AC 2009-2500: CHANGING LIBRARY VENDOR CONTRACTS: A CASE STUDY IN ACQUIRING E-BOOKS FROM AN ON-LINE BOOK VENDOR

Charlotte Erdmann, Purdue University

Charlotte Erdmann is Engineering Librarian, Coordinator of Collections and Associate Professor of Library Science at Purdue University's Siegesmund Engineering Library. She is interested in collection development, statistical analysis of library use data and intellectual property searching. Erdmann is also Purdue's Representative to the Patent and Trademark Depository Library Program and is active in the Engineering Libraries Division, American Society for Engineering Education and the Patent and Trademark Depository Library Association.

Changing Library Vendor Contracts: A Case Study in Acquiring E-Books from an Online Book Vendor

Abstract

Safari Tech Books Online offers several options for providing computer books to campus users. Purdue University used the vendor for 1 ½ years and discovered that the current titles option was not working as well as expected. The option included providing access to the current year and two previous years for all available publishers and was less labor-intensive for the library staff. A statistical analysis was done that examined the use patterns, including most heavily used books and publishers, as well as books with and without catalog records. This paper discusses the new option chosen and compares the statistical use patterns before and after the change. The university library ultimately changed its plan to more closely meet user needs by limiting the number of publishers, maintaining high use titles, and selecting specific titles as needed. Additionally, problems that were encountered before and after the change are discussed.

Introduction

Purdue University is a tier-one research university with approximately 40,000 students and supports undergraduate and graduate students in computer-related departments and schools in the Colleges of Science, Engineering, Technology, and Management. The Purdue University Libraries system is a decentralized system with 13 locations. Three libraries regularly purchase computer books. The addition of an e-book vendor had the potential to better serve clients and reduce staff selection time and maintenance issues.

Purdue University Libraries began a subscription in March 2005 after trial access of four products. This paper provides a literature review, background with a brief profile of the university library and vendor, description of first implementation, statement of problems encountered, and statistical data that influenced the request to change the e-book subscription, and implementation of those changes. During 2006, librarians expressed concerns and re-evaluation began in the fall of 2006.

This case study examines one large university library system's experience subscribing to Safari Tech Books Online, an e-book vendor. For detailed explanations of Safari's plans, the author encourages the reader to check the vendor section for details.

Literature Review

This paper does not explain features, interfaces, or functionality of the Safari Tech Online Books system. This has already been done in previous reviews that outline the e-book product. Turner¹ (2003) of the University of Southern Mississippi reviewed Safari when 1200 titles were available

and logins were necessary. He reviewed the product based on the slot system and simultaneous users. Turner was very complementary with the exception of login access.

Whang² (2006) of the University of Washington wrote an excellent more recent review which highlighted the computer and business plans, search interfaces, and how the system works. In addition to the three standard plans which are current titles file (previous two years plus current year), complete titles file, and slot system, Whang discussed the all O'Reilly Computer Plan and the Business Plan. IP authentication became available before she wrote her review. Turner and Whang discussed the statistical reports. Both universities had access to books through the slot plan. With the slot plan, Whang stated that one can add titles at any time and swap a title after thirty days. A UW librarian swapped titles quarterly.

No journal article could be found that discusses the complete plan or moving from current titles plan to the slot plan.

Stephens, Melgoza and Wan³ (2008) discussed currency, usage and book release policies in their study at the Texas A&M University Libraries. The Libraries began its "pick and choose" subscription to Safari in 2003 with two simultaneous users. Due to the time-consuming nature of selection, the Libraries moved to the current titles file (current year and two previous years) when Texas A&M renewed its contract. There were 1,258 titles when the authors did their study "to determine the overall currency of electronic books in the Safari database; to determine currency and release policies of individual publishers who contribute books to the Safari database; to compare usage patterns of Safari books to their print counterparts at Texas A&M University..."

The authors checked books on publisher web sites and on Safari and compared editions. They also surveyed publishers to determine the release dates for the print and online. The results showed that: O'Reilly and Pearson and its imprints (Addison-Wesley, Cisco, Peachpit, and Prentice-Hall) released online and print concurrently while No Starch released the online 90 days after print publication. Other publishers did not respond to the survey. Additional information acquired by Stephens, Melgoza and Wan showed that O'Reilly did not remove books while Pearson "let the market determine removal." Only .6% of the titles did not have the newest edition available.

