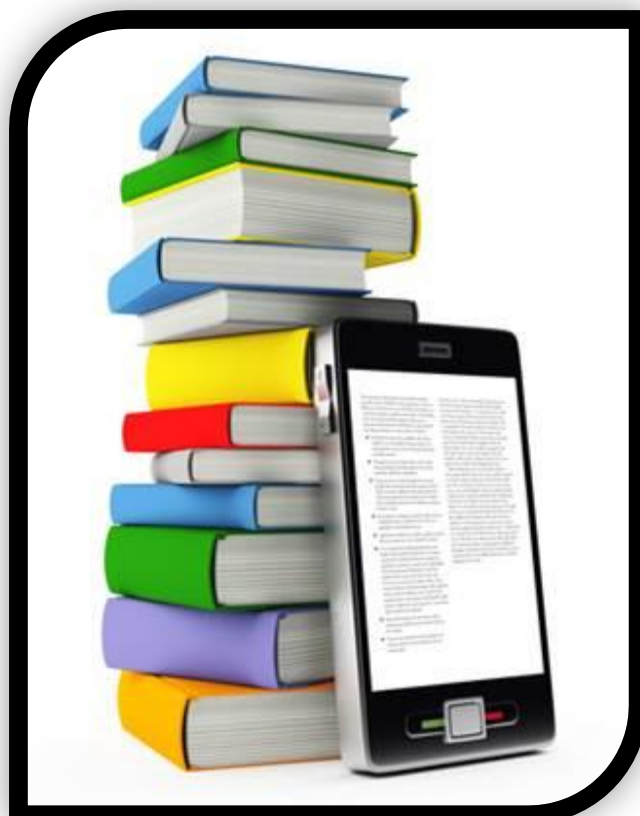# E-Book Program Development Study: Quarterly Report
December 1, 2013 – February 28, 2014

Melissa Goertzen, E-Book Program Development Librarian

Columbia University Libraries

mjg2227@columbia.edu

Executive Summary

The E-Book Program Development Study is an ambitious assessment project aimed at gathering essential data to drive the development of policies related to e-book development programs. It aligns with CUL's mission to support the development and delivery of high-quality services that facilitate research, teaching, and learning across campus and within the wider scholarly community. The results will provide a set of recommendations and policies for internal and external stakeholders as they collaborate on the development and implementation of e-book projects and programs.

The central objective of the third quarter was to collect data sets that will inform e-book collection development policies relating to preferred business models, acquisition methods, and resource allocation. Much of this work centered on a large-scale cost analysis project. Using the research framework established during the second quarter of the study, quantitative data was collected to document how funds are allocated, examine usage rates, and develop a methodology for future evaluation.

At the same time, a collaboration with Nisa Bakkalbasi, Assessment Coordinator, resulted in the development of a new and innovative means to gather information about e-book use across disciplines. The method relies on a qualitative analysis of e-book search terms harvested by Google Analytics and e-book titles from COUNTER e-book usage reports. In February 2014, the data set was collected and analysis is currently underway.

The second objective was to continue documenting the internal and external e-book landscape. For instance, the third quarter saw the completion of a collection development policy review and interviews with library staff at CUL. At the same time, discussions about e-book discovery, access and use continued with members of the greater academic community. This included meetings with the University Dean for Libraries and Information Resources at CUNY and three sales managers from YBP and EBSCO.

In summary, the third quarter focused on quantitative data collection and experimentation with innovative research methods. The results will inform e-book collection development policies relating to preferred business models, acquisition methods, and resource allocation. This work provides a context for study results and suggests how the e-book collections align with the Libraries' overarching mission to support research, teaching, and learning activities across campus. This context also creates an essential framework to craft a vision for the future direction of e-book curation and preservation at CUL.

Contents

Introduction

The E-Book Program Development Study is an ambitious assessment project aimed at gathering essential data to drive the development of policies related to e-book development programs. It aligns with Columbia University Libraries' (CUL) mission to support the development and delivery of high-quality services that facilitate research, teaching, and learning across campus and within the wider scholarly community. The results will provide a set of recommendations and policies for stakeholders as they collaborate on the development and implementation of e-book programs.

Summary of the Second Quarter

The objective of the second quarter was to determine what types of data provide evidence for how the e-book collection is discovered, accessed, and used. This work also involved establishing a research design that informs the creation of strategies and recommendations about e-book collection development at CUL.

