

The E-Book Program Development Study is an ambitious assessment project aimed at gathering essential data to drive decisions related to e-book services and collection development activities. It aligns with CUL's mission to support research, teaching, and learning initiatives across campus. Study findings create a bridge between the current information landscape and CUL's vision for the future of e-book programs and services.

E-Book Program Development Study

Results and Recommendations
2013 - 2015

Melissa Goertzen
E-Book Program Development Librarian
Columbia University Libraries (CUL)

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 collections and services at Columbia University Libraries (CUL). It aligns with CUL's mission to support the delivery of high-quality services that facilitate research, teaching, and learning across campus and within the wider scholarly community. The results provide a set of recommendations for stakeholders as they collaborate on the development of e-book projects and programs.

The primary objective of the study was to document the e-book landscape at Columbia University and understand how current challenges fit into the larger context of collection development and management within the academic community. A second objective was to create innovative and sustainable assessment methodologies that enable librarians to collect data and evaluate e-book holdings in a standardized fashion. The quantitative and qualitative data collected over the past two years serves as a benchmark for the future evaluation of e-book holdings and collection development practices at CUL.

The framework of the study was structured around the e-book lifecycle. Primary areas of focus included selection and acquisition, discovery, access and use, preservation, and evaluation. The success of the project involves its ability to explore these areas in an innovative and methodical fashion.

One highlight of the study was an extensive cost analysis project that resulted in a standardized method to evaluate e-book subscriptions and front list purchases. Over the course of six months, cost data from Voyager, title lists from publishers, and COUNTER usage statistics were collected and analyzed in order to evaluate the perceived value of 97 e-book subscriptions and 35 front list packages. After collections were ranked according to cost, the top six subscriptions and packages were ranked according to cost per use figures. The findings aided in negotiations with vendors regarding subscription renewal costs and resulted in a savings of over \$60,000 for FY15. Since that time, the cost analysis method has been adopted by the Science and Engineering Library as a regular part of their analysis and assessment program. It has continued to yield successful results that led to significant cost savings in FY16.

A second highlight was the development of an innovative text analysis methodology that provided insight into users' search and discovery behaviors across different disciplines. The method was developed during a yearlong collaboration with Nisa Bakkalbasi, Assessment Coordinator at CUL. It utilized text data from two sources: users' e-book search queries entered into CLIO and e-book title words provided by COUNTER usage reports. Using NVivo, a mixed method analysis software, word frequency queries were run using different text match options to generate lists of the most frequently occurring search terms and e-book title words. The results provided insight into the nature of e-book use across disciplines, including broad topic (e.g. history), academic level of use (e.g. introductory), and genre/type (e.g. reference).

To verify the text analysis method, Nisa and I conducted focus group and interview sessions with faculty members and graduate students to understand *for what purpose* and *why* e-books are used in conjunction with scholarly activities. Findings suggest that the decision to use e-books for academic purposes is largely dictated by the reading technique most suited to the task at hand. Participants said that they prefer print formats for continuous reading (e.g. prolonged reading without interruption for research or learning purposes, making annotations, etc.) and electronic formats for discontinuous reading and quick reference (e.g. confirming citations, referring to quotes, etc.). We also discovered that faculty and students do not have a preference for specific technologies or platforms through which to access e-books. Rather, they desire a variety of discovery, access, and functionality options that can be tailored to particular circumstances and working environments. Essentially, they desire the ability to customize e-book use to suit a specific set of needs in a given moment.

The overarching theme that has come up time and time again is the intent of e-book use. The findings suggest that collection development activities ranging from the selection of business models and formats, the negotiation of license agreements, the marketing of collections, and the development of preservation strategies hinge on one central question: what is the intent of use? Once this question has been answered, librarians can enter into meaningful conversations with stakeholders, advocate for the needs of user communities, and develop collections composed of a variety of formats that meet both immediate and long-term information needs.

The reality that the e-book landscape is constantly evolving was factored into decisions regarding the overarching assessment framework guiding this study. The research design was created so that it 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.

Table of Contents

Part One: Introduction	6
Part Two: Research Objectives and Study Design	8
2.1. Introduction to Columbia University Libraries.....	8
2.2. Research Objective	9
2.3. Study Framework.....	9
2.4. Preliminary Information Scan	10
2.5 Study Work Plan.....	11
Part Three: Summary of Results	13
3.1. Overarching Theme: Intent of Use.....	13
3.2. Selection and Acquisition.....	14
Review of E-Book Business Models.....	14
Cost Analysis Project: E-Book Subscriptions.....	15
Cost Analysis Project: E-Book Front Lists.....	17
3.3. Discovery.....	18
3.4. Access and Use.....	19
3.5. Preservation.....	20
2.7 Unresolved Areas.....	22
Part Four: Recommendations	23
4.1. A Vision for E-book Collection Development	23
4.2. Collection Development Goals.....	24
Collection Depth Indicators.....	24
Criteria for E-Book Purchases (Packages and Individual Titles) .	25
Criteria for E-Book Subscriptions and Databases	25
Criteria for Adding Free E-Books to the Collections	26
4.2. Recommended Evaluation Workflows.....	26
Evaluate E-Book Subscriptions & Packages	26

Evaluation of the Collection Development Department	27
Part Five: Conclusions	28
Appendix A: The E-Book Landscape.....	29
A.1. Survey of the Internal E-Book Landscape	29
A.1.1. Research at CUL	29
Gutenberg-e	29
A Book Online? No, An Online Book!.....	29
Users Unbound	29
2CUL EBooks Task Force – Final Report.....	30
E-Books: Who Wants What and How to ILL Them.....	30
A.1.2. SWOT Analysis of CUL.....	30
A.2. Survey of the External E-Book Landscape	33
A.2.1. E-Book Landscape in North America (Literature Review)	33
A.2.2. The International E-Book Landscape	54
Overview of Emerging E-Book Markets	54
Emerging Markets: Japanese Case Study	55
Emerging Markets: Argentine Case Study.....	58
Appendix B: Selection and Acquisition.....	60
B.1. Review of E-book Business Models	60
B.2. E-Book Cost Analysis Project.....	63
B.2.1. Development of a Cost Analysis Method	63
B.2.2. E-Book Subscription Cost Analysis Project	67
B.2.3. E-Book Front List Cost Analysis Project.....	68
B.2.4. Paper: 2014 Charleston Conference	70
B.3. E-Book Comparison Chart.....	77
Appendix C: Organization and Discovery	78
C.1. Paper: 2014 CUNY Library Assessment Conference.....	78

C.2.	Paper: Exploring E-Book Use through Text Analysis	88
Appendix D: Access and Use		101
D.1.	Faculty Interview and Student Focus Group Questions.....	101
D.2.	Focus Group Questionnaire	102
D.3.	Paper: Exploring E-Book Discovery and Use.....	104
D.4.	The Question of Intent: Building E-Book Collections.....	114
Appendix E: E-Book Preservation		116
E.1.	Blog Post: The Fine Art of App Preservation	116
E.2.	E-book Preservation Special Report	117

Part One: Introduction

Two years ago I arrived at Columbia University Libraries (CUL) and was tasked with completing the E-Book Program Development Study. The central goal was to document the e-book landscape on campus and within the academic community in order to create collection development recommendations and strategies that will shape future e-book services at CUL. When I read the project description, the opportunity seemed both exciting and daunting; the opportunities for experimentation were enormous, but I knew that the complexities involved with e-book collection development, not to mention the speed at which formats and technologies are evolving, would present significant challenges. In the beginning I felt like I had to provide solutions that would solve the collection development challenges we face on a daily basis. Now, I believe that discovering the right questions to ask in this evolving landscape is the work that guarantees long-term success of collection development policies and strategies.

