Collection Development—
Demand-Driven Acquisitions
Boom or Bust: Short-Term Loans Five Years Later

Sherri Brown, Literature and Humanities Librarian, University of Kansas
Lea Currie, Librarian/Head of Content Development, University of Kansas
Andi Back, Fine Arts and Humanities Librarian, University of Kansas

Abstract

The University of Kansas (KU) Libraries began a demand-driven acquisition program for e-books in late 2011, which included short-term loans (STLs). At that time, STLs cost 5 to 10% of a book’s list price, with libraries paying no more than 130% when actually purchasing an e-book. The literature from the early years praises the new purchasing model as cost effective, often saving libraries thousands of dollars annually. However, in 2014, the cost of STLs began to increase, with a handful of publishers charging as high as 30 to 35% per loan. In FY14, the KU Libraries saw a sudden 122% increase in the cost of their STLs and a 277% increase in the overall cost of their e-book purchases. Alarmed by this sudden increase, KU librarians began looking for solutions to save money, including lowering the number of loans from three to two before triggering a purchase. Unfortunately, STL costs at KU continued to rise the following year by 32%. STLs were no longer working as a cost-saving measure, and publisher pricing for STLs has continued to rise, some to as much as 40 to 50% of the cost of a book.

When the concept of demand-driven acquisitions (DDA) for e-books was first introduced, many libraries embraced this model as a cost-saving method for collecting monographs. Publishers agreed to try out pilot projects with certain libraries to see if this new model was profitable. E-book vendors such as EBL, Ebrary, and EBSCO offered e-book purchasing with short-term loans (STLs) costing 5 to 10% of the list price of the book, with libraries paying no more than 130% when actually purchasing the e-book. The library literature from these early years of STLs praises the new purchasing model as cost effective, often touted as saving libraries thousands of dollars annually. As opposed to the long-time standard of approval plans that allowed libraries to purchase books on a “just-in-case” basis with reduced budgets, libraries found the “just-in-time” approach of the DDA model to be highly advantageous. Even more desirable was a hybrid approach, essentially a demand-driven-preferred approval plan, which enabled libraries to provide access to more content than ever before, while spending less.

The University of Kansas (KU) Libraries began a DDA program for e-books (eDDA) in late 2011 after successfully implementing a DDA program for print books (pDDA) during the previous two years. Some librarians at KU had expressed skepticism about investing in e-books, particularly in the humanities, but were swayed into piloting an e-preferred acquisitions model by writing criteria for purchasing e-books into the Yankee Book Peddler (YBP) approval plan for the sciences and social sciences. E-books had to fit into the approval plan parameters like their print counterparts and had price limits. For the social sciences, the price was capped at $94.99 to come in on approval, and books priced $95 to $150 were loaded into the online catalog as DDA. Any book over $150 had to be selected by librarians. For the sciences, there was a $150 price cap for books on approval, and books costing $151 to $200 were loaded into the online catalog as DDA. Any book over $200 had to be selected by a librarian. By consulting retrospective lists of books that would have come in on approval, we were able to project expenditures and average prices. We estimated that in the first year, 23% of our monograph purchases would be eDDA, 50% would be pDDA, 20% would be received on approval, and 7% would be selected by librarians. These projections were fairly accurate.

When the e-book plan was created, KU Libraries set their preference for e-books to come in primarily through EBL’s nonlinear lending model, which allowed for 365 uses of a book before being bound to purchasing a second copy. If nonlinear lending was not available, single-user access was next preferred, followed by multiple user, if requested for use in the classroom. Librarians also had a week to review the approval e-books before they came in automatically.

Copyright of this contribution remains in the name of the author(s).
https://doi.org/10.5703/1288284316454
Table 1. Ten years of approval plan numbers.