At the same time, documentation of the e-book landscape at CUL continued. Much of this work involved meeting with members of the Selectors' Group and reviewing collection development policies. In tandem with these activities, a review of the external landscape began. This involved discussions with members of consortia, librarians at peer and local institutions, and e-book providers.

To accomplish these objectives, the following 5 tasks were completed:[1]

1. Research Design and Methodology;
2. Data Collection;
3. Internal review of the e-book landscape at CUL;
4. External review of the e-book landscape in the academic community and publishing industry;
5. Dissemination of first quarter results.

In summary, the results of the second quarter provide an assessment framework that informs all data collection and analysis activities. Together, the selected research methods will create a body of quantitative and qualitative data that documents the e-book landscape at CUL and serves as a baseline for future evaluation. The data will also suggest what e-book services the Libraries' can reasonably provide to stakeholders and target areas where CUL can provide leadership in the academic community through advocacy.

The reality that the e-book landscape is currently unstable was also factored into decisions regarding the assessment framework. The research design was created so that it

---

[1] For further details related to these activities, please refer to the Second Quarterly report (http://academiccommons.columbia.edu/catalog/ac%3A168771).

can be replicated regardless of how e-books evolve in the coming years. Because the design is flexible and adaptive in nature, it promotes continued assessment, evaluation, and strategic planning as a regular component of e-book programs.

Third Quarter: Objectives

The central objective of the third quarter was to collect data sets that will inform e-book collection development policies relating to preferred business models, acquisition methods, and resource allocation. Much of this work centered on a large-scale cost analysis project. Using the research framework established during the second quarter of the study, quantitative data was collected to document how funds are allocated, examine usage rates, and develop a methodology for future evaluation. For further details, please see the report section entitled "Cost Analysis Project" (p. 6).

During the same time, a collaboration with Nisa Bakkalbasi, Assessment Coordinator, resulted in the development of a new and innovative means to gather information about e-book use across disciplines. The method relies on a qualitative analysis of e-book search terms harvested by Google Analytics and e-book titles from COUNTER e-book usage reports. In February 2014, the data set was collected and analysis is currently underway. For further details, please see the report section entitled "Content Analysis Project" (p. 7).

The second objective was to continue documenting the internal and external e-book landscape. The third quarter saw the completion of a collection development policy review and interviews with library staff at CUL. At the same time, discussions about e-book discovery, access and use continued with members of the greater academic community. This included meetings with the University Dean for Libraries and Information Resources at CUNY and three sales managers from YBP and EBSCO. This work provides a context for study results and suggests how the e-book collections align with the Libraries' overarching mission to support research, teaching, and learning activities across campus.

To accomplish these objectives, the following 5 tasks were completed:

1. Data collection: Cost analysis project
    a. Discussed e-book fund allocation and budgets with Jeff Carroll, Director of Collection Development, to select a representative data sample for the study;
    b. Met with Colleen Major, Head of Electronic Resources Management: Operations and Analysis, to discuss a data collection strategy using COUNTER reports and e-book title lists;
    c. Generated a list of e-book holdings at CUL based on e-book subscription, package, and title lists;
    d. Worked with Daisy Alarcon, Bibliographic Assistant, to collect and organize data pulled from fund reports, title lists, and COUNTER reports;

e. Conducted statistical and usage analysis on data sets;
f. Mapped a sample set of e-book titles to LC Classifications for subject analysis.

2. Data collection: Content analysis of usage reports and search terms project
   a. Collaborated with Nisa Bakkalbasi, Assessment Coordinator, to collect data from Google Analytics and COUNTER reports;
   b. Co-authored a paper proposal with Nisa Bakkalbasi for the CUNY Assessment Conference.

3. Internal review of the e-book landscape at CUL
   a. Met with Zachary Lane, ReCap Coordinator, to discuss the undergraduate population and focus group strategies;
   b. Completed a review of the collection development policies and mandates in place at CUL's twenty-two libraries;
   c. Attended an H&H department meeting to hear Tom Harford, Dean of Students for General Studies, discuss the changing population and needs of GS students;
   d. Attend all Selectors' Group meetings to learn about selection and acquisition procedures.

4. External review of the e-book landscape in the academic community and publishing industry
   a. Met with Curtis Kendrick, University Dean for Libraries and Information Resources, to discuss e-book initiatives and policies at CUNY;
   b. Met with three sales managers from YBP and EBSCO to learn about e-book business models and acquisition options;
   c. Documented various business models available to libraries through major academic publishers, aggregators, and distributors.

