technical support*

No training comes with most open source products unless a commercial vendor is retained. Documentation tends to be limited and aimed at developers. There usually is limited technical support, especially for users of the software. However, a few open source products do have training and technical support available commercially for a fee.

- *Lack of participation*

Too few participants can cause the development effort to become too expensive for one or a handful of committed parties.

- *Lack of guarantees and remedies*

Unlike turnkey systems using proprietary software, there are no guarantees of quality or performance for open source software. Vendors that provide support services for open source products do offer some guarantees, but no remedies similar to those offered by vendors of proprietary products. Only purchasers of proprietary products can expect financial and other contractual remedies for poor response times and loss of functionality.

- *Scalability and speed*

Open source software may not offer the scalability and speed of proprietary software because the easy-to-use and general-purpose programming languages used are not very scalable and are slower than other languages.

There are tens-of-thousands of open source software products, many hundreds of which are in use by libraries. This TechNote focuses on only a small number of products, those that appear to be the most popular. Not included are open source integrated library systems because a TechNote on open source ILS has previously been posted.

The following criteria have been used to select the software described in this TechNote:

1. There is current development activity
2. Current source code and documentation are available for downloading under the GNU General Public License or a comparable license.
3. Scalability is not an issue—there is no risk of database size or activity levels exceeding the capacity of the software
4. The software has been successfully used in libraries.

Safe Bets

Some open source software has been so widely adopted and has so much development support that there is virtually no risk in choosing it. The safe bets include the Linux operating system, the MySQL database management system, the Firefox browser, and Open Office.

Linux, an operating system, is the most widely used open source software in the world. Approximately 40 percent of all of the world's servers are Linux-based. Not only Linux functionally rich, but it is very stable and secure. It also is supported by a number of commercial firms that offer technical support and training materials. These include Borland, IBM, Macromedia, Novell, and Red Hat. Linux is in widespread use in business and industry not only on servers, but also on clients and desktop PCs. One of its most effective implementations in libraries has been "thin clients, the mounting of Linux on low-end PCs and drawing the operating system and applications software from a Linux server via the network. Linux can be downloaded by qualified developers from www.osdl.org/ but most potential users will need to download it from any one of a score of distributors. A directory is available at www.linux/download_linux/

An attractive option for libraries that have old desktop computers is *Ubuntu*, a stripped down version of Linux that is ideal for mounting on machines that have as little as a 500 MHz processor and 64 MB of RAM. It includes a web browser, presentation, spreadsheet, and instant messaging applications. Ubuntu is not a popular choice for servers as only Sun Microsystems supports it among server manufacturers. The URL is www.ubuntu.com/

MySQL, a multi-user DBMS, was developed by MySQL AB, a Swedish for-profit company, in the mid-1990s. The company is now a subsidiary of Sun Microsystems. There are more than 11 million installations of the DBMS on a wide variety of operating systems and hardware platforms. Unlike most open source software, MySQL is written in C and C++, thus it is very stable, scalable, and efficient. The basic product is available

for free download under the GNU General Public License, but a more comprehensive, supported enterprise version is available for sale from Sun Microsystems. The URL is www.mysql.com/

Firefox, a browser developed by Mozilla, has millions of users worldwide. It not only offers a more secure browsing experience than Microsoft's Internet Explorer, it has hundreds of plug-ins that can be used to tailor the browser to local needs. It is available for download from www.mozilla.org/download/

OpenOffice.org is an office suite for word processing, spreadsheets, presentations, graphics, and databases. It can read and write files from other office software, including that produced by Microsoft. It has an open development process that allows anyone to report bugs, request new features, or actually enhance the software. It can be downloaded from www.openoffice.org/

Other Products by Application

Booking System

WUBS (Web user Booking System) was developed at Britain's Loughborough University Library as a web-based product to allow online booking of resources such as rooms and PCs directly by end users. The URL is <http://sourceforge.net/projects/wubs/>

Digital Library and Repository Software

DSpace captures data in any format—text, audio, video, and data—and distributes it over the web. It indexes the data so that users can search and retrieve the items. To submit content, the contributor must belong to the DSpace community. Submission is simple, consisting of uploading the file(s) and adding descriptive information such as author, title, publication information, and keywords. See <http://www.dspace.org/>

Fedora is a flexible tool for managing and delivering digital content. It supports multiple views of each digital object and the relationships among digital objects. It can be downloaded from www.fedora.info/