<table>
<thead>
<tr>
<th>FY</th>
<th>Approvals received: pBooks</th>
<th>Approvals received: eBooks</th>
<th>pDDA purchases</th>
<th>eDDA purchases</th>
<th>pDDA records loaded</th>
<th>eDDA records loaded</th>
<th>Total titles purchased</th>
<th>Total expenditures, including eDDA short-term loans</th>
<th>Cumulative percent change</th>
<th>Avg price</th>
<th>Change in titles purchased</th>
<th>Change in expenditures</th>
</tr>
</thead>
<tbody>
<tr>
<td>FY06</td>
<td>15,250</td>
<td>15,250</td>
<td>823,857.50</td>
<td>54.02</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>FY07</td>
<td>14,551</td>
<td>14,551</td>
<td>828,450.29</td>
<td>56.85</td>
<td>95%</td>
<td>101%</td>
<td>95%</td>
<td>101%</td>
<td>95%</td>
<td>101%</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>FY08</td>
<td>14,328</td>
<td>14,328</td>
<td>760,554.18</td>
<td>53.09</td>
<td>94%</td>
<td>92%</td>
<td>94%</td>
<td>92%</td>
<td>94%</td>
<td>92%</td>
<td>94%</td>
<td></td>
</tr>
<tr>
<td>FY09</td>
<td>12,223</td>
<td></td>
<td>614,149.36</td>
<td>50.25</td>
<td>80%</td>
<td>75%</td>
<td>80%</td>
<td>75%</td>
<td>80%</td>
<td>75%</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td>FY10</td>
<td>11,497</td>
<td>37</td>
<td>573,612.85</td>
<td>49.89</td>
<td>78%</td>
<td>70%</td>
<td>78%</td>
<td>70%</td>
<td>78%</td>
<td>70%</td>
<td>78%</td>
<td></td>
</tr>
<tr>
<td>FY11</td>
<td>11,156</td>
<td>199</td>
<td>545,972.55</td>
<td>49.33</td>
<td>75%</td>
<td>66%</td>
<td>75%</td>
<td>66%</td>
<td>75%</td>
<td>66%</td>
<td>75%</td>
<td></td>
</tr>
<tr>
<td>FY12</td>
<td>8,629</td>
<td>925</td>
<td>488,058.48</td>
<td>50.71</td>
<td>64%</td>
<td>59%</td>
<td>64%</td>
<td>59%</td>
<td>64%</td>
<td>59%</td>
<td>64%</td>
<td></td>
</tr>
<tr>
<td>FY13</td>
<td>7,869</td>
<td>1,608</td>
<td>512,344.57</td>
<td>52.85</td>
<td>64%</td>
<td>62%</td>
<td>64%</td>
<td>62%</td>
<td>64%</td>
<td>62%</td>
<td>64%</td>
<td></td>
</tr>
<tr>
<td>FY14</td>
<td>7,788</td>
<td>1,571</td>
<td>552,205.29</td>
<td>57.93</td>
<td>64%</td>
<td>68%</td>
<td>64%</td>
<td>68%</td>
<td>64%</td>
<td>68%</td>
<td>64%</td>
<td></td>
</tr>
<tr>
<td>FY15</td>
<td>7,767</td>
<td>89</td>
<td>468,177.77</td>
<td>54.88</td>
<td>56%</td>
<td>57%</td>
<td>56%</td>
<td>57%</td>
<td>56%</td>
<td>57%</td>
<td>56%</td>
<td></td>
</tr>
<tr>
<td>FY16</td>
<td>7,479</td>
<td>64</td>
<td>460,634.97</td>
<td>54.33</td>
<td>56%</td>
<td>56%</td>
<td>56%</td>
<td>56%</td>
<td>56%</td>
<td>56%</td>
<td>56%</td>
<td></td>
</tr>
</tbody>
</table>

In addition to the approval e-books, YBP also sent records for eDDA to be loaded into the online catalog without purchasing. Initially, EBL was our only provider of eDDA. We later added ebrary, EBSCO, and individual publishers to the list of e-book providers. EBL allowed browsing of an e-book for 5 minutes before a 1-day STL was triggered. An STL was also triggered by printing or by downloading a book to use offline in Adobe Digital Editions. Initially, KU Libraries chose to set our eDDA plan to have three STLs with autopurchase on the fourth use. During the first year of eDDA, KU Libraries went from approximately $65,000 in print DDA savings in FY09 to more than $118,000 in savings for the entire DDA program in FY11. While expenditures for books declined significantly for the first few years, the number of books available in the online catalog for discovery increased a great deal.

In the initial years, the eDDA and pDDA programs were deemed a success, and the number of records loaded into the catalog continued to increase. All was well until 2014, when publishers began to raise the cost of STLs, with a handful going as high as 30 to 35% of the cost of the book per loan. In FY14, the amount spent on STLs by KU Libraries skyrocketed to a 122% increase over the previous year, with a 277% increase in the overall cost of e-books. Since KU had experienced a flat collections budget since 2009, there were no funds to cover these increases, and money had to be reallocated to pay these additional costs. To try and staunch the outward flow of funds for STLs, KU lowered the number of STLs from three to two before a purchase was triggered, which cut costs significantly. In 2015, KU also stopped getting e-books automatically on approval by rejecting them and setting them all to DDA. This was also a significant cost-saving measure, but it is proving to be somewhat short-lived as more and more publishers have stopped allowing DDA for their titles altogether. Currently, when a publisher does not allow eDDA, we load the record for the print book into the online catalog where patrons can request that we order the book for them.

Looking for other ways to make eDDA more cost effective, we recently investigated how changing the number of STLs from two to one would affect spending. We also considered how changing our preference from a nonlinear lending preferred format to a single-user format might lower our e-book costs. EBL provided a year’s worth of data (February 1, 2015–February 5, 2016, see summary in Table 2). During this time period, we spent $4,408.53 to purchase 41 titles (37 nonlinear lending, two unlimited users, and two single-user) for an average cost of $107.53 per book. If we had purchased these 41 titles in single-user-only format, we found we would have saved $909.82. If we changed the autopurchase to trigger on the second use, we would have spent $3,809 more. Ebrary provided data from April 28, 2015, through May 3, 2016 (see summary in Table 2). We spent $2,615.82 to purchase 23 of the 60 titles triggered at least twice (six single-user, 10 unlimited users, and seven three-user). If we changed the autopurchase to single-user, we would have saved $585.18. If we changed the autopurchase to trigger on the second use, we would have spent $1,949.91 more. We had far less data to review from EBSCO. They sent us data from March 1, 2015, to February 26, 2016, which only amounted to four titles (see summary in Table 2). We spent $228.98 on STLs during that time, with no autopurchases. If we lowered the triggers before autopurchase, we would
have spent $367.46 more. It became quite apparent that lowering the number of triggers before a book is purchased would cost us more, while changing our purchasing preference to single-user would save us a significant amount of money.