5. Disseminated results from the second quarter to solicit feedback from stakeholders
   a. Presented second quarter results at the Management Committee meeting in January 2014;
   b. Presented second quarter results at the Selectors' Group meeting in January 2014;
   c. Posted second quarter results in the Academic Commons and CUL wiki.

E-Book Program Development Librarian Work Plan	Jun-13	Jul-13	Aug-13	Sep-13	Oct-13	Nov-13	Dec-13	Jan-14	Feb-14	Mar-14	Apr-14	May-14	Jun-14	Jul-14	Aug-14	Sep-14	Oct-14	Nov-14	Dec-14	Jan-15	Feb-15	Mar-15	Apr-15	May-15
Literature Search	█	█	█																					
Initial Overview (Depts)	█	█																						
Review Strategic Plan	█	█	█	█																				
Data Gathering	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█			
Quarterly Reviews			▮		▮		▮			▮			▮			▮			▮			▮		
Internal Review				█	█	█	█	█																
Collection Development Policy Review				█	█	█	█	█																
Iterative Process for E-Book Package Offers				█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█			
Collaborative Collection Development				█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█			
External Review						█	█	█	█	█	█	█	█	█	█	█	█	█	█	█				
Discoverability (In CLIO, MARC Record quality, etc.)					█	█	█	█	█	█	█	█	█	█										
Review Order Workflow					█	█	█	█	█	█	█	█	█	█	█	█								
Material that cannot be purchased by libraries					█	█	█	█																
Analyze Subscription vs. Purchase					█	█	█	█	█	█	█													
Evaluate PDA/DDA					█	█	█	█	█	█	█													
Functionality					█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█			
Focus Groups					█	█	█	█	█	█	█													
Usability Studies					█	█	█	█	█	█	█													
User Experience					█	█	█	█	█	█	█													
Archive/Preservation					█	█	█	█	█	█	█													
Standards Review					█	█	█	█	█	█	█													
Publisher Review						█	█	█																
Annual Review												▮												
Setting Policy																				█	█	█	█	
Summary and Recommendations																				█	█			
Wrap up																						▮	▮	▮

Table 1. E-Book Program Development Work Plan

Third Quarter: Results

Cost Analysis Project

The goal of the project is to collect quantitative data that will inform e-book collection development policies in regards to fund allocation, preferred business models, and acquisition methods.

The project is guided by the following four objectives:

1. To document how funds are allocated to build e-book collections at CUL;
2. To document usage rates of e-book purchases and subscriptions;
3. To develop a standardized data collection model that can be replicated on an annual basis;
4. To develop a body of quantitative data that will serve as a baseline for future evaluation.

The results will answer the following six questions:

1. Which items fall under e-book collection development policies at CUL?
2. How many items are in the e-book collection?
3. What e-book packages does CUL purchase/subscribe to?
4. How are funds allocated to build e-book collections?
5. What are the top five e-book packages/subscriptions/titles?
6. Are resources directed towards packages/subscriptions/titles that are widely used?

After discussions with Jeff Carroll, Director of Collection Development, and Colleen Major, Head of Electronic Resources Management, it was determined that e-books are most often purchased on the EO or EB fund codes. For this study, data collection was limited to titles, packages or subscriptions that had fund activity during the 2013 fiscal year (FY2013).

To collect financial data for all e-book purchases, a Voyager query was run for all library funds ending in EO (subscriptions) or EB (firm orders). After running the cumulative query, a base list was created for each category: subscriptions (EO), package purchases (EB packages) and individual purchases (EB firm orders). The lists include the title, fund code, price, and a brief price analysis.

Spending for each of the three categories was totaled, and calculations were made to identify the top 70% (bulk) and bottom 30% (tail) of purchases within each budget. Statistical analysis was conducted to determine the total, average, median, high, and low costs of each category. Finally, items in each category were sorted by cost, largest to smallest. The top 5 items by cost were selected for further usage analysis.

To conduct usage analysis, title lists for each of the top 5 items were downloaded from the CERM wiki. At the same time, the corresponding COUNTER report was pulled from the vendor/publisher website. Then, data was filtered to remove titles published prior to 2013.

