

F/OSS in the Library World: An Exploration

Brenda Chawner

School of Information Management

Victoria University of Wellington,

Wellington, New Zealand

+64 4 463 5780

brenda.chawner@vuw.ac.nz

ABSTRACT

Existing research into free/open source software development has largely ignored the nature of the application domain the software is for. Jørgensen and Sørensen's 'development arena' [12] provides a useful conceptual framework for grouping and comparing projects. This paper applies the framework to free/open source software projects in two library and information management application spaces: integrated library systems and institutional repositories. The preliminary analysis shows that F/OSS projects to develop integrated library systems are generally small scale, and the software is used mainly by smaller libraries. Institutional repository software, not widely available from commercial vendors, is being developed by larger libraries, often with the assistance of research funding from other bodies, and is being implemented by university and research libraries worldwide.

Categories and Subject Descriptors

J.1 [Administrative Data Processing]: Education

General Terms

Economics.

Keywords

Institutional repository software; Library management systems; Open source software development

1. INTRODUCTION

Little of the research into free/open source software (F/OSS) development has taken account of the nature of the application domain the software is for. Empirical research has found that there is a wide variation in F/OSS project activity, with a large number of projects that are either inactive, or have only one developer [6, 13]. As yet we lack a conceptual framework that helps explain this observation. Jørgensen and Sørensen's concept of a 'development arena' [12], initially used to study the development of new technologies, provides a conceptual framework to group related projects. Using this framework, it is possible to compare projects in a single development arena, or to compare patterns in different development arenas. In order to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Open Source Application Spaces: Fifth Workshop on Open Source Software Engineering (5-WOSSE) May 17, 2005, St Louis, MO, USA.

Copyright 2005 ACM 1-59593-127-9 ... \$5.00.

identify suitable development

arenas for such an analysis, we can look at a specific application domain, such as library and information management.

Members of the library and information management community have promoted free/open source software since the late 1990s. Morgan [16] noted the similarity between the 'gift culture' of F/OSS projects and the well-established tradition of freely sharing information in the field of library and information management, suggesting that there is a natural synergy between the two. Chudnov viewed F/OSS as a solution to the issues many libraries experienced with commercial library management software that was slow to evolve and expensive to upgrade [8]. An F/OSS alternative might result in lower costs, but also, more importantly, more flexibility and a better match to libraries' requirements [11]. Tennant [20] suggested that F/OSS development could be important in developing prototypes of digital libraries because of the relatively low costs associated with its use. Morgan suggested library staff could increase their technical expertise by being involved with F/OSS projects [16].

The Open Source Software for Libraries portal (<http://www.oss4lib.org>), established in mid-1999, listed some 90 library-related projects by mid-January 2005. These range from simple scripts to produce statistics to integrated library systems to institutional repository software. A closer examination reveals that while some projects are active (for example Koha, Greenstone, ePrints, and dSpace), others are relatively inactive (OSCR), abandoned (course / control), or withdrawn (FreeReserves). This suggests that library-related F/OSS projects share similar characteristics with the wider population of all F/OSS projects.

The purpose of this paper is to:

- ☐ introduce the concept of a development arena (DA)
- ☐ use the DA framework to describe and compare free/open source software projects in two library and information management application spaces: integrated library systems and institutional repositories.

2. DEVELOPMENT ARENAS

Jørgensen and Sørensen [12] define a 'development arena' (DA) as the context for a product or process development. A DA includes:

- ☐ actors (i.e. people and organizations), objects, and standards;
- ☐ logical and physical locations where changes occur; and
- ☐ translations that stabilize and destabilize relationships.

The concept of a DA provides an open-ended spatial metaphor for studying technological development, particularly where there are opportunities for cooperation and/or competition. To use an ecological metaphor, a DA represents the nature of the environment being studied, such as a wetland, a West Coast

beach, or a prairie—it gives us a sense of the ‘place’ where the development is occurring. The term ‘translation’ comes from Callon and Latour’s Actor-Network Theory (ANT) [5, 14], a sociological framework for studying the process of technological change. ANT defines translation as the process through which new actors join an existing network, taking the perspective that this involves a translation of the newcomer’s interests to match those of the network.