Our questions set us on a journey of discovery and open our minds to opportunities for innovation. When we challenge our assumptions, look at issues from multiple perspectives, and test what we believe against evidence collected along the way, we begin the process of pulling back the layers of a problem to uncover the golden thread, the core issue that ties together all of the seemingly disconnected elements of an investigation. This process provides a context for research findings and can be the starting point for strategic planning and collaborative relationships that define how collections and services will be delivered in the future.

Several months into the E-Book Program Development Study, I realized that I was operating under the assumption that users prefer electronic content for research, teaching, and learning activities. Throughout society, we rely so heavily on the Internet, mobile technologies, and social media to gather and disseminate information that I assumed users would have a preference for e-books for all scholarly activities. However, as I started to collect usage statistics, examine discovery and access trends, and speak with faculty, students, and library staff, I realized that my initial impressions of content use were far too simplistic and did not tell the full story. As I dug deeper into data sets, I started to ask more and more questions about when, how, and why users gravitate towards certain formats (e.g. print, electronic, archival materials) to support scholarly activities and build knowledge around specific subject areas.

At first, I worried that the discovery of questions rather than solutions would result in a lengthy list of proposals for future studies. But, the opposite turned out to be true. The inquiry process provided a focus and pulled everything I had observed into one overarching question: what is the intended use of e-book content? Users interact with information for a variety of reasons including course use, research pursuits, and general reference. All of these activities serve different functions within a research community, rely on different forms or reading (e.g. continuous vs. discontinuous), and support different information needs.

When I consider the intent of use from a collection development perspective, my mind automatically begins to separate activities into the categories of ‘current use’ and ‘future use’ (Yale University Library, 2013). This seems appropriate as there are few business models or collection development strategies that address both requirements at once. I think this separation points to a general shift in the way academic libraries in the 21st century must approach collection development activities; successful initiatives rely on a balance between ‘just in case’ and ‘just in time’ strategies. This balance allows information professional to determine when it makes economic sense to invest resources in high use materials for current users and when it is appropriate to purchase materials that may have low use but add to the long-term value and legacy of the collection (Yale University Library, 2013). Again, having a strong understanding of information needs and how faculty, students, and researchers intend to use e-book titles, packages, and subscriptions provides the insight required to make these calls.

My work on the E-Book Program Development Study has convinced me that going forward, collection development decisions relating to the selection of business models, the negotiation of license agreements, and the development of outreach or marketing initiatives, hinges on the Libraries’ understanding of the intent of e-book use. Once we have successfully answered this question, we can enter into meaningful conversations with stakeholders, advocate for the needs of users during conversations with publishers and aggregators, and build collections that satisfy immediate and long-term information needs.

As the E-Book Program Development Study comes to a close, I look forward to setting out on a new journey of exploration and building on the benchmarks established over the past two years.

Melissa Goertzen, E-Book Program Development Librarian
Columbia University Libraries

Part Two: Research Objectives and Study Design

2.1. Introduction to Columbia University Libraries

Columbia University in the City of New York is a world-class research institution that brings together a community of approximately 4,362 faculty members and 26,000 students. While Columbia College is highly regarded for its undergraduate programs, close to three-quarters of the total student body is enrolled in graduate and professional programs. It is also home to one of the top five academic research library systems in North America; collections contain more than 12 million volumes, 160,000 journals and serials, and extensive electronic resources, audio-visual materials, and archives. Services and collections are organized into 21 libraries that are run by a team of more than 450 professional and support staff.

Columbia University Libraries' (CUL) experience with e-books dates back to the 1990s, beginning with experiments such as the Online Books Project, the Virtual Reading Room, and pilot programs with netLibrary. In 2004, CUL began purchasing e-books in an experimental capacity. Due to positive reception by faculty and students, CUL expanded e-book collections to support research, teaching, and learning activities across campus. Over the last four years, the e-book collection has increased exponentially in size; CUL now provides access to over two million e-book titles (see Table 1) and expenditures comprise 25 per cent of the total book budget. In response to this growth, CUL is developing a unique strategy and vision for e-book programs and initiatives across campus.

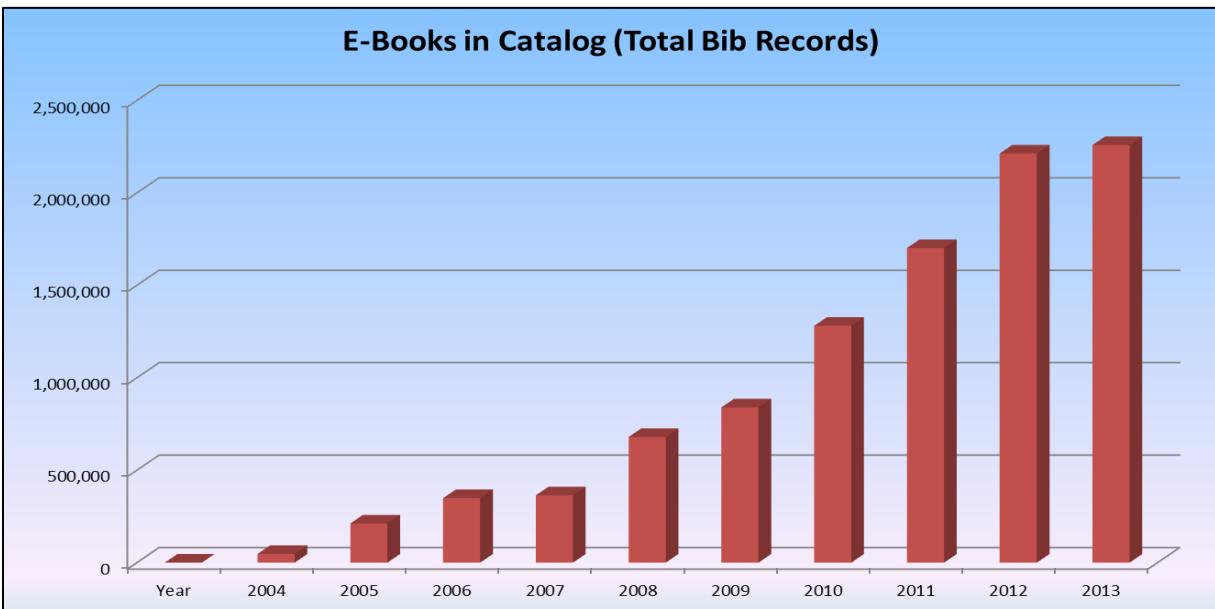


Table 1. E-book acquisitions at Columbia University Libraries.

In May 2013, I was hired to complete the E-Book Program Development Study, an ambitious two-year assessment project aimed at gathering essential data to drive the development of policies related to e-book programs and services. It aligns with CUL's mission to support the delivery of high-quality services that facilitate research, teaching, and learning across campus and within the wider scholarly community. The results provide a set of recommendations for stakeholders as they collaborate on the development and implementation of e-book projects and programs.

2.2. Research Objective

The objective of the study was to document and assess the e-book landscape at Columbia University Libraries a) internally, b) within the context of the academic community, and c) within the context of the e-book publishing industry. The research design was guided by the following four principles outlined in the CUL/IS Strategic Plan 2010 – 2013:

1. User-focused design;
2. Data-driven decision making;
3. Continuous assessment of results;
4. Flexible and adaptive response to user needs.