During the two semesters surveyed, Stephens, Melgoza and Wan stated that e-book use varied from 1 to 1040 times per book with an average of 40 while print use of books with Library of Congress QA75-76 ranged from 1 to 44 times with an average of 1.9. Texas A&M clients also experienced a 40% turnaway rate. Otherwise, use should have been higher. 42 % of the Safari titles had no use while 58% of the print had no use. Texas A&M acquired some print copies through an approval plan since the Safari current titles plan dropped a year when a new year began. Some e-books are unique and do not have new editions. Adding print titles gives the collection some "historical depth." Stephens, Melgoza and Wan discuss moving from a slot plan

to the current titles plan. They mention the disadvantage of losing older content when the new year begins.

Since the study, Texas A&M has moved from two to four and now six simultaneous users. It is considering subscribing to the complete file.

Wallace⁴ (2006) wrote a journal article discussing the use of Safari for selected course reserves with an existing slot system at the University of North Texas, Denton. He also did a similar conference presentation⁵ at ASEE in 2005. At the beginning of project, the university library had three simultaneous users and 750 slots. To cover the anticipated use increase, the computer science department paid for an additional simultaneous user. To avoid access challenges, professors explained to students the “importance of logging out to assure colleagues’ access to the required readings.” Also students were instructed to report any problems to Wallace who monitored usage. Based on the Safari BackOffice report “Rejected Session Requests” and student feedback, the increased reserve use did not impact campus Safari usage.

Background

Purdue University Libraries’ Collections Team reviewed several electronic book vendors for computer science materials in 2002-2003. These included ebrary, Books24x7, NetLibrary, and Safari Tech Books Online. Budget constraints made the Libraries a late adopter of e-books. As users’ needs increased, action occurred in 2004.

In September, Jane Kinkus, Mathematical Sciences Librarian, took the lead and prepared a memorandum⁶ to the Engineering and Sciences Cluster which outlined the rationale for acquiring electronic books. The librarian emphasized that “computer programming skills are required by curricula that span all of the University’s academic areas” and prepared a synopsis of the advantages and disadvantages of print and electronic formats. This list is a combination of memorandum and the cluster librarians’ discussion:

Access to print format limits use to the number of copies owned. Materials are in heavy demand and unavailable. Books tend to be large and flimsy volumes which are easily damaged by frequent use. Users must browse to find specific instruction and pieces of code. Materials take a lot of shelf space and require cataloging and processing. They also become obsolete within two to three years requiring “constant collection development, including weeding.”

Electronic format is more “equitable” and permits more people to use the materials because licensing permits simultaneous, multi-user access. Searching “full-text keywords or strings of code” is possible which makes it easier find needed content. Collection development is also simpler. Licensing is based on IP address which is typical of most library databases. MARC catalog records are available from e-book vendors.

Kinkus also prepared a summary of features of e-book vendors. As a result, the Engineering and Sciences Cluster members endorsed the memorandum and recommended that it be sent to the Libraries' Administration requesting trial access to several vendors' products.

The administration approved trials to ebrary, books 24x7, and Safari Tech Books Online. As a result of the trials, users and librarians alike agreed that Safari had the most user friendly interface as well as available keyword and code fragment searching. The vendor also supplied content from publishers that produced high quality materials and used a web browser with no additional viewer software. The university libraries staff decided that Safari Tech Books was the best fit and very useful for all types of clients.

Vendor

Safari Tech Books Online made e-books available since 2002. It is owned by Proquest and specializes in computer and business e-books. In 2005, Safari provided access to books from 25 publishers. Major publishers included: O'Reilly, Addison-Wesley, Que, Sams, Riders, Prentice-Hall, Microsoft, and Cisco.

Safari offered three major options: complete file of all available e-books; current titles file which is two years plus current year, and the title/slot system based on the value of the book. A book is worth .5 to 3 slots.