Another cost-saving measure was offered to us by YBP to help cut more of our STL costs. Entitled Apex, this new program is an option in our approval plan to have titles that would most likely trigger an eDDA purchase come in automatically using an algorithm from YBP. Purchases only amount to one or two titles a week, but since we started the program in July of this year, we have potentially saved almost $250 in STLs during the first three months. We will continue to watch the statistics for the next year or so to see if this program is really saving us money in the long run. Finally, to further find STL savings, we have identified several publishers in our approval plan that charge 35% or more in STL costs to turn off. In most cases, we choose to load records for these books in print and forego the cost of the e-books and STLs.

KU is currently collaborating with ProQuest and YBP to negotiate a collaborative eDDA program for KU and Kansas State University (KSU) where a purchase by either campus will provide access to the e-book at both universities. ProQuest is contacting and negotiating with a list of publishers that KU and KSU developed to create a collaborative collection development project. While several publishers on our list have not agreed to participate in this project, ProQuest has had some success in negotiating with a handful of publishers to start an eDDA program. KU and KSU may not see cost savings with this project, but we have long wanted to collaborate on a collection development project, and this option seems like an opportunity to better share our collections.

It remains to be seen if there truly is a future in demand-driven acquisition, with so many publishers pulling out of providing this option. Whereas libraries continue to see the benefits of the eDDA model, publishers, who were accustomed to making money off their front list, are having a hard time reconciling the idea that with this new DDA model they are more likely to make money off their back list. As we have seen at KU, the option of providing more records in our online catalog for discovery has drawn the support of our researchers. They have embraced this option as the overall expenditures for books has receded, and we are hopeful that many publishers will see the opportunity to make profits off back lists as their profits from front lists are likely to diminish as well.

Table 2. STL and autopurchase expenditures from three major eDDA vendors.

<table>
<thead>
<tr>
<th>Vendor</th>
<th># of titles with 2+ STLs</th>
<th>Amt. spent on STLs &amp; autopurchases</th>
<th># of STLs</th>
<th>% spent on STLs</th>
<th>Avg. cost per STL</th>
<th># of autopurchases</th>
<th>% spent on autopurchases</th>
<th>Avg. cost per autopurchase</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBL</td>
<td>91</td>
<td>$7,903.07</td>
<td>171</td>
<td>44</td>
<td>$20.44</td>
<td>41</td>
<td>56</td>
<td>$107.53</td>
</tr>
<tr>
<td>ebrary</td>
<td>60</td>
<td>$4,834.24</td>
<td>114</td>
<td>46</td>
<td>$19.46</td>
<td>23</td>
<td>54</td>
<td>$113.73</td>
</tr>
<tr>
<td>EBSCO</td>
<td>4</td>
<td>$228.98</td>
<td>8</td>
<td>100</td>
<td>$28.62</td>
<td>0</td>
<td>0</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Rolling with PDA and DDA: How Academic Libraries Can Use Patron-Driven and Demand-Driven Acquisition Techniques to Build Library Collections With Minimal Management and Budget

Kerry A. Falloon, Head of Acquisitions, College of Staten Island Library–CUNY

Abstract

Patron-driven and demand-driven acquisitions (PDA/DDA) have been utilized for some time in academic libraries, but some university libraries are still new to the process. With changes in the last couple of years regarding short-term monograph loans becoming cost ineffective, the popularity of streaming video PDA, and library materials budgets increasingly being cut, the conundrum of successfully implementing a DDA/PDA program and how to evaluate its effectiveness is a question many libraries need to answer. In the fall of 2015, the College of Staten Island Library-CUNY, implemented a small DDA monograph pilot program with Yankee Book Peddler (YBP) and EBSCOHost e-books to build its collections in academic areas new at the college. The goal was to offer patrons immediate on-demand selections in areas that might need additional collection building in depth and breadth but will take some time to accomplish by collection development (CD) librarians. The implementation and the results of this DDA pilot project, as well as in the fall of 2016, a new streaming video PDA program with Kanopy Streaming, will be discussed.

Introduction

This shotgun session with multiple presentations was well attended by librarians and vendors alike, on the overall topic of collection development/demand-driven acquisition and e-book threads. This segment of the presentation, “Rolling with PDA and DDA . . . ,” is based on pilot projects that were completed at the College of Staten Island (CSI) Library with a total budget of $5,000 allocated per year toward each proposal. With budget cuts of over $100,000 in the year the implementation of the first project began, this paper discusses small-budget and manageable PDA and DDA programs that can be supportive of collection building in academic libraries with a minimal budget and plan management needed. These projects were in direct contrast to a larger scale CUNY DDA plan, which the College of Staten Island (CSI) benefitted from in 2014. One of the 11 senior colleges in City University of New York (CUNY) system, the CSI Library serves full-time equivalent (FTE) of over 11,500 students from associate’s to doctoral level academic programs, making its curricula comprehensive in scope. The general size of its collections includes roughly 475,000 monographs, 15,500 DVDs and videos, 28,850 streaming video titles, and 224,00 e-books, which includes CUNY e-book purchases. The CSI library had an $800,000 materials budget from multiple funding sources in recent years. The combination of these factors makes collection development and maintenance challenging yet rewarding in this library setting.