(CUL/IS Strategic Plan 2010-2013, p. 8)

Study findings create a baseline for future evaluation, establish innovative methods that promote ongoing evaluation, and gather together a rich body of quantitative and qualitative data that provided evidence for when, how, and why users discover, access, and use e-books for research, teaching, and learning activities. Finally, the set of recommendations that result from study findings create a bridge between the current landscape and a future vision for e-book initiatives on campus.

2.3. Study Framework

After an initial scan of the e-book landscape at CUL and within the academic community, it was evident that information professionals across North America grapple with similar challenges in regards to e-book collection development activities. Due to the vast and complex nature of the current information landscape, addressing each challenge is well beyond the scope of this study. In order to pinpoint the most pressing issues at CUL, seventy informal interviews were conducted with library administrators, directors, selectors, and library staff.

Given CUL's active involvement with consortial groups (e.g. 2CUL, MaRLI, NERL, and IviesPlus), consideration was given to how recommendations may impact decisions regarding collaborative collection development activities and the assessment of materials acquired through joint purchases. In order to maximize the potential for future partnerships within the academic community, all data was collected from sustainable sources available to all consortial partners.

Finally, the fact that the e-book landscape is evolving at a rapid pace was taken into heavy consideration. The business models, formats, devices, and platforms evaluated over the last two years may be obsolete within the next decade. As such, the study results identify a number of future studies that promote continual evaluation of the e-book landscape in order to stay abreast of changes in the market place and users' information needs.

Based on this context, the following framework guided the E-Book Program Development Study (see Figure 1).

1. Develop a set of recommendations that address challenges related to selection and acquisition, discovery, access, and preservation at CUL;
2. Examine how study findings provide opportunities to evaluate and strengthen collaborations with partner institutions and content providers (e.g. vendors, publishers, aggregators).
3. Establish an evaluation framework that facilitates regular assessment and planning so that collection development strategies can be updated and revised as the e-book landscape evolves.

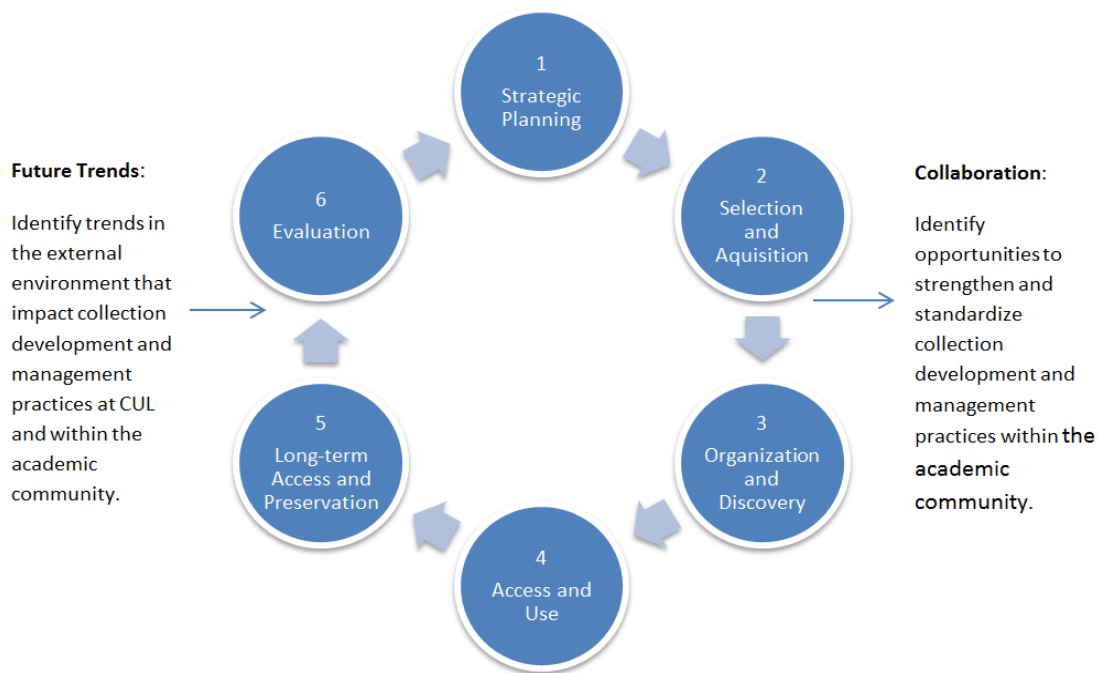


Figure 1. E-Book Program Development Study Framework.

2.4. Preliminary Information Scan

In order to build on a body of e-book investigations that have taken place at CUL over the past twenty years, I conducted an initial information scan to target areas where additional inquiry would be desirable. Below is a list of nine areas where knowledge gaps exist. E-Book Program Development Study addressed these topics. To view the full results of the information scan, please see Appendix A.

1. Document discipline-specific needs and expectations regarding e-books and their role in research library collections;
2. Determine how vendor statistics can be used to make collection development decisions;
3. Investigate how resources are allocated to build e-book collections and whether the library receives a return on investment;
4. Document and evaluate e-book lending services and business models;
5. Identify and assess trends in e-book discovery;
6. Investigate e-book usage trends;
7. Investigate and provide evidence for desired e-book functionality;
8. Document challenges related to e-book preservation;
9. Develop sustainable assessment methods that promote evaluation of e-book collections.

2.5 E-Book Program Development Study Work Plan

An ambitious work plan was established at the start of the E-Book Program Development Study. The goal of these activities was to build a body of knowledge that informs collection development strategies and recommendations (see Table 2). The work plan divided study activities into five categories:

1. Internal review of the e-book landscape at CUL (green);
2. External review of the e-book landscape in academia and publishing industry (yellow);
3. Observation of e-book workflows (blue);
4. Data collection and analysis (red);
5. Dissemination of preliminary results to internal and external stakeholders (orange).

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 2. E-Book Program Development Study Work Plan

Internal and External Review of the E-Book Landscape

The purpose of the internal and external reviews were to document the e-book landscape at CUL and understand how the needs and challenges across campus fit into the larger context of the academic community and publishing industry. This work involved reviewing previous studies and projects completed by CUL staff, interviewing seventy library staff (e.g. administrators, directors, and selectors), reviewing collection development policies, conducting a literature review, evaluating e-book business models, developing relationships with colleagues at partner institutions (e.g. Cornell, NYU, NYPL, etc.), and discussing the e-book landscape with a variety of publishers, vendors, and aggregators (e.g. ProQuest, JSTOR, Taylor & Francis, Ingram, 3M, etc.).

Observation of E-Book Workflows

The purpose of observing e-book workflows was to determine how e-book titles are acquired, discovered, accessed, and preserved at CUL. This work involved examining fund allocations, studying e-book workflows and procedures developed by the Continuing and Electronic Resources Management Division (CERM), determining how metadata records are acquired, analyzing usage statistics collected from COUNTER reports, and examining existing preservation practices.

Data Collection and Analysis

The purpose of data collection and analysis was to develop rich quantitative and qualitative data sets that inform e-book collection development activities and provide baselines for future evaluation. Much of this work centered on three large-scale projects: a cost analysis of 96 active e-book subscriptions and 35 front list packages; an analysis of users' search and information retrieval behaviors; a series of faculty interviews and student focus group sessions; and a number of case studies to document trends on e-book creation and distribution.

Dissemination of Preliminary Results to Internal and External Stakeholders

The purpose of disseminating study results to stakeholders was to solicit feedback from the professional community and promote discussion about the current e-book landscape. Over the course of the study, research posters and papers were presented at local, national, and international conferences (e.g. CUNY Library Assessment Conference, the 2013 & 2014 Charleston Conferences, and the 7th International Conference on Qualitative and Quantitative Methods in Libraries). The strengthening of relationships and collaborations between academic institutions and publishers may result in best practices that work to standardize e-book collection development practices at CUL and within the research community as a whole.

