

OPEN SOURCE IN LIBRARIES – WHY AND HOW?

Irena Petrijevcanin Vuksanovic, Bojan Sudarevic
*National and University Library in Zagreb,
Center for Open Source,
Hrvatske bratske zajednice 4, HR-10000 Zagreb, Croatia
ivuksanovic@nsk.hr, bsudarevic@nsk.hr*

Abstract. Libraries and open source movement share some of the basic values and principles: openness, transparency and culture of sharing. Also, open source software offers some practical advantages particularly important for libraries, such as the focus on open standards. This has been recognized by individual libraries that use open source software and by IFLA Information Technology Section that has established the Open Source Working Group. Regardless, most public libraries still mainly use proprietary software. While some library-specific open source applications are widely used, such as various digital libraries software, others – such as library management systems – are not, mainly because of lack of technical support providers, perceived code immaturity and market saturation. However, when it comes to general purpose applications – word processors, graphics editors, database systems, spreadsheet applications and even desktop operating systems – there are no reasons not to use open source, considering that their quality and features are at equal or even higher level, in comparison with proprietary applications. Therefore, it is necessary to inform and educate library community on the benefits of open source software, in order to create prerequisites for its successful implementation and use in libraries.

Keywords. Libraries, open source software, open standards, software licensing, open source business model, cooperation, sharing.

1. Introduction

Open source software is software that can be used – without licensing cost – for any purpose, studied by examining the source code, modified, improved and distributed,

with or without modifications.

Roots of the open source movement can be traced to early 1980, to the Free Software movement and computer science departments of the few prominent US universities, mainly Massachusetts Institute of Technology and University of California, Berkeley. Since then, open source has advanced from purely idealistic approach to software development used by computer hobbyists and academia, to viable information technology strategy for businesses and governments.

Nowadays, open source software is used almost everywhere – in homes, private companies, government and public institutions – often by users that don't even know that they are using it. Mozilla Firefox, Mozilla Thunderbird, Android operating system for smartphones and tablet computers, and many other popular and widely used programs are, in fact, open source.

Some open source applications are routinely used in mission-critical environments, such as Linux operating system, that is used by, among others, NASA (Loftus, 2006), CERN, New York Stock Exchange, London Stock Exchange, The White House, German city of Munich, French National Police, French Parliament, IBM, Hewlett-Packard (Focus Editors, 2010), and even companies that are not only producing proprietary software, but also engaging in vilification campaigns against open source, such as Microsoft (Wagner, 2003; Moody, 2006; Goodin, 2012).

Yet, libraries still mostly rely on proprietary software, for performing library-specific tasks, as well as for general tasks, such as writing, web browsing and sending emails. While open source software is not the perfect solution for every task, in many situations libraries can benefit from using it. However, successful use of open source software requires deep understanding of its specifics.

2. Understanding open source

There is no inherent technical difference between open source and proprietary software. Indeed, status of the source code can be changed from closed to open and vice versa. Well-known examples of that are two popular web browsers: Mozilla Firefox and Google Chrome.

Mozilla Firefox is an open source application based on Mozilla Application Suite which was, in turn, based on Netscape Navigator, proprietary web browser whose source code was opened in 1998 by its owner, Netscape Communications Corporation. On the other hand, Google Chrome is proprietary web browser, whose new versions are first

released as an open source application, Chromium. When new code and features are thoroughly tested and declared bug-free and safe for use by general public, Google includes them in Google Chrome, which is then published as proprietary software. Key areas in which proprietary and open source software differ are licensing, costs and role in the society, with licensing being the foundation for the other two. Understanding these areas, and especially licensing, is a precondition for quality decision-making when acquiring open source software (Petrijevcenin Vuksanovic & Sudarevic, 2012).

2.1. Simplifying acquisition of software and reducing costs (at the same time!)

While proprietary applications are commercial products, sold in the same way as computer hardware, open source applications are available on the Internet for download without acquiring costs and without usual restrictions on using (on any number of computers and for any use), modifying, copying and distributing.