The advantages to the complete and current titles files were that more books were available and librarian involvement is limited. Many libraries have chosen the current titles plan while others have chosen the slot plan. With the slot plan, a staff member at a subscriber organization chose each book that is added to Safari. The company provided an announcement service as new books are added to Safari.

Safari's pricing model was based on the plan (complete file, current titles file, or number of slots), number of simultaneous users, and size of university.

Safari offered considerable capabilities to track usage. The BackOffice reports were very helpful. It was a complicated system so it was necessary to pick and choose which reports worked best. Reports may be received in Microsoft Excel format.

University Libraries' Subscription

The Libraries' Administration chose the current titles file of all publishers for two years plus the current year and two simultaneous users. The decision to purchase the Safari current titles file made the largest number of books available to our clients and eliminated involvement in selection. As a result, the Mathematical Sciences Library and the Engineering Library discontinued the purchase of many print books on computer applications.

The new subscription was greeted with enthusiasm by library users and staff.

Kinkus took the lead in monitoring the Safari access and communicating with the Library Administration, Electronic Resources Licensing Librarian, and Safari.

Rejected Sessions

Initial high turnaway rates made obtaining access to Safari difficult. The rejected session data revealed issues. The representative reported a turnaway rate of 48% for two simultaneous users. According to Figure 1, one observed the high number of rejected sessions:

Figure 1: 2005 Safari Rejected Session Summary for Purdue University

Another factor that influenced the simultaneous user problem may have been timeouts. When a session was inactive, the Safari standard time to end a session was 20 minutes. A library staff member reported that he had seen a student leaving Safari open while doing his homework. This certainly made it impossible for another student to use Safari. The Libraries asked Safari to change the timeout to 15 minutes.

Complaints had been received that students could not get into Safari. In August 2005, Kinkus reported the user access problems to the Library Administration who obtained Safari statistics on turnaways. In early October, the Associate Dean of Libraries, Cheryl Kern-Simirenko found funds to add two more simultaneous users. After October 2005, the rejected session rates dropped significantly. (See figure 1).

Problems Observed

Kinkus observed that there were 2000 titles shortly after the subscription started. She noticed that titles dropped to 1800 within two months. It appeared that one cannot assume that the number of titles was stable and increased as new titles were added since some titles are

withdrawn. The only explanation for changes in numbers of books available was the withdrawal of titles.

A client reported that a Safari title with a record in the library catalog had a dead web link. Searching Safari directly revealed that the title was no longer available on Safari and had probably been withdrawn.

With these issues, it was determined that it is good practice to completely reload the MARC catalog records on a regular basis and avoid inconveniencing users.

Current Plan Limitations

The current titles collection had an inherent problem by the volume of changing titles at the end of the year. Safari changed titles at the end of calendar year based on the publication year. For example, a library subscribed to Safari in July 2005 and had access to books from 2003, 2004, and 2005. In January 2006, all books from 2003 were no longer accessible. The collection became 2004 and 2005 plus all books added in 2006. This put the responsibility on the library to do a complete reload of the MARC records as soon as possible after the new year began.

Additionally, some books are published in one year and not added until the next year. For example, ABC title was published in November 2003 and was not inserted (also known as “added”) until June 2004. ABC title was removed January 1, 2006. It made no difference whether a book was published in January 2003 or December 2003, it was removed January 1, 2006. The Safari representative confirmed these details.

In January 2006, a faculty member emailed the Libraries regarding books that had been removed from Safari. He required his students to use several books for a course that had been removed from the Libraries’ plan in January. The library administrator talked to the faculty member and contacted Safari staff but was unable to resolve the issue. With the current titles plan, there is no arrangement to select inaccessible books because they do not fall in the date range covered by the plan. The Libraries’ alternatives were to subscribe to the complete collection, find another e-book vendor for the specific titles or purchase paper copies of the books. For Purdue University, the cost of subscribing to the complete collection was prohibitive. No alternate e-book vendor could be found. As a result, Libraries bought print copies of the books and put some books on reserve.