Part Three: Summary of Results

3.1. Overarching Theme: Intent of E-Book Use

The overarching theme discovered through the E-Book Program Development Study is the intent of e-book use. The findings suggest that collection development activities ranging from the selection of business models and formats, the negotiation of license agreements, the marketing of collections, and the development of preservation strategies hinge on one central question: what is the intent of use?

In reality, there is no e-book solution that simultaneously meets both the ‘current use’ and ‘future use’ requirements (Yale University Library, 2013). In some cases, it makes economic sense for libraries to purchase titles in electronic formats without consideration for long-term access. In other cases, it is appropriate to purchase materials for preservation purposes despite levels of current user demand (Yale University Library, 2013). Today, it seems that the success of collection development initiatives relies on a balance between ‘just in case’ and ‘just in time’ strategies. As such, it is imperative for librarians to have a comprehensive understanding of how faculty and students intend to use specific e-book packages, subscriptions, and even titles. With this knowledge, we have the insight required to determine when it makes economic sense to invest resources in high use materials for current users and when it is appropriate to purchase materials that may have low use but add to the long-term value and legacy of the collection.

It is important to note that as the information landscape continues to evolve at a rapid pace, user intent will evolve as well. Below are the indicators that were used in the E-Book Program Development Study to uncover intent of use. These nine indicators will serve as a starting point for future evaluation of collection development policies and documentation of users’ information needs.

1. Solicit feedback from faculty and students to monitor evolving information needs and the ways in which collection materials are used to support research, teaching, and learning activities;
2. Develop strong channels of communication with liaison librarians to monitor information needs of faculty and student;
3. Consider whether materials will be used for continuous or discontinuous reading;
4. Monitor usage trends across e-book subscriptions and packages. Consider how usage trends match against collection depth indicators;
5. Monitor how e-books are searched and discovered by studying search queries and keywords entered into the OPAC;
6. Develop a method to document titles placed on reading lists and course reserves;
7. Document research trends across disciplines (e.g. citation studies, grants).

3.2. Selection and Acquisition

Note: To view the full results of all research activities related to selection and acquisition, please see Appendix B.

Review of E-Book Business Models

To explore emerging trends that will impact e-book collection development activities at CUL, an ad hoc group came together in 2014 to evaluate a number of e-book platforms, services, and business models that may offer new acquisition and access options for users at Columbia University. This work subsequently resulted in the development of an evaluation framework for new potential offers from vendors and partnerships with partner institutions. The task force consisted of the following six members:

1. Bob Wolven, Associate University Librarian for Bibliographic Services and Collection Development;
2. Rob Cartolano, Associate Vice President for Digital Programs and Technology Services;
3. Jeff Carroll, Director of Collection Development;
4. Breck Witte, Director of Library Information Technology Office;
5. Susan Marcin, Head of Electronic Resources Management: Technologies and User Experience;
6. Melissa Goertzen, E-Book Program Development Librarian.

Representatives from 3M Cloud Library, BiblioBoard, MyiLibrary, and Taylor & Francis were invited to give presentations on campus. After each meeting, the task force reviewed content offerings, proposed license agreements, the platform interface, and acquisition workflows. In addition trial access was set up for each platform and selectors were invited to preview content and submit feedback.

At the same time, the ad hoc group held meetings with Terry Kirchner, Executive Director of the Westchester Library System, and Denise Hibay, Head of Collection Development at the New York Public Library, to learn how each of the above mentioned platforms functions in a practical library environment. The conversations revealed that a significant percentage of e-book budgets are directed towards platform and maintenance fees, which significantly reduces the percentage of funds that are available to purchase or license titles.

After reviewing content, business models, acquisition workflows, and feedback from selectors, it was determined that at this time the above mentioned platforms do not have applications at CUL. As an alternative to licensing these products, CUL and NYPL discussed a library card pilot that would promote e-book content at both institutions to the user community at Columbia University. To review full results, please see Appendix B.1.

Cost Analysis Project: E-Book Subscriptions

One of the initial goals of the E-Book Program Development Study was to create an inventory of all e-book holdings at CUL. It was soon apparent that due to the volume and complexity of the data set, this project was outside the scope of the study.

At this point, I began looking at available data sets and decided to experiment with a new methodology that would compile a snapshot of e-book purchases (i.e. subscriptions, packages, and firm orders) made throughout the 2013 fiscal year (FY2013). The goal was to document how funds are allocated to purchase e-book content, and determine if acquired content is of value to the user community.

After discussions with the Director of Collection Development, and the Head of Electronic Resources Management, it was determined that e-books are most often purchased on the EO fund code. To collect financial data for all e-book subscriptions, a Voyager (CUL's integrated library system) query was run for all library funds ending in EO. After running the cumulative query, I created a base list of all e-book subscriptions at CUL. Given the enormous size of this data set and the time restrictions placed in the study, data collection was limited to subscriptions that had fund activity during the 2013 fiscal year (FY2013). This limited the base list to 96 e-book subscriptions. Next, subscription fees in FY2013 were totaled, and calculations were made to identify the top 70 per cent (bulk) and bottom 30 per cent (tail) of purchases within the budget. Statistical analysis was also conducted to determine the total, average, median, high, and low costs.

To examine usage trends, the top four subscriptions (ranked by cost) were selected and corresponding title lists were collected from the Continuing & Electronic Resources Management (CERM) Division. At the same time, the corresponding BR2 COUNTER report was pulled from the vendor website. At this point, I encountered an unexpected challenge; in several cases, multiple collections from the same vendor are purchased as separate items on the EO fund code. However, there is no apparent way to filter COUNTER reports by collection. At this point, I considered analyzing the data by vendor instead of by collection, but decided that this method would skew results because of the discrepancies in cost, size, and use. Instead, I filtered the data for a second time by matching the 2013 title lists with COUNTER report data.

Next, I adapted a cost analysis framework used by the University of Western Australia to analyze DDA models (Davies & Morgan, 2013, p. 172). After examining our data set, I 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 (Table 3).

EO Fund Analysis	Subscription A	Subscription B	Subscription C	Subscription D
2013 Cost	\$54,788.40	\$21,280.00	\$16,071.33	\$57,005.85
No. of titles	89,529	6,631	116	14,339
No. of sections viewed	33,834	2,185	85	1,649
No. of loans	2,581,195	119,725	21,899	6,219
% of titles without use after purchase	62.21%	67.05%	26.72%	88.50%
Average cost of e-book	\$0.61	\$3.21	\$138.55	\$3.98
Cost per use	\$0.21	\$0.18	\$0.73	\$9.17

Table 3. Results of the cost and usage analysis of e-book subscriptions (EO funds.)

After analyzing the cost and usage data of the top e-book subscriptions, it was determined that the cost per use of Subscription D was high (\$9.17 per use) compared to Subscription A (\$0.21 per use), Subscription B (\$0.18 per use), and Subscription C (\$0.73 per use).

The results were presented to the E-Resource Usage Data Working Group (ERUDWG) at CUL. The consensus was to conduct a second analysis of Subscription D based on the following criteria: evaluation of content, overlap analysis, and interface analysis. The results indicated that Subscription D contained a large number of outdated technical manuals (96 per cent published before 2011), a high number of titles available through other platforms, and incomplete multivolume sets.