Jørgensen and Sørensen illustrate the DA framework by examining the development of high-definition television (HDTV), which had separate actor-networks in Japan, Europe, the United States, and Denmark working on HDTV development projects; the multinational company Philips was also involved. They found that each site used a different strategy in their HDTV project:

- the Japanese used resettling and inclusion to extend the existing analog network;
- * the European initiative involved extension and differentiation, by setting up a competing network in collaboration with satellite broadcasters rather than traditional national broadcasters;
- the U.S. had a strong technological and standards focus, and used a strategy of exclusion to overtake the Japanese and European projects;
- the Danish project resulted in a working television, with a strategy of re-framing the problem; and
- Philips had a strategy of multiple engagements, by being involved in both the European and U.S. initiatives, to deal with the uncertainty of the final outcome.

The main focus in this paper will be the actors and locations involved in the various projects. Two types of actors, developers and users, will be identified for each project, where possible.

3. THE INTEGRATED LIBRARY SYSTEM DA

An integrated library system (ILS) provides “a wide variety of housekeeping activities for the management of libraries. Modules in general use include cataloguing, acquisition, circulation, OPAC [online public access catalogue], interlibrary loan, and periodicals control” (*Harrods librarians’ glossary*, 8th ed., Aldershot: Gower, 1995). ILSs have their origins in stand-alone software developed in the 1960s and 1970s, and began to take their current form in the mid-1980s [2]. The ILS DA is well-established, with annual U.S. revenues in the order of \$500 million [4]. A mature ILS can have more than 1 million lines of code [3]. Roughly 30 companies are active in the global ILS market, some of whom concentrate on specific sectors, such as school or corporate libraries, or focus on particular countries. No F/OSS ILS projects have been identified as significant components of the ILS marketplace [4].

Breeding [3] has argued that there is little possibility of a F/OSS ILS having much impact on the market; this is in part because switching ILS systems is both costly and time-consuming [1]. While some ILS vendors have migrated their systems to Linux platforms, none has made a commitment to F/OSS development of their ILS package. Despite this apparent lack of potential of F/OSS projects in this DA, there are at least 10 F/OSS ILS projects, in different stages of development. The following section lists each project, along with their key characteristics.

Avanti (U.S.): version 1.0 due in “early 2005”. 1 developer. No information about users.

Emilda (Finland): version 1.2.1 released on 21 January 2005. 6 developers. No information available about users.

Gnuteca (Brazil): version 1.5 released on 8 September 2004. Information provided only in Portuguese.

Koha (New Zealand): version 2.2.0 released on 5 January 2005. Koha is the oldest F/OSS ILS, first released under the GPL in early 2000. It was originally developed for the Horowhenua Library Trust, a small rural public library system with 3 branches. Koha has the largest number of developers of the F/OSS ILS projects (49 in 5 countries, but only a small proportion are active at any given time). It has become an international project, with coordinators for different releases being located in New Zealand, Canada, France, and the United Kingdom. The software is being used by at least 49 libraries in 15 countries, with the majority very small libraries, who are not significant users of commercial ILS software. One U.S. public library has adopted Koha, and is active in the ongoing development of the software.

LearningAccess ILS (U.S.): no code released and appears to be inactive as of February 2005

Obiblio (U.S.): version 0.5.1 released on 13 March 2005. 6 developers. No information available about users.

OpenBiblio (Belgium): version 2.0.3 released on 4 February 2005. 1 developer. No information available about users.

PhpMyBibli (France): version 2.0.1 released on 10 February 2005. 3 developers. Information provided mainly in French.

PhpMyLibrary (Philippines): version 2.0.3 released on 16 November 2005. 6 developers 15 known sites, most in the Philippines

PINES ILS (U.S.): in early development, first code released in early January 2005, with an alpha version scheduled for mid-2005. A consortium of 249 Georgia public libraries is funding the project, and their stated reason for starting the project is that they would “have complete control over the functionality of the system, and can truly tailor the ILS to the very specific needs and wants of the PINES community (public and staff)” [17]

This shows that F/OSS activity in the ILS DA is scattered, with no evidence of significant cooperation between the various projects. Several projects, particularly Gnuteca and PhpMyBibli, have a strong local focus. These systems are used at the lower end of the ILS market, by libraries who are not perceived by commercial vendors as potential customers. Most projects have started as the classical F/OSS single developer (or in some cases an organization) who is “scratching a personal itch” [18]. It is too early to see whether the PINES ILS project will have a different outcome, but this project has the most potential to challenge Breeding’s prediction, given its large potential user base.