MARC Record Delays

Purdue staff observed that MARC catalog records had been delayed since the beginning of the Libraries’ Safari plan. At any given time, at least ten to twenty percent of the books did not have available records. Information Technology staff sometimes loaded new MARC records and found few titles available. When IT staff members reloaded from scratch, they found more

records available. After a new books was added, it was supposed to take 30-45 days for new MARC record to be issued by Safari.

The author sent a list of titles with publication date and insert date to Safari staff with hopes that the situation would improve.

Statistical Analysis of 2004-2006

In early 2007, the author of this paper completed a statistical analysis of 2004-2006 titles to identify use patterns, publishers, most heavily used books, and number of MARC records available. Erdmann used Safari BackOffice reports, Microsoft Excel, and Microsoft Access. Purdue librarians wished to answer the following questions:

1. "Is it possible that people are using books from some publishers more than others?"

Using Microsoft Access, the author prepared a query grouping publishers and counting titles to determine the number of titles available from each publisher and then exported data to Microsoft Excel for charting. The same process was done with the Sections Read.

Figure 2 shows the publishers and percentage of titles in Safari's current titles plan.

Figure 2: Current Titles Plan by Publisher and Percentage of Titles for 2004- 2006

The author also grouped titles by publisher and sections read. Figure 3 shows publishers and percentage of sections read during the same time. The publisher list is collapsed.

Figure 3: Safari Sections Read by Publisher and Percentage for Purdue University Users

The largest amount of use was of O'Reilly, Addison-Wesley, Sams, and Prentice-Hall title. Compared with the number of titles on the previous table, one expected higher usage of Que and Peachpit Press titles.

Figure 4 shows the same data by number of sections read, giving the reader another view. The publishers are collapsed. Publishers may not show if the titles from the publisher had little use.

Figure 4: Safari Sections Read by Publisher and Number for Purdue University Users

Based on Figures 3 and 4, the evidence showed that Purdue users made heavy use of four publishers: O'Reilly, Addison-Wesley, Sams, and Prentice-Hall. The author ran several Access queries which showed publishers that had low numbers of section read or no sections read. For the purposes of this paper, this data is not included. Other publishers had several books on certification that were popular. This was noted for later consideration.

2. "With the two year plus current titles plan, how many heavily used titles are lost when the calendar year changes?"

This question was important since librarians were concerned about the titles that were dropped due to the faculty member's problem with missing books in early 2006. Figure 5 provided a sample of highly used titles that were lost in January 2007. No newer editions were found for these books. There were 85 titles.

Title	Imprint	Publication Year	Sections Read
PHP and MySQL Web Development, Third Edition	Sams	2004	1334
Security Warrior	O'Reilly	2004	1178
Sams Teach Yourself C++ in 21 Days, Fifth Edition	Sams	2004	1133
The LATEX Companion: Second Edition	Addison Wesley Professional	2004	1047
Build Your Own ASP.NET Website Using C# & VB.NET	SitePoint	2004	771
XML in a Nutshell, 3rd Edition	O'Reilly	2004	764
Introduction to Computer Security	Addison Wesley Professional	2004	637
Core Java(TM) 2 Volume I - Fundamentals, Seventh Edition	Prentice Hall	2004	616
Visual Basic .NET Programming Language, The	Addison Wesley Professional	2004	604
Internet Denial of Service: Attack and Defense Mechanisms	Prentice Hall	2004	603
VBA and Macros for Microsoft Excel	Que	2004	578
Linux Cookbook	O'Reilly	2004	564
COCOA PROGRAMMING FOR MAC OS X SECOND EDITION	Addison Wesley Professional	2004	536
Cisco ASA and PIX Firewall Handbook	Cisco Press	2005	505
Learning Python, 2nd Edition	O'Reilly	2003	478
Sams Teach Yourself Java(TM) 2 in 21 Days, Fourth Edition	Sams	2004	461
Oracle Essentials: Oracle Database 10g, 3rd Edition	O'Reilly	2004	454
Core Java(TM) 2 Volume II - Advanced Features, Seventh Edition	Prentice Hall	2004	451
Unified Modeling Language Reference Manual, The, Second Edition	Addison Wesley Professional	2004	439

Figure 5: Sample of Highly Used Books from 2003-2005 Dropped in 2006

These books merited continued access but due to the restrictions of the current titles plan this was not possible. The librarians discussed purchasing print copies of these titles but access from users' desktops was preferable to the library check-outs.