Thanks to that, acquiring open source software is significantly different from acquiring proprietary software – acquiring of proprietary applications must go through procurement procedures, acquisition of open source software is usually done by downloading software from the web (Petrijevcenin Vuksanovic & Sudarevic, 2012). This can significantly simplify acquiring of the software in the libraries.

2.2. Licensing

Open source software is copyrighted by its authors, but, unlike proprietary software, it is published, with source code provided, under copyright licenses that are approved as compliant with the Open Source Definition by the Open Source Initiative (Open Source Initiative, 2001).

Definition of Open Source, based on the Debian Free Software Guidelines, has been created in 1998. Since then, it is maintained by the Open Source Initiative. It defines requirements that a software license must meet in order to be considered as an open source license. Software distributed under any of the licenses approved by Open Source Initiative is considered open source software.

Currently there are 69 licenses approved by the Open Source Initiative. Besides *being compliant with the Open Source Definition, most of them are also compliant with the international legislation on copyright, to ensure their enforceability in legal jurisdictions*

worldwide.

Two main types of open source licenses are copyleft licenses and permissive licenses. Licenses of both types have certain advantages, so a license can play important role when choosing which open source application to use.

Copyleft licenses allow binary redistribution, but only if source code is also available. They allow modification of the software, but in most cases integration with other open source software is possible only if both applications are published under the same license. Some copyleft licenses, usually called weak or partial copyleft, exempt the source code from some copyleft provisions, allowing distribution of some modifications under non-copyleft terms. Using software released under copyleft license ensures that all future versions of the same application will be available as open source

Permissive licenses do not impose restrictions on what a user can do with the software, including relicensing it as proprietary. The only obligation of the distributor is to recognize the author of the original software. Use of software released under permissive license enables user to be more flexible, as applications can be relicensed and their source code can be combined with source code of other applications, but it gives no guarantee that future versions of the same application will be available as open source.

Multi-licensing is a widespread model of licensing open source software under multiple licenses. Applications licensed under multiple licenses can be distributed under any single of the used licenses. This has become very popular lately, as it enables software companies to offer enhanced versions of their own open source applications in commercial, proprietary versions. It also reduces problems caused by incompatibility between open source licenses.

2.3. Is open source really without cost?

Question that often confuses computer users that are used to proprietary software – and used to paying for it – is: can open source really be free? To understand the economics of open source, it is necessary to understand open source community and business model of the open source companies.

Most open source developers are enthusiasts, working for free. Their motivations are self-improvement, desire for peer recognition, future job offers, personal practical interest (as users of the application they develop) and enjoyment of programming (Torvalds L. & Diamond D., 2002; Lerner J. & Triole J., 2002). Communities

surrounding open source applications are open for everyone to join. Their main values are openness, transparency, cooperation and sharing.

Open source companies, like all other companies, are trying to achieve profits, but they cannot base their business model on selling software licenses, so it is based on providing technical support services to the users of the software. As there is usually more than one competing company providing support for the same application, price for support services for open source applications is usually lower than licensing costs of the proprietary software.

Of course, libraries that are using open source software are not obliged to buy support services – they can decide not to buy support services from commercial vendor, and may opt to rely on their own IT staff for providing support. That way there are no initial costs of software.

3. Open source and open standards

Open standard is a standard that is independent of any single institution or manufacturer, and to which users may propose amendments. Also, key characteristics of open standards are that anyone can use them to develop software, anyone can acquire them for free or without a significant cost, and they have been developed in a way in which anyone can participate. (Corrado, 2005)

Example of an open standard widely used by libraries is Dublin Core, vocabulary consisting of 15 metadata elements that offer expanded cataloging information and improved document indexing for search engines, maintained by Dublin Core Metadata Initiative and approved as standard by ISO. More general-use examples are OpenDocument Format (ODF), file format for text documents, spreadsheets, charts and presentations, developed by Organization for the Advancement of Structured Information Standards (OASIS) consortium and approved as standard by ISO, and Office Open XML (OOXML), like ODF, file format for text documents, spreadsheets, charts and presentations, developed by Microsoft, maintained by European Computer Manufacturers Association (ECMA) and approved as standard by ISO.