There are many competing philosophies and ideologies regarding collection development (CD) in academic libraries. Academic libraries were viewed as warehouses of owned information and academic librarians as key players in CD practices. However, DDA and PDA can transform libraries into providers of access to immediate information. Using these tools in conjunction with traditional CD practices, libraries can still be considered the institutional upholders or bastions of knowledge while providing patrons with access to immediate informational needs. The concept of demand-driven acquisitions (DDA) or patron-driven acquisitions (PDA) was initially a challenge for CSI librarians who were used to a traditional “just-in-case” purchasing practices, with a set amount of funds spent yearly. They later understood they could be a part of the process at the title or profile level, seeing themselves as gatekeepers of this demand-driven information, as selected by patrons based on their “just-in-time” needs. These concerns were recognized during the initial e-book DDA project, with CD librarians vetting each subject area title by title. The library’s CD committee also examined each of Kanopy’s subject and publisher area collections for inclusion in the PDA program. For the purpose of this paper, although DDA and PDA are synonymous with each other, the author is using both in deference to the terminology the vendors use.

Copyright of this contribution remains in the name of the author(s). https://doi.org/10.5703/1288284316455
Background

Most academic libraries have experienced flat to severe budget cuts in the last decade. In fiscal year 2015, the CSI library had a decent size budget with $111,758 spent on print books, $34,640 on e-books, $25,334 on streaming video platforms, and $3,171 on DVD’s. The CSI library also benefits from joint e-book collections shared across CUNY, including ebrary’s Academic Complete. Our print budget still heavily outweighed digital monograph purchases, although physical DVD’s, due to low circulation, is considered an underutilized and fading collection. In July 2015, CSI library received $126,000 in budget cuts at the same time we were developing EBSCO’s DDA pilot project. In fiscal year 2016, CSI library spent $50,000 on print books, $17,430 on e-books, $27,744 on streaming videos, and $3,983 on DVD’s. For the first time, e-book purchases represented roughly half of all monograph purchases chosen.

The College of Staten Island’s library patrons did participate in a CUNY-wide DDA, which ran from November 2014 to March 2015, in which $75,000 in funds dedicated to the project were exhausted, with 363 titles chosen across all CUNY campuses (Egan, Yearwood, & Kendrick, 2016). COUTTS’ MyiLibrary platform, CUNY’s main book vendor, was chosen, and the allocation formula was based on the total value of titles (multiplied by a consortia factor of three) should be equal to 15 times the amount dedicated to the project (Egan et.al., 2016). Almost 5,000 titles were selected, and the CSI library’s patrons accounted for around 4% of all triggers (Egan et.al., 2016). Since titles were chosen across all subject areas, the reasoning behind the lower number of selections CSI had versus other CUNY campuses is surmised as the continued reticence of CSI patrons to choose an e-book over a print book, as seen in librarian and patron interactions. Another potential reason was not fully utilizing CUNY’s new discovery tool launched a month earlier, which other CUNY campuses were doing at the time.

Implementation of DDA/PDA

The original plan with the DDA project with YBP and EBSCOhost e-books was for $10,000 to be spent across fiscal year 2015. With the recent budget cuts, the $10,000 placed into a deposit account with YBP would be split into a two-year period, or $5,000 per year toward anticipated expenditures. Although YBP’s collection development manager suggested that thousands of titles be loaded, the CSI library wanted to stretch as much as possible its deposit account funds and to have a sustainable DDA project across a timespan of two years, since no additional funding was expected the following year. The CSI library decided there was no magic allocation formula, just trial and error, generating a baseline the first year and readjusting the second year based on certain factors. Factors that affected expenditures were primarily considered to be supply versus demand or number of titles loaded; capping the cost of titles loaded to between $150 to $250 based on subject area; the quality of the titles loaded based on publisher and recent publications (less than five years old); and patron needs (e.g., new academic programs with collections of less depth and breadth). Factor considerations were partially based on CUNY’s experience with their shortened DDA program. The CSI library’s main criteria for success was a sustainable DDA project over the time specified, with the right amount of funds allocated and titles loaded.

The CSI library decided to choose EBSCOhost e-books due to patron familiarity of EBSCOhost as a platform and the superior quality of their PDFs with optical character recognition (OCR), which are easily accessible to patrons with disabilities. A three-user license model was also chosen. Folders were created in YBP’s GOBI in areas identified as East Asian studies, ESL/ TESOL, legal studies, Arabic studies, social work, physical therapy, data analytics, and nursing studies, so librarians could review title by title. In all, 421 full machine-readable cataloging (MARC) records were batch-loaded into CUNY’s catalog, although 500 titles were chosen to supplement traditional CD practices for new academic degrees at the college. Titles were pulled after publishers determined a title would no longer be available via DDA. After the catalog load, an average of three titles per month were pulled, creating ongoing catalog maintenance. Further catalog maintenance would also include changing the MARC status field of LOCAL, which was done so DDA records would not be submitted to Online Computer Library Center (OCLC) every four months, to VSI after a purchase.