3. Are MARC records being loaded on a regular basis? Is this phenomenon affecting usage?

The availability of MARC records was a continuing problem with Safari and the Libraries' Information Technology department. When the statistical data was analyzed in early 2007 for 2004-2006 data, there was a marked use of the older materials rather more than newer materials. Typically, one expected the newest materials to be used most frequently but this was not true.

Pub_Date Year	AvgOfSections Read
2004	65
2005	71
2006	47

Figure 6: Safari Average Use of Titles by Year

At the same time, the number of records in the Library Catalog was low for 2006 materials. This may have influenced lack of use of newer materials. Work demands may have prevented regular loading in 2006. The IT Department did regular reloads in 2007 and the number and accuracy of the records was better. The author found that 10-20% of titles still did not have records. In conversation with Safari staff over the years, MARC records were supposed to be ready within 30-45 days of addition to Safari.

Rethinking the Current Plan

Safari offered excellent content from reputable publishers. There was never any question about that fact. Given the problems with the current titles plan, the librarians have been rethinking the options for Safari access for some time. The statistical analysis showed that users were not using many of the books from certain publishers.

The desire to change the plan built slowly over time. Challenges had been recorded via email messages and consultation with colleagues in other libraries. In August 2006, the author did more analysis and started formalizing a proposal. The librarians in the Physical Sciences, Engineering, and Technology Division (previously known as the Sciences and Engineering Division) reached consensus and decided that offering fewer titles served our users well and may keep the simultaneous user problems to a minimum. A new option also helped solve the problem of the removal of frequently used materials at the beginning of each year.

The one hesitation was that the Libraries may no longer automatically get the new titles. It was necessary for a librarian to select new books. Someone would have to deselect some titles to get more titles. Ideally, the swap could be done once a year to make de-selection less time-consuming.

Selection librarians agreed to add new materials from four publishers. On request, selectors were encouraged to add new titles as long as the books were available from Safari.

New proposal

In February 2007, the author proposed a new plan that more closely met user needs by limiting the number of publishers, maintaining high use titles, and selecting specific titles as needed. The Physical Sciences, Engineering, and Technology Division endorsed the plan and sent it to the

Libraries' Administration. The Libraries' new proposal was a hybrid of Safari's plans. In June 2007, the librarians learned that Safari was not willing to accept the hybrid model so the only option available was the slot system. In August 2007, Safari and Purdue implemented the 1080 slot plan. When all slots were used, this was an estimated 875 books.

Based on statistics, the librarians chose four publishers, i.e. O'Reilly, Addison-Wesley, Sams, and Prentice-Hall as well as 100 most heavily used books from all publishers in 2002-2004. They also chose not to acquire certification titles unless the books came from the four publishers.

Implementation

Library staff members found the August 2007 implementation of the new plan to be time consuming but it ran fairly smoothly. Safari staff loaded the 100 titles and books from the four publishers for 2005-2007. The author selected most new books from the four publishers. At the end of 2007, she did the end of the year use analysis. Most of the titles used were from 2005-2007. The whole list was examined for new editions and obsolete versions of software and 2005 titles were examined more closely. Most of the titles dropped were from 2005 and earlier. The author did the de-selection in January 2008.

When adding new titles in January 2008, the author discovered that over 100 titles from 2007 had not been part of the reload in August 2007. In hindsight, it may have been better to verify all 2007 titles. Because of the additional 100 titles, the author found slots in short supply in 2008. A proposal to add more slots could not be funded due to budget constraints. The author de-selected titles in October 2008 so that new selections could continue until year's end.

2008-2009 Year In Review

There was good news in 2008. The author honored two professors' requests to add twenty older books from O'Reilly to meet course needs.

The addition of MARC records appeared to be improving. The author met with the Safari representative this year. She voiced the customers' concerns to Safari staff members. In January 2009, the author is pleased to report that there were only 11 records that did not have MARC records. There had been 125+ after the December 2008 reload.