While open standards can be implemented in both proprietary and open source software, practice has showed that open source applications are more reliable in this matter. Example of problematic implementation of an open standard in proprietary applications is already mentioned OOXML. While it was developed by Microsoft and later approved by ISO, there is no implementation of the standard in any application –

not even in Microsoft Office (Paul, 2010).

On the other hand, ODF standard is implemented in several open source office suites (About OpenDocument Format, <http://opendocumentformat.org/aboutODF>).

Open standards are a precondition for preservation of long-term access to electronic data stored in libraries. As current information technology become obsolete, it will be necessary to migrate data to new systems. Without the assistance of the software manufacturer, proprietary software may make migration practically impossible. By using open source software and open standards, libraries can ensure continued access to their data, long into the future. (Corrado, 2005)

4. Open source in libraries

Values and culture of the open source community – openness, transparency, cooperation, sharing, encouraging community interaction and contributor involvement – also reflect values and culture of the libraries (Jaffe & Careaga 2007; Jones, 2001).

While libraries are focusing on providing access to as much information as possible to the citizens, open source community is creating software with source code available for anyone to use, modify and distribute.

Values of these two communities are not only compatible – they also mutually complement each other in a practical manner. The only way in which libraries can be sure that in the future they will have access to data stored in their computer systems and be able to provide that access to the citizens, is to have complete control over those systems – not only technical control but also legal. The best way to achieve control over those systems is to have full access to their source code, and unrestricted right to adapt it according to their own needs. The easiest way to achieve that is to use open source software.

Importance of open source software for libraries has been recognized by the IFLA IT Section, that has established the working group on Open Source. The role of this group is to foster support and development of open source systems for libraries (IFLA, 2011).

4.1. Library-specific software

Due to their needs, libraries are using some specific applications, the most important ones being library management systems (LMS) and applications for building and maintaining digital libraries.

Competing with proprietary library management systems is very difficult for open source applications. Considering the fact that this market is very mature – some proprietary library management systems exist for more than three decades – and saturated (most libraries, especially national and other large libraries ones, are already using some LMS), there isn't much space for new systems. Also, providers of technical support for open source systems cannot compete with proprietary vendors, nor in the size, nor in the market coverage. Nevertheless, open source library management systems do exist, and the two most prominent are Koha (used mainly by individual libraries) and Evergreen (used by library consortia). In most libraries, Koha and Evergreen are supported by the system librarians and IT staff of the libraries. Their market share is on the rise as they are becoming more feature-rich.

Situation is quite different in the field of applications for building and maintaining digital libraries, institutional repositories, digital archives and digital asset management in general. Since the market of digital library systems is younger than LMS market – rapid advancement in this field has started in the late 1990's – open source applications were present from the beginning. As a result, open source systems such as DSpace (developed by Massachusetts Institute of Technology and Hewlett-Packard), Greenstone (created by New Zealand Digital Library Project and UNESCO) and EPrints (created by the University of Southampton) are widely used (Patil and Kanamadi, 2009).

Some open source applications are used by large national and academic libraries, such as VuFind, library search engine that can be used for advanced searching and browsing of library catalogs and digital collections. VuFind, used by National Library of Australia, National Library of Ireland, London School of Economics, and others (VuFind Installations. http://vufind.org/wiki/installation_status), serves as an example of open source application that complements already existing systems – both proprietary and open source – in an unobtrusive manner.

4.2. Other software used by libraries

Libraries also use non-specific software, such as operating systems, web browsers, word processors, graphics editors, presentation programs, database systems, spreadsheet applications, email clients, multimedia players, various server applications, etc.