Due to the poor usage of CSI’s DVD and perpetually owned streaming video purchases, the library’s next goal was a PDA project with Kanopy streaming video. From 2015 to 2016, three films were purchased, totaling $435 for only three plays during the year,
although the titles were recommended by faculty. These would have free under Kanopy’s trigger of four plays needed within a 12-month period. Seven other CUNY libraries already had a PDA with Kanopy, although CSI was the first CUNY to license with Kanopy several years early for their media education foundation publisher package, which does not participate in PDA. Baruch College’s library, with a budget of $5,000 for the entire use of the 26,000 title Kanopy PDA collection, came closest to what CSI library wanted to achieve, although few films were triggered since they were not load PDA records into the CUNY catalog. In early October 2016, CSI decided to load 11,564 PDA records across 30 subject packages and one publisher package. Kanopy provided de-duplication across all packages. The CSI library committed to subsequent loads of new videos added into the collections every six months at minimum and to possibly consider adding new collections. At the Charleston Conference 2014, librarians from UMass Amherst and Simmons the reported a significant increase in their return on investment (ROI) by over 95% by comparing the cost per play using Kanopy PDA, purchased streaming collections, and the cost per lifetime checkouts of their DVD collections (Ferguson & Stangroom, 2014). The CSI library is hoping to see the same return on investment as this pilot project continues to generate meaningful data for analysis.

DDA/PDA Project Analysis

Between September 2015 and September 2016, 21 EBSCOhost e-book titles were triggered or about 5% of all titles loaded, at a cost of $2,487 to the library. This was half of the projected expenditures for the year. The average cost per e-book selected was $118. The total value of all 421 titles loaded was originally $36,992. The maximum cost per title was $150 with a four-year retrospective publication time period. The first objective was to load more titles to see more activity. In September 2016, another 790 titles were loaded in new degree areas. We also changed to a one-user license model, for more title choice and lower cost. Total value of the over 1,200 titles is now $126,168, with six titles triggered in a month’s time. The maximum cost per title was also increased to $250 maximum cost per title, especially in nursing and physical therapy titles. The CSI library did not use YBP for collection development during year two but utilized EBSCOhost’s Collection Manager to choose subject-sets of areas identified. In an additional analysis, Taylor & Francis Group was the predominate publisher chosen, next Wiley, and Harvard University Press, and 76% of the triggers were through page turns of 10 pages or more. Eight titles were looked at but not triggered. Half of the titles selected were ESL/ TESOL titles, and the rest were East Asian and Arabic studies, amounting to 21 titles triggered the first year. If the subject areas had less than a 3% divergence between the percentage of records loaded and the percentage of records triggered by patrons, this would indicate that the amount of title records loaded met the needs of the patrons in a satisfactory manner. From this analysis, legal studies and Arabic studies were identified as needing less of the overall percentage of titles loaded, while East Asian studies and ESL/TESOL studies as potentially needing a greater percentage titles loaded based on patron need, as defined by trigger. Big data analytics, physical therapy, and nursing studies seem to have satisfactorily met the needs of patrons in the overall percentage of titles loaded.

In a preliminary analysis of cost per use, the CSI library randomly sampled 21 print titles in the same subject areas as were triggered for purchase by DDA, during the same time period. The conclusion is that the cost per use for DDA titles is generally less than the cost per use of print, as chosen by librarians, or the total loss of an investment when there was no use of a print title.

In the first month of Kanopy’s PDA program, the CSI library had 63 plays or eight hours of play time at no cost except $150 for one trigger. The CSI library was eager to receive the useful data analytics the platform provides, including what devices patrons were using to access the videos, what browser the patrons were using, hours of the days the patrons accessed, patrons’ search keywords, and the referring URL patrons are using to access the content. It gave a better overall view of patron behavior than had been seen with EBSCO’s DDA project. Whether patrons are accessing titles via our catalog or CUNY’s discovery tool is a question that can be answered in further analysis.

The CSI library CD goals were much different for digital monographs versus digital streaming videos. In terms of monographs, we did not opt for short-term loans due to their increasing cost but wanted perpetual ownership. With videos, the turn-over of content is much higher, so an initial one-year license model was chosen. Then, in an evidence-based
model scenario, we could choose to outright purchase or do another multiyear license for the content based on usage statistics. With monographs, the library wanted to supplement traditional collection development practices and still purchase print books, but with videos, it was felt that the DVD format is almost obsolete, and streaming videos are preferred for both students and faculty. For both digital monographs and digital streaming videos, the platforms needed to have accessibility features for the disabled and good usage data analytics.

**Future Considerations for PDA/DDA**

The shortfalls of the PDA/ DDA pilot projects, lessons learned, and future prospects at CSI library for DDA/PDA are being considered. The first year of the DDA pilot project was not without challenges. For example, the ability to get good usage data from EBSCO was quite challenging. Knowledge of patron behavioral triggers per title must be manually entered once a month per the DDA statement generated by EBSCO. Any other data besides basic usage per title, which is mixed in with non-DDA titles in EBSCO’s administrative module, must be requested through EBSCO’s e-book team. This made it difficult to immediately analyze patron behaviors and needs. EBSCO is also not able to give the library the referring URL from which patrons accessed the DDA titles. Likewise, DDA MARC records generated by OCLC can take over two weeks to retrieve per large batch-load. The removal of DDA titles by first publishers and then from our catalog was another unexpected inconvenience.