Rejected sessions have also decreased. Figure 7 shows 2006-2008 rejected sessions. Rejected sessions have remained less than fifty per month in 2008. In comparison with previous years, the author feels that the rate of rejected sessions is acceptable

Figure 8: 2006-2008 Safari Rejected Session Summary for Purdue University

In January 2009, the author did the January selection and one slot remains available. She also did the 2008 data analysis, de-selected the lesser used titles from 2006 and earlier. The analysis showed that 60 high-use books that were part of the original 100 titles are in Purdue's Safari plan. In previous years, the author looked at cumulative sections read. In this analysis, she took a year by year view of sections read for each title. In some cases, use has dropped. In other cases, use is up in 2006, down in 2007, and up in 2008.

Lessons Learned

The author has gathered some lessons from this experience. This list is not complete but it contains some thoughts that may help libraries with future vendor discussions.

- Get support from Library Administration. Go through appropriate channels.
- Become knowledgeable about products and discuss products with peer librarians.
- Acquire expertise in data collection. It does take time. Use vendor reports that work well. Document procedures so that it is possible to delegate.
- Approach vendors with customer needs. Teach vendors that happy customers are returning customers.
- Balance content with the number of simultaneous users.
- Ask for more simultaneous users when it is needed.
- Seek cooperation with vendors and librarians. Hold vendors responsible.

- Be persistent when negotiating with vendors. It may take time for a vendor to act.
- Verify that the library is getting its customers' needs met. Sometimes a "good deal" is not the "best deal."
- Communicate changes to customers.

These lessons may not apply in every situation but are certainly applicable to this situation.

Conclusion

This conference paper has described one university libraries' implementation of Safari and its move from the current titles plan (two years plus current) to a slot plan. Safari is a quality product that fits the university needs and supplies quality information from reputable publishers. The slot system with specific selection criteria gives many options that the current titles option did not. The implementation took time and has been a successful experience. Some work on behalf of the librarians went a long way toward increasing user satisfaction. While the previous plan required little work from the library staff, the patrons were not served as well. The solution has also addressed typical problems faced by all Safari customers, e.g. turnaway rates, simultaneous users, and MARC records. The slot plan solved several problems faced by current titles plan customers.

Special Thanks

The author would like to thank Jane Kinkus, Mathematical Sciences Librarian who worked with the original implementation; Emily Mobley, Emeritus Dean of Libraries and Cheryl Kern-Simerenko, former Associate Dean of Libraries who originally funded Safari; Patricia Kantner, Head of Information Technology and Resource Services Operations and Bill Corya, Electronic Resources Licensing Librarian who blessed the new proposal; and Dean Lingley, Senior Administrator, Library Enterprise Applications and Sue Long, Database Maintenance who are responsible for loading the MARC records. Employees from the Engineering Library have also helped with year-end analysis. The author also thanks the Safari representative and staff who worked with Purdue University Libraries during our subscription to Safari.

¹Turner, Steven. "Safari Tech Books Online: A Valuable Resource." *Mississippi Libraries* vol.67, no.2 (Summer 2003): 48-50.

²Whang, Linda. "Electronic Resources Reviews: Safari Tech Books Online." *Issues in Science and Technology Librarianship* no. 47: (Summer 2006): 8 pages. <http://www.istl.org/06-summer/electronic1.html> [Accessed 2/7/2009]

³Stephens, Jane, Pauline Melgoza and Gang (Gary) Wan. "Safari Books Online: Currency, Usage and Book Release Policies of an E-Book Database." *Collection Building* vol. 27, no.1 (2008) : 14-17

⁴Wallace, Randy. "Safari Tech Books Online as Supplementary Reserve Materials." *Science & Technology Libraries*, vol. 26, no. 3/4 (2006): 7-11.

⁵Wallace, Randy. "Utilizing Safari Tech Books Online as Supplementary Reserve Materials." Presented at *Engineering Libraries Division, ASEE Annual Conference*, Portland, Oregon, June 15, 2005

⁶Kinkus, Jane. *Memorandum to Sciences and Engineering Cluster Re: Electronic books. Purdue University Libraries, Sept. 8, 2004.*