There are high-quality open source applications for performing all of these tasks –

Linux operating system (in variants for desktop workstations and servers), LibreOffice office suite (that includes word processor, spreadsheet application, presentation program and simple database system), VLC Media Player, Mozilla Firefox web browser, Mozilla Thunderbird email client, Apache web server, PostgreSQL and MySQL relational database systems, etc. All aforementioned applications are comparable to their proprietary alternatives, in terms of quality, ease of use and capabilities. If individual library uses the application that depends on a proprietary operating system, Linux probably cannot be used, but other applications still can.

Using general-purpose open source software can significantly reduce costs, considering that all these applications are free, in contrast to their proprietary counterparts.

4.3. To implement or not to implement?

First stage of implementation process is deciding whether to implement open source application in particular situation, as well as choosing the appropriate application. In the process of making those decisions, certain factors should be taken into account: evaluation of current state and needs of the library, features of available open source solutions (in contrast to proprietary solutions), availability of technical support and capabilities of library's own IT staff.

If a library is already using proprietary application for particular task, and if it is satisfied with those applications and capable of securing funds for licensing cost for the foreseeable future, switch to the open source may not be the best solution, especially in the case of mission critical systems.

On the other hand, if there are problems with currently used proprietary applications, or if funding is not secured, or if a library simply wants to achieve greater control over its computers systems and data – Implementation of open source alternatives should be considered.

Also, potential drawbacks should be taken into account. While open source can bring significant cost reduction and help libraries to achieve greater independence from software vendors, it is not without its own issues. In the past, drawbacks of open source software included lower quality of documentation, fragmentation into multiple projects through “forking”, less user-friendliness than commercial software, and lower functionality than commercial equivalents (Chawner, 2004). While these are all true for new open source projects, key open source applications have significantly matured

during the last decade.

4.4. Gradualism, cooperation and staff training – keys to success

Implementation of open source software – especially if it involves switch from existing, proprietary, application – should be well planned and gradual, as it is important not to disrupt currently existing services. It would be wise to use both systems – old, proprietary, and new, open source – in parallel, until new system becomes fully operational, and for as long as is necessary for the library and IT staff to master the operation of the new software.

Cooperation with other libraries using the same software should be realized during early phases of implementation process. Through cooperation with other users of the same application, it is possible to achieve a level of support comparable to the commercial support, but without cost or at very low cost. Skills acquired in this way can be transferred to the rest of the library staff.

Also, it would be useful for members of the IT staff responsible for administering the new system, to join community of developers of the application and to acquire deep understanding of inner workings of the application.

5. Conclusion

There is no inherent technical difference between open source and proprietary software. Key areas in which they differ are licensing, costs and role in the society. Understanding these areas, and especially licensing, is a needed when acquiring open source software.

In many situations libraries can benefit from using open source software, although it is not the perfect solution for every task. Market of library management systems is very mature and saturated, so there isn't much space for new, open source systems. On the other hand, situation is quite different in the field of applications for building and maintaining digital libraries, where open source systems are widely used. Also, there are high-quality open source applications for performing tasks that are not specific for libraries.

The easiest way to ensure permanent access to data stored in computer systems of libraries is to use open source software. In that way, libraries can achieve both technical and legal control over their data and systems. In addition to that, this is an

example of mutually complementary values, shared by libraries and open source community in a practical manner.

When the library makes a decision to use certain open source application, keys for success are gradual implementation, cooperation – both with open source community and with other libraries – and continual staff training, preferably through direct involvement in development of the application.

References

About OpenDocument Format. Retrieved May 30, 2012 from

<http://opendocumentformat.org/aboutODF>.

Chawner, B. (2004). Free/Open Source Software: New Opportunities, New Challenges.

Retrieved May 30, 2012 from

<http://www.vala.org.au/vala2004/2004pdfs/33Chawn.PDF>.