We discovered that in several cases, multiple collections from the same vendor are purchased as separate items on the EO or EB fund codes. However, there is no apparent way to filter COUNTER reports by collection. At this point, we considered analyzing the data by vendor/publisher instead of by collection, but decided that this method would skew results because of the discrepancies in cost, size, and use. Instead, we filtered the data for a second time by matching the 2013 title lists with COUNTER report data. This process was completed in Excel using the MATCH formula. Based on these results, we calculated the number of titles loaned, number of loans, percentage of titles without use after purchase, the average cost of an e-book, and cost per use.

While data collection for the cost analysis project is complete, analysis is currently underway. Final results will be available by the end of the fourth quarter (May 31, 2014).

Content Analysis Project

During the second quarter, it was proposed that a survey be distributed to faculty and students to collect data about e-book discovery and access trends across disciplines. However, based on the low response rate from a recent survey conducted in the Fall semester, a decision was made to look for alternative approaches to gather data and potentially schedule the e-book survey for Fall 2014. Through collaboration with Nisa Bakkalbasi, Assessment Coordinator, an alternative approach to the survey method was developed to gather information about e-book use across disciplines. The method relies

on a qualitative analysis of e-book search terms harvested by Google Analytics and e-book titles from COUNTER e-book usage reports.

Data collected through COUNTER usage statistics and the LibQUAL+ service quality assessment survey tell us that faculty, graduate students, and undergraduates value access to the growing e-book collection at CUL. While the aggregate results indicate that e-book use continues to increase, usage rates are not uniform across disciplines. Anecdotal evidence suggests that while e-book use has grown in the sciences and social sciences, scholars in the arts and humanities rely heavily on print books. Given the highly diverse research needs of the university community, CUL is keen to understand scholarly e-book usage in various disciplines.

The study method utilizes data from two sources: readers' e-book search terms harvested by Google Analytics and requested e-book titles provided by the COUNTER e-book usage reports. The data was analyzed using NVivo, a qualitative analysis software, to examine popular scholarly e-book topics and the correlation between search and delivery.

Additional Activities

Throughout the third quarter, I continued work with the ERUDWG committee. Our projects focused on analysis of the new COUNTER 4 report format. It was helpful to discuss the format with colleagues and I gained a deeper understanding of how the reports can be used to analyze and compare usage trends across e-book collections.

In December 2013, Dr. Cristina Pattuelli at the School of Information and Library Science at the Pratt Institute invited me to attend the Human Information Behavior class to speak about the E-Book Program Development Study. I also had the opportunity to watch students present the results of their own research projects. It was a wonderful opportunity to learn about up-an-coming topics related to electronic resource management and I look forward to future opportunities to connect with Pratt students.

Since July 2013 I have been a member of the ACRL/NY Committee, and we recently hosted the 2013 Symposium. It was an excellent opportunity to connect with information professionals in the area and learn about innovative projects taking place at other institutions. In addition, I wrote two articles summarizing talks by keynote speaker Susan Gibbons, University Librarian at Yale University, and Tom Scheinfeldt, Associate Professor of Digital Media at the University of Connecticut. Both were published in the January 2014 ACRL/NY newsletter.

In February 2014, I attended the CUL/IS Communication Workshop in order to learn about various ways to connect with project stakeholders. The session provided me with ideas regarding how to raise awareness about e-book programs at CUL. I look forward to attending similar communication events in the future.

Conclusions

The third quarter focused on quantitative data collection and experimentation with innovative research methods. The results will inform e-book collection development policies relating to preferred business models, acquisition methods, and resource allocation. Also, the third quarter saw the completion of a collection development policy review and interviews with library staff at CUL. This work provides a context for study results and suggests how the e-book collections align with the Libraries' overarching mission to support research, teaching, and learning activities across campus. This context also creates an essential framework to craft a vision for the future direction of e-book curation and preservation at CUL.

Next Steps

1. Complete the cost analysis project
    a. Submit results to ERUDWG for feedback in April 2014
    b. Generate graphs to highlight significant results
    c. Draft a paper on this project and search for publication opportunities
2. Complete the content analysis of usage reports and search terms project
    a. Upload datasets to NVivo for analysis
    b. Draft a paper on this project and search for publication opportunities
3. Submit an IRB application to Columbia University to prepare for summer focus group sessions
4. Continue documentation of e-book policies and workflows within the academic community and publishing industry
5. Disseminate results from the third quarter to solicit feedback from stakeholders