Board accepted

The Importance of Open Access, Open Source, and Open Standards for Libraries

Edward M. Corrado
Systems Librarian
The College of New Jersey
corrado@tcnj.edu

Abstract

The open access, open source software, and open standards concepts have been garnering increased attention in the field of librarianship and elsewhere. These concepts and their benefits and importance to libraries are examined. Benefits include lower costs, greater accessibility, and better prospects for long-term preservation of scholarly works.

Introduction

Open access, open source software, and open standards are three concepts that have been receiving increased attention lately in the library world. Open access is seen by some as a possible solution to the increasing price of serials and as a way for governmental funding agencies to receive a better return on investment. Open source software can benefit libraries by lowering initial and ongoing costs, eliminating vendor lock-in, and allowing for greater flexibility. Open standards allow for interoperability to exist between diverse library resources and eases data migration between systems. All three of these concepts are important to libraries individually and they can be even more beneficial when they are leveraged simultaneously.

Open Access

Open access to scholarly information has been a hot topic for debate among librarians, scholars, and publishers over the last few years. Recent proposals by

the National Institutes of Health (NIH) in the United States (requiring for scholarly works that come out of NIH funded research to be made available via NIH's PubMedCentral open access database), by the government in the United Kingdom (requiring that all UK government-funded research to be available via open access), and by others has expanded this debate. Various different, though similar, definitions of open access exist with the *Budapest Open Access Initiative* definition being the most widely used (Goodman 2004). Other definitions include the *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities,* the *Bethesda Statement on Open Access Publishing*, and the *Washington DC Principles for Free Access to Science*. While there are multiple definitions and flavors of open access, open access basically calls for scholarly publications to be made freely available to libraries and end users.

Willinsky (2003) identified nine flavors of open access. The flavors are: 1) e-print archive (authors self-archive pre- or post-prints), 2) unqualified (immediate and full open access publication of a journal, 3) dual mode (both print subscription and open access versions of a journal are offered), 4) delayed open access (open access is available after a certain period of time), 5) author fee (authors pay a fee to support open access), 6) partial open access (some articles from a journal are available via open access), 7) per-capita (open access is made available to countries based on per-capita income), 8) abstract (open access available to table of contents/abstracts, and 9) co-op (institutional members support open access journals).

The growth of the open access movement is partially in response to the enormous costs of many scholarly journals. With traditional journal publication methods it is not uncommon for an institution to have to pay for an article twice. First they pay scholars to produce the work and then the institution's library pays to purchase the work back from the journal publisher. Anderson (2004) is correct that there is no such thing as free information and that there are costs involved in producing scholarly information. However, with the advent of new technologies and software programs, it is becoming increasingly less expensive to compile and distribute scholarly information. By using different funding methods and electronic delivery of journals, the costs can be absorbed by alternative means to subscription fees. One of the great benefits to open access is that libraries in smaller institutions or in economically disadvantaged areas around the world can have greater access to these scholarly resources.

Open access helps to ensure long-term access to scholarly articles. Unlike articles that are licensed in traditional article databases, libraries and others can create local copies and repositories of these resources. Libraries, by working together to make repositories of open access literature, can ensure continued access to these scholarly publications into the distant future.

Open Source

Open source software is software that includes source code and is usually available at no charge. There are additional requirements besides the availability of source code that a program must meet before it is considered open source including: the software must be free to redistribute; derivative works must be allowed; the license can not discriminate against any persons; and the license cannot discriminate against any fields of endeavor. Software that is licensed under an open source license allows for a community of developers from around the world to improve the software by providing enhancements and bug fixes.

Libraries can realize many advantages by using open source software. One of the most obvious advantages is the initial cost. Open source software is generally available for free (or at a minimal cost) and it is not necessary to purchase additional licenses for every computer that the program is to be installed on or for every person who is going to use the software. Open source software not only has a lower acquisition cost than proprietary software, it often has lower implementation and support costs as well.

It is easier to evaluate open source software then proprietary software. Since open source software is typically freely available to download, librarians and systems administrators can install complete production-ready versions of software and evaluate competing packages. This can be done not only without any license fees, but also without having to stick to a vendor's trial period, evaluate a limited version of the software, or deal with the vendor's sales personnel. If the library likes an overall open source package but would like a few added features they can add these features themselves. This is possible because the source code is available. Even if a library does not have in-house expertise they can benefit from source code availability because another library may be able to provide them the fix or they can hire a consultant to make the changes that they desire. Fuchs (2004) points out that if a proprietary program "is deficient in some way [the user] must wait until the vendor decides it is financially viable to develop the enhancement -- an event that may never occur." With open source software the user can develop the enhancement themselves.