Corrado, E.M. (2005). The Importance of Open Access, Open Source, and Open Standards for Libraries. *Issues in Science and Technology Librarianship*, 42.

Chicago: ISTL. Retrieved June 16, 2012 from <http://www.istl.org/05-spring/article2.html>.

Focus Editors (2010). 50 Places Linux is Running That You Might Not Expect. Focus.

Retrieved June 29, 2012 from <http://www.focus.com/fyi/50-places-linux-running-you-might-not-expect>.

Goodin, D. (2012). Skype replaces P2P supernodes with Linux boxes hosted by

Microsoft. *Ars Technica*. Retrieved June 15, 2012 from

<http://arstechnica.com/business/2012/05/skype-replaces-p2p-supernodes-with-linux-boxes-hosted-by-microsoft>.

IFLA (2011). Open Source Working Group. Retrieved May 30, 2012 from

<http://www.ifla.org/en/it/opensource>.

Jaffe, L.D., & Careaga, G. (2007). Standing up for Open Source. *Library Philosophy & Practice*, Summer 2007, Vol. 9 Issue 3, p. 1-17.

Jones, P. (2001). Open(source)ing the Doors: for Contributor-run Digital Libraries.

Communications of the ACM; May 2001, Vol. 44 Issue 5, p45-46.

Lerner J. & Triole J. (2002). Some Simple Economics of Open Source. *Journal of Industrial Economics*, 52 (p. 197-234).

Loftus, J. (2006). Linux sends NASA rovers to Mars, among other things. *Search Enterprise Linux*. Retrieved June 30, 2012 from

- <http://searchenterpriselinux.techtarget.com/news/1157924/Linux-sends-NASA-rovers-to-Mars-among-other-things>.
- Moody, G. (2006). A Brief History of Microsoft FUD. LXR. Retrieved June 15, 2012 from <http://lxe.com/module/newswire/view/57261>.
- Open Source Initiative (2001). The Open Source Definition (Annotated), Version 1.9. Retrieved May 30, 2012 from <http://www.opensource.org/osd.html>.
- Patil, M.S. & Kanamadi (2008). Digital Library Open Source Software: A Comparative Study. In: *Proceeding of International Conference on "Knowledge For All: Role of Libraries and Information Centres"*. Mumbai: ILA and TISS. Retrieved June 13, 2012 from http://kud.academia.edu/m spatil/Papers/191628/Digital_Library_Open_Source_Software_A_Comparative_Study.
- Paul, R. (2010). Microsoft: Office 15 – not 2010 – to be fully OOXML compliant. Ars Technica. Retrieved June 11, 2012 from <http://arstechnica.com/information-technology/2010/04/microsoft-office-15not-2010to-be-fully-ooxml-compliant>.
- Petrijevcenin Vuksanovic, I. & Sudarevic, B. (2012). Toward a Framework for Implementing Open Source Software in Public Administrations: Legal, Economic and Societal Aspects. In: V. Luzar-Stiffler, I. Jarec, Z. Bekic (ed.), *Proceedings of the ITI 2012 34th International Conference on Information Technology Interfaces* (p. 293-298). Cavtat: University of Zagreb, University Computing Centre.
- Richmond, G. (2008). The Large Hadron Collider switches on. If it's the end of the world, it will be powered by GNU/Linux. Free Software Magazine. Retrieved June 29, 2012 from http://www.freesoftwaremagazine.com/articles/large_hadron_collider_switches_if_its_end_world_it_will_be_powered_gnu_linux.
- Torvalds L. & Diamond D. (2002). *Just for Fun*. New York: HarperCollins Publishers.
- VuFind Installations. Retrieved June 30, 2012 http://vufind.org/wiki/installation_status.
- Wagner, M. (2003). Microsoft Uses Linux To Publish Its Own Web Site. CRN. Retrieved June 15, 2012 from <http://www.crn.com/news/applications-os/18839747/microsoft-uses-linux-to-publish-its-own-web-site.htm>.