4. THE INSTITUTIONAL REPOSITORY DA

An institutional repository (IR) is a digital collection “capturing and preserving the intellectual output of a single university or a multiple institution community” [8]. The concept emerged in the fall of 2002 [15], growing out of the eprint and open access/open archives movements. At that time, no commercial software provided full IR functionality. While commercial IR packages are now available (largely from ILS vendors) [7], open source IR packages have a high profile in the library practitioner literature. The IR DA is therefore less mature than the ILS one, and we might reasonably expect a different pattern of development.

The Open Society Institute guide to institutional repository software [10] identifies 9 open source packages that provide IR functionality. The following section summarizes their characteristics:

Archimede (Canada): version 0.7.5 released on 2 February 2005. Developed by Laval University, with three developers and 1 reported site.

ARNO (The Netherlands): version 1.1 released on 18 June 2004. Funded by a Dutch scientific body, and with 6 reported sites, all in the Netherlands.

CDSware (CERN Document Server Software, Switzerland): version 0.5.0 released on 17 December 2004; number of sites not reported, but 5 listed in OSI guide [x]

DSpace (United States): version 1.2.1 released on 4 February 2005. Developed jointly by MIT and Hewlett Packard, with a grant from the Andrew W. Mellon Foundation. 7 committers, 15 contributors. Over 70 reported sites, in North America, Europe, the United Kingdom, Asia, Australia, and South America.

Eprints (United Kingdom): version 2.3.8 released on 16 February 2005. Developed by the University of Southampton with the assistance of funding from the UK Joint Information Systems Committee. 5 developers, and 150 reported sites in the United Kingdom, the United States, Australia, Europe, South America, and Asia. Eprints is the oldest IR project, having been first released under the GPL in mid-2002.

Fedora (United States): version 2.0 released on 31 January 2005. Developed by the University of Virginia Library and Cornell University, with a grant from the Andrew W. Mellon Foundation. 10 member development team. 22 reported sites, and one ILS vendor has chosen Fedora as the basis of its digital asset management software. (The name Fedora comes from 'Flexible Extensible Digital Object and Repository Architecture', and is not related to Red Hat's Fedora project.)

i-Tor (The Netherlands): version 1.1.3a released on 3 January 2005. Developed by the IT section of the Netherlands Institute for Scientific Information Services. i-Tor can be used as a website content management system as well as providing the foundation for an IR. 8 developers. 2 reported IR users, both in the Netherlands; 29 reported website users.

MyCoRe (Germany): version 0.9.3 released on 23 November 2004. Developed by a consortium of German universities. 6 sites reported in OSI guide, in Germany and Sweden. Site and documentation in German.

OPUS (Germany): release information not available. Based on a University of Stuttgart project to allow staff to manage electronic publications, 37 users reported in OSI guide, all in Germany.

All of the users identified for the various projects are at the larger end of the library population, with many being significant universities, national libraries, or research institutions.

The IR DA shows both similarities and differences in the patterns of development to that of the ILS DA. Like the ILS DA, there is little cooperation between projects, and several projects that have a strong local focus, particularly in non-English speaking countries. Unlike the ILS DA, though, all of the projects started with some, and in many cases significant, institutional support. Universities (either as individuals or consortia) and research institutions have taken the initiative in developing IR software and making it available under F/OSS licenses. The absence of viable commercial alternatives at the time the need for an IR was recognized, combined with the availability of initial

funding from external sources, meant that the different projects got off to a strong start. One project, DSpace, involved a commercial partner. The implementation of these packages outside their originating institutions in a range of countries shows that there is widespread acceptance of a F/OSS approach to this application space. The adoption of the Fedora software by a commercial vendor suggests that there is a different perception of the position and significance of F/OSS developments in this area.

Compared to the ILS DA, most of these projects are relatively active, with new versions of the software appearing in the last few months. This might be in part because of the relative newness of this application space resulting in changing requirements, but it might also reflect that availability of funding support from larger institutions and funding bodies allowing more staff to work on the projects.