Next, a team of librarians from the Science and Engineering Library examined the title list and identified 394 high use titles (more than 20 page views). Of this subset, they discovered that 196 titles (49.75 per cent) are available through other platforms at CUL. Then, they searched GOBI for the remaining 198 high use titles and discovered that the vast majority are available for individual purchase. Based on this analysis, it was determined that Subscription D does not contain a significant amount of unique content. A decision was made to cancel the subscription.

One unexpected outcome of this analysis was the opportunity to speak with the Vice President and a team of sales representatives managing Subscription D on three separate occasions. The company requested feedback from CUL regarding how to improve the platform interface and content, and were provided with study findings. After a series of negotiations, Subscription D was renewed for one year at an 80 per cent discount, resulting in a cost savings of \$50,000.

At the end of FY2016, Subscription D was assessed using the same methodology. A decision was made to cancel the subscription based on calculated value to the user community at CUL.

The cost analysis methodology described above was adopted by the Science and Engineering Library Division (SEL) as a part of their regular collection analysis and assessment program. It has proven an effective method to examine the value of e-book subscriptions and packages to the user community. Results of subsequent evaluations have resulted in the successful negotiation of subscription costs, annual increase, and terms of license agreements. To full the full method and results, please see Appendix B.2.2.

Cost Analysis Project: E-Book Front Lists

At CUL, a large percentage of resources are directed towards front lists. When I analyzed the top five packages ranked by cost (Table 4), the cost per use appeared high (averaging at \$36.00). A closer examination of the data revealed that many front list titles are not available to CUL users until the end of the year (largely due to publication dates). It seemed that evaluating the cost per use of 2013 front list titles based on 2013 COUNTER usage reports did not accurately reflect their value.

EB Fund Cost Analysis	Package A	Package B	Package C	Package D	Package E	Package F
2013 Cost	\$59,681.40	\$62,515.00	\$113,377.81	\$216,888.00	\$114,950.00	\$66,907.63
No. of titles	704	446	408	5,988	1,553	1,945
No. of titles loaned	153	79	111	4,509	294	686
No. of loans	2,937	486	2,143	119,085	8,620	11,360
% of titles without use after purchase	78.27%	82.29%	72.79%	24.70%	81.07%	64.73%
Average cost of e-book	\$84.77	\$140.17	\$277.88	\$36.22	\$74.02	\$34.40
Cost per use	\$20.32	\$123.63	\$52.91	\$1.82	\$13.34	\$5.89

Table 4. Results of cost and usage analysis of e-book packages.

In order to develop a method to evaluate the cost per use of e-book front lists, the Package F front list was selected for evaluation because of the fact that it provided access to a rich data set (e.g. title lists, usage statistics) spanning several years. I located the Package F title list and matched it against BR2 COUNTER reports ranging in date from January 2011 to April 2014. Then, I experimented with a method to observe how usage and cost change over time.

It was expected that the number of loans would also increase over time, but the results indicate a different trend. Between 2011 and 2012, loans increased by more than 80 per cent. In the following year, the number of loans dropped by more than 50 per cent. After considering collection content and usage trends, it seems that there are two possible explanations: 1) the titles were included in course reading lists and/or course reserves, and 2) users downloaded Package F titles in 2012 when they became available through CUL. During this analysis, an attempt was made to identify all Package F titles that were included in course reserves over the past three years. However, the time involved to extract this data is not conducive to the time frame for the E-Book Program Development Study. The topic has been flagged for a future study and may inform recommendations made at the end of the project, as findings would provide information about how patrons use the e-book collection.

To view the full method and results, please see Appendix B.2.3.

3.3. Discovery

To view the full results of all research activities related to selection and acquisition, please see Appendix C.

In 2014, a collaboration with Nisa Bakkalbasi, Assessment Coordinator, resulted in the development of an innovated method that investigates users' e-book search behavior and information needs across different disciplines. The study covered the period from January 1, 2014 to December 31, 2014. The research method utilizes text 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 sets present CUL with an accurate, continuous, and objective picture of e-book use. After data clean-up and formatting, requested e-book titles from COUNTER reports and e-book search terms from our Google Analytics account were loaded into the qualitative analysis software, NVivo to identify frequently used words and explore recurring patterns. Then, we performed text analysis to generate word frequency tables and word clouds for each of the frequency sets to graphically display how each of the collections, at least in terms of the titles used, covers a different sector of the e-book platform universe.

We discovered that the users' searching behavior when using a library catalog is different than users' searching behavior on the Web. The average search query length when searching CLIO ($m=3.62$ terms) compared to the average search engine query length ($m=2.4$ terms) may indicate that library users are more sophisticated in how they structure their queries when they are looking for very specific items or specific answers.

Running e-book search queries and requested e-book title words through a text analysis tool revealed new ideas and concepts related to what types of materials our users search for and use. Text analysis of search terms and requested title words provided insight into the nature of e-book use across disciplines, including broad topic (e.g. history), academic level of use (e.g. introductory), and genre/type (e.g. reference). It is challenging to deduce reader intent from word frequency analysis, as text data remain widely open for interpretation. However, responses to open-ended questions from the most recent LibQUAL+ survey are consistent with our findings that e-book collections are widely used across all major disciplines to support instruction and learning.

The study of search phrases provided a good deal of information about what types of e-books users search for, but much less information about why searches are conducted or how satisfied users are with the discovery process as a whole. Despite its limitations, there are two significant strengths of the methodology that drive us to utilize it in future studies: (1) real queries from real users with actual information needs and usage, and (2) the volume of the data analyzed and time-span it covered. To view the full results of this study, please see Appendices C.1. and C.2.

3.4. Access and Use

To view the full results of all research activities related to selection and acquisition, please see Appendix D.

To verify the text analysis method discussed in the previous section, Nisa and I conducted focus group and interviews sessions with faculty members and graduate students to understand *for what purpose* and *why* e-books are used in conjunction with scholarly activities. Participants answered eight pre-determined questions during one-on-one interviews and dual moderator focus group sessions. They were also invited to complete a questionnaire regarding e-book discovery, access, and use. All sessions were transcribed and the data was analyzed using grounded theory approach to examine emerging themes.

One of the central themes that we discovered in both the focus group and interview sessions was that the decision to use e-books for academic purposes was largely dictated by the reading technique most suited to the reading task at hand. Participants said that they prefer print formats for continuous reading (e.g. prolonged reading without interruption for research or learning purposes, making annotations, etc.) and electronic formats for discontinuous reading and quick reference (e.g. confirming citations, referring to quotes, etc.).

Through discussions with faculty and students, we discovered that the decision to use e-books for discontinuous reading and quick reference is related to a large advantage provided by the format: convenience. Based on our discussions with faculty and students, we determined that convenience is composed of three criteria: content availability (e.g. not having to “wait in line” for an e-book), accessibility (e.g. remote access, availability of multiple file formats, etc.), and usability (e.g. search, print, and download capabilities).

When we pursued the topic of convenience further, we discovered that faculty and students do not have a preference for specific technologies or platforms through which to access e-books. Rather, they desire a variety of discovery, access, and functionality options that can be tailored to particular circumstances and working environments. Essentially, they desire the ability to customize e-book use to suit a specific set of needs in a given moment.

When we analyzed the qualitative data collected from interview sessions, we discovered a theme that was entirely unique to faculty members. This is the idea of perpetual access to content and information, not only in regards to e-books themselves, but also in terms of annotations and notes. The faculty members we spoke with do not feel confident that electronic annotations will be preserved for years to come. They said that in many cases, they refer to annotations over the course of their entire career. If an e-book is pulled from a platform or a publisher ceases to exist, faculty are at risk of losing work that is irreplaceable. In many cases, we found that faculty prefer to annotate print formats because they know the text, along with their work, will remain safe on their shelves for years to come.