Open source software allows for more support options. Proprietary software vendors often package service with the product. This is particularly true of proprietary library-specific software. When support from a vendor is inadequate it is an additional expense to purchase another tier of support, assuming that it is even available. Open source software allows for different vendors to compete for support contracts based on quality of service and on price. Access to the source code also allows for self-support when practical and desired.

The amount of vendor lock-in is dramatically reduced with open source software. The large initial costs often associated with proprietary software makes it difficult to reevaluate the choice of software when it does not live up to expectations. Proprietary software can lead to a single point of failure. If a vendor goes out of business or decides not to support a program anymore there is often nothing an

user can do. Organizations using the software could provide self support or other vendors can come in and fill the void left by the previous vendor if the program were available as open source software.

Open Standards

Pountain (2003) defines an open standard as "a standard that is independent of any single institution or manufacturer, and to which users may propose amendments." This definition is a good starting point, but in reality the term "open standard" means different things to different people. Three key characteristics of open standards identified by Coyle (2002) are 1) that anyone can use the standards to develop software, 2) anyone can acquire the standards for free or without a significant cost, and 3) the standard has been developed in a way in which anyone can participate. When a standard has the first two of these characteristics (the ability to use the standard and to obtain it with out a significant cost) it can be said to be an open standard in an utility sense. That is to say that an open standard is a standard that is not encumbered by a patent, does not require proprietary software, and can be utilized by anyone without cost. Proprietary standards can sometimes be expensive and it may be cost prohibited to purchase access to a proprietary standard if it is ever needed. Many people consider a standard to be sufficiently open as long as it is open in a utility sense. Others take this a step further and consider a standard to be open only if the process meets the criteria of being created and modified in an open process as well. An example of a standard that fits the definition of a standard that is open in utility but not in process is XHTML. In order to help develop the XHTML specification one has to be a member of W3C. In order to become a member of W3C businesses pay between $5,000 and $50,000 per year (Coyle 2002). Conversely, Dublin Core is a completely open standard that is open both in utility and in process. All one has to do is show up and participate in order to contribute to the development of Dublin Core.

It is important for libraries and other cultural institutions to ensure long-term access to digital information. The rapid growth in digital technologies has led to new and improved applications for digital preservation. However at the same time it has also led to some problems as well. Two of these problems are obsolescence and dependency issues. The obsolescence problem is caused by the advances in hardware and software making many computers obsolete within three to five years (Vilbrandt et al. 2004). Dependency problems can arise if tools that are needed to communicate between systems or read file formats become unavailable. In order to account for obsolescence and dependency problems organizations must be able to migrate data into new systems. Data migration, however, cannot occur without access to data file formats.

Properly created open standards for file formats are less likely to become obsolete (Vilbrandt et al. 2004) and are more reliable and stable then proprietary formats

(Breeding 2002). In the event that an open standard file format does become obsolete, having access to the file format would allow anyone to easily, and legally, create a data conversion utility. File formats that use open standards can assist in long-term archiving because they allow for software and hardware independence. Open standards help alleviate issues caused by obsolescence or dependency problems since files created in formats that adhere to open standards are "more likely than proprietary formats to be readable twenty or fifty years from now" (Baker 1999). This allows for greater flexibility and easy migration to different systems in the future.

The use of open standards can help assure interoperability of diverse systems. There are various software packages that are being used to create digital libraries, online library catalogs, and other resources that libraries relay on. These various systems need to be able to interact in order to provide the best possible service to patrons. The way to make certain that these diverse systems, and any future systems, can communicate with each other is by using open standards to help achieve the "free flow of information through interoperability" (The Open Group 2005).

Many different organizations are advocating open standards. One of the most prominent organizations is The Open Group which created the *Developer Declaration of Independence*. The hope is that the *Developer Declaration of Independence* will help pull together the information technology industry in support of open standards. Some library-centric initiatives, including the Open Archives Institute (OAI), also support open standards. OAI's mission is to develop and promote "interoperability standards that aim to facilitate the efficient dissemination of content" (Open Archives Institute 2005). OAI has created a Protocol for Metadata Harvesting (OAI-PMH) that provides an application-independent interoperability framework based on metadata harvesting. Other common open standards for information retrieval relevant to libraries include Digital Object Identifier System (DOI), Dublin Core Metadata Initiative (DCMI), and OpenURL.