5. CONCLUSION AND FUTURE RESEARCH

The DA concept provides a framework that can expose different patterns of F/OSS project development and adoption, as these two library-related examples show. In the ILS DA, the presence of mature commercial software results in limited interest and support for F/OSS options, while in the IR DA, there is considerable institutional support for both the development and use of F/OSS applications. Future research will initially involve more detailed analysis of the users of both types of software. A longitudinal study, particularly of the IR DA, would show whether the patterns identified here continue, or whether the availability of commercial IR software reduces interest in F/OSS versions. Comparing other relatively mature application spaces, for example document management systems, with newer ones, such as web content management systems, would show if similar patterns to those seen in the ILS and RA development arenas can be identified.

6. REFERENCES

- [1] Barreau, Deborah. 2001. The hidden costs of implementing and maintaining information systems. *The Bottom Line: Managing Library Finances* 14 (4):207-212.
- [2] Borgman, Christine. 1997. From acting locally to thinking globally: a brief history of library automation. *The Library Quarterly* 67 (3):215-249.
- [3] Breeding, Marshall. 2002. The open source ILS: only a distant possibility. *Information Technology and Libraries* 21 (1):16-18.
- [4] Breeding, Marshall. 2004. Automated system marketplace 2004: migration down, innovation up. *Library Journal* 129 (6):46-50, 52, 54, 56-58.
- [5] Callon, Michel. 1986. The sociology of an actor-network: the case of the electric vehicle. In *Mapping the dynamics of science and technology: sociology of science in the real world*, edited by M. Callon, J. Law and A. Rip, Macmillan, Basingstoke, 19-34.
- [6] Capiluppi, Andrea, Patricia Lago, and Maurizio Morisio. 2003. Evidences in the evolution of OS projects through changelog analysis. In *Taking Stock of the Bazaar: Proceedings of the 3rd Workshop on Open Source Software Engineering*, Portland, Oregon, 3 May 2003, edited by J. Feller, B. Fitzgerald, S. Hissam and K. R. Lakhani.

Available from: <http://opensource.ucc.ie/icse2003/3rd-WS-on-OSS-Engineering.pdf>

- [7] Cervone, H. Frank. 2004. The repository adventure. *Library Journal* 129 (10):44-46.
- [8] Chudnov, Daniel. 1999. Open source software: the future of library systems. *Library Journal* 124 (13):40-43.9.
- [9] Crow, Raym. 2002. SPARC institutional repository checklist & resource guide. Washington, D.C.: The Scholarly Publishing & Academic Resources Coalition. Available from http://www.arl.org/sparc/IR/IR_Guide_v1.pdf
- [10] Crow, Raym. 2004. *A guide to institutional repository software*. 3rd ed. Open Society Instituten Budapest. Available from http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf
- [11] Hattery, Maxine. 1999. The free digital library. *Information Retrieval and Library Automation* 34 (11):1-7.
- [12] Jørgensen, Ulrik, and Ole Sørensen. 2002. Arenas of development: a space populated by actor-worlds, artefacts, and surprises. In *Shaping technology, guiding policy: concepts, spaces, and tools*, edited by K. H. Sorenson and R. Williams. Elgar ,Cheltenham, UK. 197-222.
- [13] Krishnamurthy, Sandeep. 2002. Cave or community? An empirical examination of 100 mature open source projects. *First Monday* 7 (6). Available from http://firstmonday.org/issues/issue7_6/krishnamurthy/index.html.
- [14] Latour, Bruno. 1986. The powers of association. In *Power, action, and belief: a new sociology of knowledge?*, edited by J. Law. Routledge & Kegan Paul. London. 234-280.
- [15] Lynch, Clifford. 2003. Institutional repositories: essential infrastructure for scholarship in the digital age. *portal: Libraries and the Academy* 3 (2):327-336.
- [16] Morgan, Eric Lease. 2002. Open source software in libraries. In *Open source software for libraries*, 7-18. Chicago: LITA.
- [17] Open-ILS.org. 2005. *Open-ILS.org frequently asked questions* 2005 [cited 20 February 2005]. Available from <http://open-ils.org/faq.html>.
- [18] Raymond, Eric S. 2001. The cathedral and the bazaar. In *The cathedral and the bazaar: musings on Linux and open source by an accidental revolutionary*. O'Reilly, Sebastopol, CA, 19-63
- [19] Suber, Peter. 2005. *Timeline of the open access movement*, 8 February 2005 [cited 20 February 2005]. Available from <http://www.earlham.edu/~peters/fos/timeline.htm>.
- [20] Tennant, Roy. 2000. The role of open source software. *Library Journal* 125 (1):36.