As we conducted the focus group sessions, we also discovered one theme that was unique to graduate students: they are willing and capable of creating e-book experiences that suit their needs, even if this means working around the library. Students indicated that through word of mouth, they often find online learning environments that are suited to their research and learning needs. These can include subscription, foreign, and in some cases, illegal databases of electronic content. WE found that many students are also willing to pay subscription fees to e-book databases that provide desired functionality, including the ability to do full text searches, make annotations, and pulling citations in a variety of styles. The results suggest that student place high value on user experience, access, and ultimately convenience.

The findings suggest that faculty and graduate students use e-books on a regular basis for discontinuous reading and quick reference, and value convenience and the ability to customize research and learning environments to suit immediate needs. They acknowledged that it is frustrating when availability, accessibility, and usability are hindered by the limited functionality of academic interfaces and platforms. Participants readily acknowledged that these challenges are not linked directly to the library. However, they said the library can advocate for the user community and work with publishers to develop flexible business models that promote greater convenience and flexibility online.

3.5. Preservation

To view the full results of all research activities related to selection and acquisition, please see Appendix E.

Due to the rapid rate at which technology is evolving, attempts to preserve e-books are likely to prove messy, uncoordinated, inconclusive, and incomplete. However, engaging multiple stakeholders in small-scale experiments that address technical, legal, or financial challenges may foster a sense of collective ownership and build towards systemic solutions. The recommendations that follow are offered, not as resolutions, but as explorations.

Focus on Outcomes

Preservation is about ensuring the survival of e-books so that they may continue to be available for future use. However, different types of content may suggest different types of use, by different groups, at different times. Focusing on one standard, ideal model of preservation for all types of e-books is likely to impede incremental progress that may ultimately get us closer to the ideal. For any group of e-books, libraries will need to consider:

1. Who is likely to want access in future? For what purposes? When?
2. Is preservation of the content (text and images) sufficient? How important are other features?

3. Are there classes for which a semi-dark archive, with limited or embargoed access, is sufficient?

Build Consensus on Priorities

A high percentage of the e-books purchased by libraries are still also produced in print and preserved in that form. At the other extreme, many works produced only as e-books are not purchased by libraries at all. With limited resources to devote to e-book preservation, should libraries focus on value (protecting investment, broad interest) or risk (ephemeral interest, entrepreneurial publishing)? Should they seek quick results (low-hanging fruit), or concentrate on difficult questions with potential for broad impact? Different stakeholders may well answer these questions differently. Finding like-minded partners may help organize action on multiple fronts.

Clarify Roles

There is broad consensus within the library community that preservation is important, but little discussion of who can and should take responsibility for different aspects – negotiating rights; storing content; maintaining format integrity; migrating access; supporting costs. Different stakeholders are likely to have different perspectives, and roles may vary for different types of e-books. Two types of action are needed:

1. Sectors within the library community organized by type (public, academic), size (large, small) and function (preservation, collection development) need to organize discussions and build consensus around willingness to act and effective locus of activity.
2. Libraries, individually and through organizational channels, need to propose a variety of potential models to publishers and providers of e-books to test their willingness and ability to support specific preservation actions.

Experiment

A major barrier to e-book preservation is that it has been difficult to see how anyone can take action without massive changes in the way e-books are licensed, distributed, and stored. While systemic change is still difficult, small scale experimentation is possible.

1. CLOCKSS, LOCKSS, and Portico are already preserving tens of thousands of e-books. Libraries can take a more active role in encouraging new publishers to participate in these programs.
2. New open-access ventures such as Knowledge Unlatched have secured agreement from HathiTrust, Portico, and others to preserve and provide access to their e-book publications, a model that might be extended to additional content.
3. Some of the e-books most at risk are the output of innovative, start-up ventures (some of which have already suspended activity). These firms might be amenable to a “rescue” operation in which one or more libraries would secure the right to preserve (and possibly provide access to) their e-book content.

4. Research libraries might explore a model used for some print archives, in which a particular publisher's e-book output would be deposited with the library for preservation, with access embargoed for an interval, and/or restricted to researchers on the archive's premises.
5. Many self-published books may ultimately be of more interest as a corpus than for reading as individual titles. Libraries might seek agreement from providers of self-publishing platforms to secure and preserve the content for non-consumptive use.

Reward Success

If innovation requires a willingness to fail, it is also true that limited successes, however imperfect, must be encouraged. Publishers and libraries who demonstrate a willingness to take risks in order to preserve e-books ought to be recognized, with encouragement to others to follow and build on their efforts.

2.7 Unresolved Areas

It is important to note that e-books are still an evolving format; libraries and content providers are all grappling with the development of sustainable business models, discovery strategies, and accessibility issues, and will continue to do so for the foreseeable future. In this environment, there exist a number of central challenges that do not have answers and will require a great deal of experimentation to arrive at solutions.

Below is a list of challenges that are currently unresolved in the academic community. It is recommended that continued assessment, evaluation, and discussion with stakeholders continue on a regular basis to monitor the development of these areas.

1. How will CUL address the publication of e-books that librarians would like to acquire but currently do not have the technical capacity or legal rights to do so?
2. Increasingly, e-books are made available as PDF files that are sold by publishers or authors. Currently, CUL does not have a mechanism to host these files or make them widely available to the user community. How will CUL address this challenge going forward?
3. E-book ILL is an issue that creates access barriers and impacts how e-book collections can be built in collaboration with partner institutions. This is an area that should be monitored for potential systemic solutions as various pilot programs and ILL services are developed (e.g. the current ILL collaboration between Springer and Occam's Reader).
4. How will CUL address e-books that are available through Open Access initiatives? What types of financial commitments and collection development activities will we take on?

Part Four: Recommendations

4.1. A Vision for E-book Collection Development at CUL

The Collection Development department oversees the development, management, and storage of collections in a variety of formats. The overarching goal of all departmental activities is to support the distinctive and distinguished learning environment at Columbia University by providing access to information that supports research, teaching, and learning activities. Collection Development also serves the research community by participating in cooperative efforts to collect, access, and preserve knowledge at the local, national, and international levels.

CUL's collection represents a mix of traditional formats and a growing body of electronic resources, particularly in regards to e-book collections. Our challenge is to provide the most comprehensive coverage of collection areas in formats that are suitable to the diverse information needs of our user community.

In the current information landscape in which we work, print and electronic formats complement each other. The E-Book Program Development Study, along with findings presented in the literature review, indicate that these formats continue to serve specific research and learning needs in user communities. In this landscape, it is recommended that 'just in case' collection development activities focus on e-books and 'just in time' activities focus on print.

Collection Development recommends that selectors consider whether materials will be used for continuous (e.g. reading for extended periods of time, conducting in-depth research, exploring subjects in depth) or discontinuous reading (e.g. reference, citation confirmation, searching for keywords, skimming chapters) at the time of purchase. The answer to this question will be a strong indicator regarding which format is most appropriate. Based on study results, it is recommended that print serve continuous reading needs and electronic serve discontinuous reading needs. The results were consistent across the major disciplines observed during this study (i.e. humanities, social science, science, and fine arts).

The Collection Development department will continue to solicit feedback from faculty and students regarding collections that support research, teaching, and learning activities. Based on this feedback, we will monitor trends that may impact how we acquire print and electronic materials to fulfill 'just in case' and 'just in time' priorities. This work involves developing strong communication channels with department heads and liaison librarians at the 21 libraries in the system. In addition, it is recommended that interview and focus group sessions continue periodically as resources allow to measure shifting discovery, access, and usage trends across campus.