While open standards have garnered increased attention in libraries recently, the use of open standards in librarianship is not new. The use of open standards in librarianship can be traced all the way back to the first American Librarian Association meeting in 1877 when the dimensions of the card catalog were standardized to 7.5 x 12.5 centimeters (Coyle 2002). A more modern example of an open standard used by libraries is the Machine-Readable Cataloguing (MARC) record. Other common open standards for bibliographic data include Metadata Object Description Schema (MODS) , Metadata Encoding & Transmission Schema (METS), and the XML Organic Bibliographic Information Schema (XOBIS).

Putting Open Access, Open Source, and

Open Standards Together

Open access, open source software, and open standards each individually offer a number of significant benefits to libraries. When they are combined the results can be even greater. Open source and open standards can help libraries provide patrons with easier access to open access materials and other resources. There are literally thousands of open access titles available and without open standards it would be very difficult to find what one is looking for or to view various articles. Imagine the difficulty, and costs involved, in maintaining a library's information technology infrastructure if each electronic journal required a separate, proprietary piece of software to read or search the journal. Open standards make it possible to create interoperable systems to access the literature in various open access journals seamlessly.

Open standards and open source can help preserve long-term access to open access and other types of electronic journals. Libraries working together can use open source software such as LOCKSS to ensure continued access to these scholarly publications long into the future. LOCKSS (short for "Lots Of Copies Keeps Stuff Safe") is a system that caches copies of digital collections around the world. As current computers, software, storage media, file formats, and other types of information technology become obsolete, it will be necessary to migrate open access articles and other data to new systems. Without the assistance of the software manufacturer (who may or may not even still be in business, let alone willing to help) proprietary software and file formats may make migration practically impossible. By utilizing open source software and open standards from the beginning, libraries can assure that this type of systems migration will be possible years down the road.

Not only has the growing cost of serials caused libraries to drop journal subscriptions, it has also factored into a 26% decrease of monograph acquisitions by the typical research library between 1986 and 1999 (Create Change 2002). Library budgets can be reallocated to monographs and other areas because of the lower costs typically involved with open access, open source, and open standards.

Conclusion

These benefits of open access, open source, and open standards are numerous. The benefits include lower costs, great accessibility, and better prospects for long-term preservation of scholarly works. Libraries should embrace all three of these concepts now and in the future. By supporting open access, open source, and open standards libraries not only can help ensure that their current and future patrons will have easier and more comprehensive access to scholarly research, they will also be helping other libraries around the world, including those in disadvantaged areas, to have access to important scholarly research.

References

Anderson, R. 2004. Open access in the real world: confronting economic and legal reality. *College and Research Library News* 64(4). [Online]. Available: **{http://dlist.sir.arizona.edu/351/}** [Accessed March 25, 2005].

Baker, T. 1999. *TIAC White Paper on Appropriate Technology for Digital Libraries*. Bangkok: Technical Information Access Center.

Breeding, M. 2002. Preserving digital information. *Information Today* 19(5): pp. 48-49.

Coyle, K. 2002. Open source, open standards. *Information Technology and Libraries* 21(1): 33-36.

Create Change. 2002. *Coping Strategies*. [Online]. Available: **{http://www.createchange.org/librarians/issues/coping.html}** [Accessed May 11, 2005].

Fuchs, I. 2004. Learning management systems: are we there yet? *Syllabus*. [Online]. Available: **{http://campustechnology.com/articles/39862/}** [Accessed: April 1, 2005].

Goodman, D. 2004. The criteria for open access. *Serials Review* 30(4). [Online]. Available: **{http://dlist.sir.arizona.edu/798/}** [Accessed: March 11, 2005].

Open Archives Initiative. 2005. Organization. [Online]. Available: http://www.openarchives.org/organization/index.html [Accessed: March 17, 2005].

Pountain, D. 2003. *The Penguin Dictionary of Computing*. New York: Penguin Putnam.

The Open Group. 2005. *Developer Declaration of Independence*. [Online]. Available: http://www.opengroup.org/declaration/declaration.htm [Accessed: March 29, 2005].

Vilbrandt, T., et al. 2004. Cultural heritage preservation using constructive shape modeling. *Computer Graphics Forum* 23(1): 25-41.

Willinsky, J. 2003. The nine flavors of open access scholarly publishing. *Journal of Postgraduate Medicine* 49: 263-267.

Previous Contents Next