Despite these challenges, we did feel were able to achieve an easy to manage demand-driven acquisition project that ran itself with minimal record maintenance and catalog batch-loads. The CSI library also was able to supplement traditional collection development practices with DDA and be sustainable throughout the project’s time period. So far, we are also seeing a good ROI and a cost-effective service, with only $3,000 spent to date but access to over $126,000 worth of titles.

The CSI library continues to explore doing more with less in terms of a minimal budget and maintenance. At this time, the CSI library’s collection development committee has re-allocated $10,000 of librarian firm order funds to ProQuest’s E-book Central DDA program, across all subject areas, which should be implemented by the spring of 2017. The goal of this pilot project is to expand DDA across all subject areas and become a permanent part of collection development practices at the library. The CSI library will need further historical data on Kanopy’s PDA to make any further considerations or changes to this second project.

<table>
<thead>
<tr>
<th>Subject Area</th>
<th>Records Loaded</th>
<th>% of Records Selected</th>
<th>Records Selected Or Triggered</th>
<th>% Records Selected Or Triggered</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Asian</td>
<td>161</td>
<td>38%</td>
<td>9</td>
<td>43%</td>
</tr>
<tr>
<td>Arabic Studies</td>
<td>87</td>
<td>21%</td>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>Data Analytics</td>
<td>51</td>
<td>12%</td>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>ESL/ TESOL</td>
<td>82</td>
<td>20%</td>
<td>6</td>
<td>29%</td>
</tr>
<tr>
<td>Legal Studies</td>
<td>30</td>
<td>7%</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Physical Therapy</td>
<td>5</td>
<td>1%</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Nursing</td>
<td>5</td>
<td>1%</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Figure 1. Analysis of the EBSCOhost DDA e-books selected and triggered.**
Conclusion

In all, the CSI library succeeded at achieving what it desired, which was a sustainable DDA within the specific time period without overspending the amount of funds allocated. In a preliminary analysis of circulation usage of DDA versus librarian-selected print titles during the same period, DDA appeared to give the library a greater return on its investment. Since success is the achievement of something desired, CSI Library succeeded at creating a sustainable DDA within the period specified with a minimal amount of funds allocated. CSI Library also saw success with a preliminary cost analysis per use of librarian “just-in-case” traditional purchases versus patron-driven “just-in-time” purchases. Future considerations for expanding DDA/ PDA practices utilizing typical firm order funds for print and DVD purchases is being planned as a resolution to doing more with less with a materials budget on the decline.

References


Ferguson, J., & Stangroom, S. (2014, November 7). Putting your patrons in the driver’s seat: Online video and PDA. Presentation at the Charleston Conference, Charleston, SC.
A Model for Patron-Driven Acquisition of Print Music Scores: From Conception to Reality

Alan Asher, Librarian, University of Florida

Trey Shelton, E-Resources Librarian, University of Florida

Jason Heckathorn, Access Support Assistant, University of Florida

Aimee Barrett, Acquisitions, University of Florida

Abstract

This paper and presentation will explore the process of developing a unique patron-driven acquisition program for print music scores and monographs from concept to reality at an American Association of Research Libraries institution. Areas to be discussed include collection development considerations, information technology infrastructure needs, acquisitions workflows, and plan evaluation. The paper and the presentation will examine how partnering with a vendor to implement an innovative collection development plan can support the needs of the library users and the goals of library collection development officers and increase access to music scores and monographs in a fiscally responsible way. Readers can expect to learn about the opportunities and challenges that the library and vendor faced in implementing the plan, outcomes and evaluation, and steps for the future.

The concept of patron driven acquisition of library materials has become a widely accepted practice in college and university libraries in North America. Now also known as demand-driven acquisition or user-driven acquisition, patron-driven acquisition has been largely driven by materials in electronic format. Public domain sheet music and scores are available in PDF formats from various sources on the Web, but copyrighted music scores have not transitioned well into electronic formats that are conducive to library patron-driven acquisition models.

The Smathers Libraries at the University of Florida have been active in developing different models for patron-driven acquisition of materials, including e-books and print resources for some academic disciplines such as art and architecture. The patron-driven acquisition of print music scores program at the University of Florida is modeled after an earlier experiment with patron-driven acquisition of art materials. The novelty of the University of Florida program is that library patrons can initiate a library purchase of a print music score directly from the online public catalog without the mediation of a librarian. Patrons are actively involved in shaping and developing the collection.

The School of Music at the University of Florida has a sizeable graduate student enrollment and offers the terminal degree in musicology, ethnomusicology, music education, and music theory and composition. Analysis of the music collection indicated there were deficiencies in the score and sheet music holdings related to contemporary music scores and for monographs related to music education and musicology. Capitalizing on a long-standing and successful relationship with the Harrassowitz Company, an agreement to develop a patron-driven acquisition model for print music scores and books was arranged with Harrassowitz in 2014, and the program was made available in the University of Florida’s library’s online public catalog in 2015.