Greenstone is a suite of software for building and distributing digital library collections. It has been developed in cooperation with UNESCO. For that reason, it is available in multiple languages. UNESCO is running training workshops around the world on the use of the software. The URL is www.greenstone.org/

Instant Messaging

Pidgin (formerly GAIM) is a multi-platform instant messaging client. It supports many commonly used instant messaging protocols, therefore, allowing the users to log into various different services from one application. Some three million users were estimated to be using Pidgin as of late 2007. It is released under the terms of the GNU General Public License. The URL is <http://pidgin.im/>

Media Player

VLC Media Player is a multimedia player for audio and video formats; including MPEG 1, 2, and 4; DivX, mp3, ogg, DVDs, VCDs, and various streaming protocols. It can be used on any one of a score of operating systems, including Windows, Mac OS X, and several versions of Linux. The URL is <http://www.videolan.org/vlc/>

Metasearch

Library Find is an open source application that was developed by the Oregon State University Libraries. It includes not only federated searching, but also built-in OpenURL resolver. It can also be used to locally index collections. It can be downloaded from <http://libraryfind.org/>

PC Reservation/Time Management

Groovix PAC has been designed as a complete solution for library public computers, including SIP2 authentication, reservations, and time management. It can be downloaded from http://groovix.com/solutions_public_access.html/

Reading List

LORLS 4.9, a development by Loughborough University, supports the creation and maintenance of reading lists. The URL is <http://bookworm.lboro.ac.uk/>

Single Sign-on

Shibboleth is an open source software package for web single sign-on across or within organizational boundaries. It allows its users to make authorization decisions for individual access of protected online resources in a manner that protects privacy and avoids each resource requiring its own ID and password. The URL is <http://shibboleth.internet2.edu/>

Sound Editor

Audacity is software for recording and editing sounds. It includes filters for removing static, hiss, hum, and other constant background noises from analog recordings. It is available for a wide variety of operating systems. The URL is <http://audacity.sourceforge.net/>

Statistics Tracking

Libstats is a web-based statistics tracking and knowledge base for libraries developed by the Engineering Library at the University of Wisconsin. It can also be used

to retain reference questions and answers. It is available for download under the GNU General Public License at <http://swik.net/Libstats/>

Textual Citations

TextCite 1.0.2 is software for organizing and commenting textual citations from books, articles or other published works for use in producing publications. A user can organize by publication, author, category, or outline. The URL is <http://textcite.sourceforge.net/>

User Authentication see Single Sign-on

Sources

Open source software can be found on the web by searching for specific software names or software applications, but the best sources for information about a variety of open source software packages are:

The *Apache Software Foundation* provides support for open-source software products by publishing open-source news, holding open sources conferences, offering training, and facilitating downloads of projects by members of the Apache Software Foundation. The URL is <http://www.apache.org/>

code4lib is a conference of open source developers and users that has gone beyond periodic meetings to the maintenance of a web site for sharing information about open source products among libraries. The web site is <http://code4lib.org/conference/>

The *Free Software Foundation* is the developer of the GNU open source operating system. There are many variants or distributions of the operating system, most of them using the Linux kernel. The FSF is also the creator of the GNU General Public License, the most widely used license for open source software. Finally, it is the co-

sponsor of the FSF/UNESCO Free Software Directory. The URL for the FSF is www.fsf.org/ and the URL for the directory is <http://directory.fsf.org/>

OCLC Research offers open source software developed by it or sponsored by it. The offerings include useful utilities for the development of library-oriented software. The software is written mostly in the Java language, although the web utilities are written in Perl. OCLC used its own public license, one similar to the GNUGPL, for several years, but now issues software under the Apache License. The URL is www.oclc.org/research/software/

The *Open Source Initiative* is a non-profit corporation formed to educate about and advocate for the benefits of open source and to build bridges among different constituencies in the open source community. The URL is <http://www.opensource.org/>

oss4lib started at the Yale Medical Library in 1999 as a listing of free software and systems designed for libraries. It also tracks news about projects and adoptions of open source by libraries. The URL is <http://oss4lib.org>

SourceForge.net, which claims to be the world's largest open source software development web site with some 90,000 open source projects, provides free hosting to open source software projects. It offers software developers tools for managing their projects; facilitates finding and downloading software by providing indexing and browsing tools, including a topical software map; and providing statistics on the number of downloads. The URL is <http://sourceforge.net/>

Final version, May 8, 2008