Over the coming years there will be greater pressure placed on available financial resources as information is made available through in a growing variety of formats and services. It is not

expected that the budget will grow alongside demand for these services. As such, the Collection Development department recommends continuous evaluation of collection materials. This is particularly relevant to new and current subscriptions. It is recommended that new subscriptions remain 'on probation' for a period of three years. This will provide sufficient time to monitor usage trends, cost, duplication, and quality of content.

4.2. Collection Development Goals

The following goals and recommendations support activities related to resource allocation, communication across CUL, cooperative collection development, and collection management activities.

Collection Development Goals

1. Allocate the materials budget and perform selection in a systematic manner than maximizes coverage, minimizes gaps, and avoids unnecessary duplication.
2. Promote communication between collection development, technical services, acquisitions, and selectors to make the most informed decisions.
3. Anticipate and respond to users' needs.
4. Consider the complete life-cycle of e-books at the time of purchase.
5. Continually evaluate collections and monitor research trends on campus.
6. Participate in cooperative collection development agreements with partner institutions.

Collection Depth Indicators

Basic Collection: E-books Recommended

Supports lower-division undergraduate research; includes the core of the discipline or sub-discipline as it relates to the curriculum. This level describes materials that serve to introduce and define subjects including selected databases, fundamental materials, introductory works, historical surveys, and reference works. ILL is expected to augment the collection.

Extensive Collection: E-books Recommended

Supports graduate course work; information is adequate to maintain knowledge of a subject required at less than research intensity. Examples of content include primary and critical resources, reference resources, specialized databases, and bibliographical resources. ILL is expected to augment the collection.

Research Collection: Print Recommended

Supports research leading to a doctorate, faculty research, or independent study. It includes resources supporting the framework for the methodology and implementation of original doctoral research. ILL is expected to augment the collection.

Recommended Criteria for E-Book Purchases (Packages and Individual Titles)

Based on study results, below are the criteria to consider in the decision to purchase e-book packages and titles:

1. Intent of use
 - a. Research: Print format recommended
 - b. Teaching and learning: Electronic format recommended
2. Intended audience
 - a. Basic collection: Electronic format recommended
 - b. Extensive collection: Electronic format recommended
 - c. Research collection: Print format recommended
3. Access is at least equivalent to print.
4. There is sufficient access (e.g. multiple user access for course reserves).
5. Cost and cost history.
6. Usage and availability of usage statistics for evaluation purposes.
7. Reliability and usability (e.g. stable links, demonstrates desired functionality).
8. Relevance of the content to the library's collection and university's curriculum.
9. Range of use (e.g. narrow range of use or interdisciplinary scope).
10. Content is preserved in stable repositories like Portico or HathiTrust.

Recommended Criteria for E-Book Subscriptions and Databases

Many e-books are available through databases, which are often subject to change. Access to titles may change depending on contractual agreements between publishers and aggregators, making it difficult for CUL to guarantee continued access to some e-books in our collection. If a title becomes unavailable through a database or subscription package, CUL will consider purchasing it directly from the publisher.

Recommendation: New subscriptions are considered 'on probation' for a period of three subscription years. At the end of that time, the subscription will be evaluated and a decision to continue or cancel will be made on the basis of use, quality of content, and cost.

Criteria to consider when ordering new subscriptions:

1. Degree of overlap.
2. Courses and/or programs supported.
3. Availability (e.g. full-text, stable links, print equivalent).
4. Holdings of peer libraries.
5. Price and price comparison with competitors.
6. Requestor (e.g. full-time faculty, instructor, student)
7. Content is preserved in trusted repositories like Portico and HathiTrust.

Recommended Criteria for Adding Free E-Books to the Collection

It is important to note that free materials bring costs to the library in terms of staff time and maintenance. One way to evaluate free content is to consider the following question: if you were required to pay for a free e-book from your collection budget, would you still want it as part of the collection? Below are four additional criteria to consider when acquiring free e-books.

1. CUL does not have access to the title through any other format.
2. Full bibliographic information is available.
3. Content is made available through a stable provider.
4. CUL has the means to host content and make it widely available to the user community.

4.3. Recommended Evaluation Workflows

Recommended Method to Evaluate E-Book Subscriptions and Purchases

Because there are a variety of stakeholders linked to each e-book subscription and package, a number of steps that must be taken to ensure appropriate individuals are brought into content evaluation projects and involved in collection development decisions. Below are the recommended steps for future evaluations of e-book collections.

1. Form a committee of librarians representing major disciplines on campus.
2. Determine what is an acceptable cost per use for disciplines (take into account how unique the content is, scope and cost of the resource, interdisciplinary scope).
3. On an annual basis, committee members should designate selectors to run cost analysis evaluations on major subscriptions and front list packages.
 - a. Present findings to the committee and flag subscriptions or packages that require further evaluation.
4. Alert the Collection Development Department, CERM, and the division head that a subscription/package has been flagged for evaluation.
5. Present cost analysis findings to the Electronic Resource Usage Data Working Group (ERUDWG) for further evaluation.
 - a. Evaluate content overlap, the platform's interface, the availability of high use titles as individual purchases through GOBI.
 - b. Gather feedback from faculty and students (delegate to the appropriate selector).
6. Present results of the evaluation to the Collection Steering Committee, Collection Development Department, and division head to approve large cancellations.
7. When appropriate, present findings to publisher, vendor, or aggregator.
 - a. Locate opportunities for price/content negotiations.
8. Notify the division head, the Collection Development department, CERM, and Selectors' Group when cancellation has been completed.

Internal Evaluation of the Collection Development Department

The Collection Development department is staffed by two librarians. As such, time it at a premium and we are continuously evaluating our processes to find ways to streamline work and build strong collaborative relationships with internal stakeholders. Below is an outline of next steps to create efficiencies and promote communication between departments.

Creating Efficiency:

1. Create a fund transfer form to internally manage fund transfers.
2. Manage requests for expensive item approvals through the email alias items4approval@columbia.edu.
3. Consolidate e-book funds for the purchase of Big Deal packages.
4. Systematically update FTE numbers to receive accurate quotes from vendors and publishers.
5. Systematically review profiles for major approval plans.
6. Schedule on-site vendor visits to hear updates regarding how the e-book landscape and associated business models are evolving.

Promoting Communication and Professional Development:

1. Create a Collection Development wiki and post training documentation, post updates, and disseminate information regarding subscriptions and/or package deals.
2. Update Collection Development policies available through publically facing websites.
3. Develop a means to disseminate information regarding license restrictions linked to e-book collections.
4. Develop documentation outlining Big Deal contracts.
5. Disseminate Selector Fund Reports and Director Fund Reports on a monthly basis.

Part Five: Conclusions

The work completed over the past two years suggests how the e-book collections align with CUL's overarching mission to support research, teaching, and learning activities across campus. More specifically, the efforts of the past year have resulted in the development of methodologies that examine how e-book resources are allocated, evaluate current subscriptions and packages, examine usage trends, and observe how users search and retrieve e-book content from the collection. The data that was gathered while developing these methodologies will be used as a baseline for future evaluation.

Over the course of this two year study, the overarching theme that came up time and time again is the issue of intent of use. That is, do users require requested information for research, teaching, or learning purposes? Will they engage with materials to support continuous or discontinuous reading? I am convinced that going forward, collection development decisions ranging from format selection, business models, license agreements, acquisition workflows, the marketing of collections, and preservation hinges on one central question: what is the intent of use? Once this question has been answered, librarians can enter into meaningful conversations with stakeholders and partners, advocate for the needs of user communities, and develop collections composed of a variety of formats that meet both immediate and long-term information needs.