In consultation with key members of the Smathers Libraries acquisitions department, the music librarian worked with staff from Harrassowitz to modify their standing order profiles for music materials to target areas of interest for the University of Florida. The Harrassowitz Company then sent brief catalog records, matching our profile needs, to be loaded into the online catalog. At the present time, there are about 7,300 titles loaded into the catalog. Music scores and sheet music account for about 73% of these titles and number over 5,000. Books and monographs, numbering nearly 2,000 titles, account for the remaining 27% of titles in the program. The profiles created for the program offer coverage of music-related books and
scores from all of Europe, the United States, Canada, Japan, Australia, and Israel, and represent all of the major languages used in musicological scholarship. Since much contemporary music is published in nonstandard formats, the profiles allow for inclusion of spiral bindings, large scores over 37 centimeters, and holograph reproductions of manuscript scores if a print edition is not available.

Smathers Libraries information technology staff created an online order form that is embedded in each catalog record. When a patron identifies a book or score of interest and clicks on the “request this item” link, the link generates an authentication page which auto-populates with information from the user’s integrated library information system. Once authenticated, the user may then initiate a purchase of the book or the score. Materials purchased in this way are sent to the Smathers Libraries in regular, biweekly shipments of materials from Harrassowitz. Once on campus, the materials are rush processed, and the patron is notified via e-mail that the book, sheet music, or score is ready for pick up. In most instances, music materials ordered through the patron-driven acquisition model are available for use within four to six weeks of the placement of the order.

Once the work of creating the online order form and loading of catalog records from Harrassowitz was completed, the day-to-day operation of the program transitioned into the normal workflow for Smathers Library acquisitions staff. The volume of materials received from orders initiated by the program has not created stress in the processing or handling of materials in the workflow chain from receipt of a title to delivery to the patron.

Accounting of expenditures for the program have been tracked through standard workflows in the acquisitions department.

Reception of the program by the students and faculty of the University of Florida School of Music has been favorable. In the first full year of operation, 91 titles have been purchased, 53 of the purchases were music-related book titles at a cost of $3,460 with an average cost per book of about $65, and 38 music scores have been purchased at a cost of $1,740 with an average cost per score of about $46. The total expenditure for the program for the first year was $5,200 with an average cost per title among all formats of about $65. As was hoped at the outset of the program, most of the titles purchased have been directly related to areas such as contemporary music scores and books about music education, teaching, and pedagogy, which were identified through collection analysis as areas of need.

The program has allowed music faculty and students to order needed resources directly from the online catalog without the mediation of a librarian, and it has greatly expanded the availability of research materials related to music education, composition, and musicology. Staff workflows in the acquisitions department have been able to absorb the additional materials without stress, and the budget allocation for music resources available to the music librarian has been expanded with the addition of budget dollars for patron-driven acquisition of materials. The patron-driven acquisition of print music scores and books has been successful enough that the Smathers Libraries has expanded the program for print materials to include other academic disciplines, including foreign language books and monographs.
DDA Management With Predictive Modeling

John Vickery, Analytics Coordinator and Collections & Research Librarian for Social Sciences, North Carolina State University Libraries

Abstract

Demand-driven acquisitions (DDA) programs have become an integral part of academic libraries’ collecting strategies. While DDA programs provide an effective way to build a just-in-time collection, it can be difficult to anticipate how many titles will be triggered for purchase and what the financial impact will be. This presentation will describe a project to build a predictive model to flag DDA titles that are likely to be triggered for purchase within the first year of being added to the catalog. By implementing a predictive model, collections and acquisitions departments can better plan the yearly DDA budget. In addition, titles with a high probability of being triggered for purchase can be purchased if they become ineligible for DDA. We will discuss how we combined text analytics and structured data as inputs to the model using a combination of Statistical Analysis System (SAS) and Python.

Introduction

This presentation describes a project to build a predictive model to flag DDA titles that are likely to be triggered for purchase within the first year of being added to the catalog. As demand-driven acquisitions (DDA) programs have become an integral part of academic libraries’ collecting strategies it is important to investigate ways to anticipate how many titles will be triggered for purchase. In addition, as publishers consider reducing their participation in DDA programs, a predictive model can be used to simulate a demand-driven acquisition strategy by identifying titles that have a high probability of being used.

About the Data

The data used to train the model started with approximately 49,000 bibliographic records from the North Carolina State University (NCSU) libraries’ Sirsi ILS. These records represented DDA titles added to the catalog between December 2011 and June 2015. Each record was identified as either currently untriggered or as having already been triggered for purchase. Purchase dates for each title were compared to the date the title was added to the catalog in order to calculate a binary variable indicating if the title was triggered for purchase within 12 months of being added to the catalog. This was represented as the variable WithinYear and was coded as 0 for nonpurchased and 1 for purchased. It was this variable that was used as the modeling target variable. Of the titles in the initial training data set, 11% had been purchased within 12 months of being added to the catalog. The majority (95%) of the titles were published between 2011 and 2016. The two most common publishers represented in the data set were Routledge and Palgrave Macmillan.

In order to accommodate the relatively rare frequency of a title being purchased within 12 months, the initial data set was over sampled to a WithinYear = 1 frequency of 50%. The oversampling process resulted in a final training data set of 11,138 records. 100% of the WithinYear = 1 records (n = 5569) was combined with a random sample of 5,569 of the WithinYear = 0 records.

Data sources used to compile the training data set consisted of local bibliographic data from the libraries’ Sirsi ILS and MARC records as well as data from external sources. The two sources for external data were Online Computer Library Center’s (OCLC) Classify API and Proquest-Syndetic Solutions book summary data. The Classify web service available at http://classify.oclc.org/classify2/api_docs/index.htm was used to augment records where the library of congress classification was missing. The Syndetics book summary data was accessed via API. Note that a subscription to the Syndetics service is required to access the API. Book summary data and how it was used is discussed in more detail in following sections.

Software Used for the Project

Three software products were used to build the training data set and develop the predictive model. Base SAS was used for the majority of the data extraction and cleaning of Sirsi ILS data. Python
programs were used to retrieve data from both the OCLC Classify and Syndetic API's. Finally, SAS Enterprise Miner (with text miner) was used for the model training and validation. SAS Enterprise Miner was also used to score a hold-out data sample as a final test of the model.

**Model Inputs**

Model inputs were purposefully restricted to data that would be apparent to a user while viewing a title in NCSU Libraries’ catalog. Figure 1 highlights data used as potential model inputs. Specifically, the following data items were selected as potential inputs: Publisher, publication year, Library of Congress subclass, and Syndetic Solutions book summary. The Syndetic book summary data can be considered unstructured, text data while the other inputs are structured. In order to be used as model input, the unstructured book summary data must be converted to a structured format. The following section describes the process of deriving structured topics from the book summary.

**Topic Modeling Process**

In order to use the Syndetic book summary data as input to the predictive model, the unstructured text must be converted into structured numeric data. The method used for this project is referred to as topic modeling. Topic modeling is a process of automatically associating a document with collections or terms characterizing a theme or idea. In this case, each summary represents a “document.” The topic modeling algorithm assigns a score to each document and term. If the association is strong enough, the document is assigned to a particular topic. Each document can be assigned to one or more topics or even no topics. It is these numeric scores representing topics or themes that are then used as inputs to the predictive model. As mentioned in a previous section, the text miner module of SAS Enterprise Miner software was used to model topics from the Syndetic book summary data. Chakraborty, et al. (2013) provide an excellent overview and case study of the topic modeling process. Blei (2012) also provides an overview of topic modeling.

![Figure 1. Potential model inputs.](image)
Extracting topics from unstructured text is a multistep process. Figure 2 illustrates the main steps involved. Summary data for each DDA title was gathered using a Python program to fetch the book summary from the Syndetic Web service. Note that a subscription to their service would be required. In the next step, the software (SAS Enterprise Miner with Text Miner) parses the collection of book summaries into a term-document frequency matrix where each element of the matrix equals the number of times that term appears in a document. During this step, low-value words such as the, and, a, etc. are dropped from the analysis. The following step involves refining the term list by interactively dropping or combining terms. For example, in the case of this project, the term book could be dropped as each summary is referring to a book, and the term, therefore, does not add value to the analysis.

In the next step, the SAS software algorithm reduces the term-document matrix dimensions using the singular value decomposition method and extracts topics from the document collection. While topics can have many terms associated with them, the top five terms for each topic are displayed as output. Table 1 lists examples of topics as represented by their top five terms. The final step is to interpret the generated topics and determine if they make sense. If the topics do not make sense, it may be necessary to refine the list of terms that are combined or dropped. The numeric score for each topic can then be used as inputs to a predictive model.

### Table 1. Example topics represented by their top five terms (+ symbols indicate that the term has been stemmed or synonymized).

<table>
<thead>
<tr>
<th>Top five terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>engineering</strong>, +engineer, +engineer, +design, +application</td>
</tr>
<tr>
<td><strong>clinical</strong>, +treatment, +medical, +therapy, +professional</td>
</tr>
<tr>
<td><strong>mathematics</strong>, +application, +mathematical, +problem, +solution</td>
</tr>
<tr>
<td>+disease, +molecular, +biology, +cell, +protein</td>
</tr>
<tr>
<td><strong>software</strong>, +design, +basic, +guide, +learn</td>
</tr>
<tr>
<td>+woman, +man, +male, +identity, +sexual</td>
</tr>
<tr>
<td>+introduction, +accessible, +concise, +comprehensive, +philosophy</td>
</tr>
<tr>
<td>+material, +engineer, +polymer, +technology, +material science</td>
</tr>
<tr>
<td><strong>food</strong>, +food, +nutrition, +product, +food</td>
</tr>
<tr>
<td>+law, +legal, +court, +right, +legal</td>
</tr>
<tr>
<td>data, big data, +process, +application, +technique</td>
</tr>
</tbody>
</table>
Results

A step-wise logistic regression model (Cramer & Howitt, 2004) and a decision tree model (Salkind, 2010) were trained on the data. The models’ performance was then compared based on misclassification rate. Based on misclassification rate, a decision tree model was selected as the best performing. The final model relied solely on inputs derived from the topic modeling process described previously. These topics are shown in Table 1. Figure 3 shows a graphical representation of a portion of the final decision tree model.

Figure 3. A portion of the final decision tree.

Figure 4. Caption: An 8% misclassification rate.
References