The reality that the e-book landscape is constantly evolving was factored into decisions regarding the overarching assessment framework guiding this study. The research design was created so that it 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.

Appendix A: The E-book Landscape

A.1. Survey of the Internal E-Book Landscape

A.1.1. Research at CUL

Since the early 1990s, a number of discussions, presentations, and research projects surrounding e-books have taken place across campus. Together, this work provides a context for the E-Book Program Development Study. Below is a brief summary of these projects.

Gutenberg-e (late 1990s)

Gutenberg-e (<http://www.gutenberg-e.org/index.html>) is the results of a collaborative effort between Columbia University Press and the American Historical Association; the central goal is to promote the electronic publication of scholarly monographs and journals. This open access site showcases award winning monographs by emerging scholars and provides opportunities for online publication. Traditional narratives are combined with digitized primary sources and historic artifacts. All content is DRM free and contains hyperlinks to supplementary literature, images, videos, music, and related web sites.

A Book Online? No, An Online Book! (2009)

Rebecca Kennison, the former Director of the Center for Digital Research and Scholarship, published a blog post that addressed the fact that there are two distinct forms of electronic books available within the academic community – books online and online books. She argued that a book online follows the basic principles of a book in print: it has a table of contents, a bibliography, notes, and an image on the landing page that is often a recognizable book cover (Kennison, 2009). Navigation is linear, beginning with the acknowledgements and ending with appendices. In many cases, a book online is the digital equivalent of its print counterpart.

An online book provides a different experience. Namely, it has a continuing life, is interactive (e.g. links to related web pages, inclusion of multimedia content, ability to contribute comments), and allows for modifications to be made at any time (Kennison, 2009). Navigation is often based around a series of keyword searches entered into search engines that direct users to the middle of the text. In most cases, an online book does not have a print equivalent.

Kennison's article points to the need for preservation strategies that ensure not only long-term access to electronic information, but also to the ongoing accessibility and usability of electronic formats as technologies evolve and devices become obsolete.

Users Unbound – Reading, Libraries, and the Digital Humanities (2011)

Barbara Rockenbach, Director of History and Humanities, gave a presentation that discussed what libraries can do with new reading technologies to support user expectations and reading behaviors. She said that users want access to print and solitary study space, but also desire technologies that enable new ways to read and work with texts. Reconciling these multiple

desires may be one of the most exciting challenges libraries face in the coming decade (Rockenbach, 2011).

In a joint study conducted between Cornell and Columbia in which 45 in-person interviews were conducted, Rockenbach said the findings highlighted the distinction between continuous reading, which most preferred to do in print, and discontinuous reading (e.g. searching and browsing), which many want to do with a digital copy of a text (Rockenbach, 2011). Most preferred to read print content but valued the convenience of having books available digitally anywhere, anytime, especially when faced with limited space or the need to travel extensively for research (Rockenbach, 2011).

2CUL EBooks Task Force – Final Report (2012)

The 2CUL EBooks Task Force, reporting to the 2CUL Steering Committee and comprised of members from both Columbia and Cornell Libraries, was charged with surveying the diverse and rapidly growing landscape of e-books in libraries, including the areas of collection development, discovery, bibliographic access, usability, preservation, and emerging technologies, with a special eye to any synergies and potential opportunities between the two institutions. The report outlines the complexities of building collaborative collections and acquisition systems, even between two institutions that have a strong working relationship (Major et al., 2012).

E-Books: Who Wants What and How to ILL Them (2013)

Peter Bae gave a presentation at the 13th IFLA ILDS Conference that discussed the current challenges associated with e-book interlibrary loan (ILL) initiatives, including licensing restrictions and technical hurdles. Traditionally, the library has served as the central “platform” for access to print books – the collections are accessible through the stacks, reading rooms, or the general circulation system. Bae said there is no reason libraries should abandon this role in the current electronic environment (Bae, S. P., 2013).

In order to develop sustainable ILL systems, Bae said that libraries must first understand the information needs of the user community. This involves collecting evidence regarding how and why e-books are used for research, teaching, and learning purposes. Currently, there is no significant research or data answering those questions. The results of the E-Book Program Development Study fill this gap.

A.1.2. SWOT Analysis of CUL

After reviewing e-book research that was previously conducted at CUL, a SWOT analysis was conducted in order to survey the current e-book landscape on campus. The objective was to identify internal and external forces that help or hinder the implementation of e-book strategies and recommendations. The analysis was based on informal interviews conducted with seventy librarians at CUL, a review of the CUL/IS Strategic Plan 2010-2013, a review of existing

collection development policies and acquisition workflows, and a literature review that examined e-book trends in the academic community and publishing industry.

The results indicate that CUL is positioned within the academic community to develop innovative methodologies, propose business models, and initiate workflows that help to standardize e-book collection development practices and preservation strategies on a local, national, and international level. It is also in a strong position to advocate for the information needs of users during discussions with vendors, aggregators, and publishers.

The analysis also pointed to a number of challenges that may impact collection development and preservation activities. These include the threat of hardware and software obsolescence, the reality that many e-book collections are licensed, restrictions on use by way of Digital Rights Management (DRM), and the threat of company mergers and closures. Also, the current e-book market in general is geared towards consumers. This provides library users with expectations (e.g. content access, functionality) that are often not met by academic platforms.

Strengths:

- CUL is progressive and innovative
- Strong international reputation as a research library and academic institution
- Authority/leadership in the academic community and professional associations
- CUL has the resources and drive to collect deeply (e.g. purchasing back files from major academic publishers)
- Prioritizes users' needs and is driven to provide highly accessible and usable e-book collections
- Strong collaborative relationships with partner institutions and consortiums
- Location in New York provides opportunities to develop relationships with large publishers located in the city
- Faculty have international reputations and are leaders in their respective fields
- Columbia is focused on graduate studies, and students produce high quality research through thesis and dissertation projects

Weaknesses:

- There isn't a standardized definition of the term "e-book" across campus which leads to confusion in terms of expectations and functionality
- There isn't an e-books workflow that specifically addresses their complex management needs and challenges
- Terms of licensing agreements are not in a location that is easily discoverable
- There is not a clear understanding of how/why CUL patrons use e-books for research, teaching, and learning purposes
- A large amount of staff time is spent tracking down content in e-book collections (e.g. broken URLs, items pulled from databases by vendors) instead of evaluating the content
- Not enough staff/time/budget to find solutions to vendor generated metadata problems, e-book workflow issues, etc.

Opportunities:

- Free social media initiatives make it possible to create metadata based on "the wisdom of the crowd" (e.g. crowdsourcing)
- Consortiums and collaborative relationships are viewed as the most effective means to negotiate license agreements, prices, fix MARC records
- Industry trends are moving towards open access and self-publishing
- E-books are gaining a reputation as a new and innovative research and reference tool, not just digital versions of print monographs
- New technologies are being developed to work around DRM issues
- The Portico preservation strategy is viewed as a viable model, but has not been tested in a practical setting

Threats:

- Publishers are anxious to change licensing models because of piracy concerns
- Online book lending initiatives (e.g. Amazon Lending Program) may change how patrons interact with libraries
- The e-book landscape evolves so rapidly that it is difficult to predict what the challenges will be in a year from now
- The e-book market is focused on consumer needs, and reasons regarding how/why e-books are used in academic environments are largely undocumented
- There isn't a national strategy regarding preservation for e-books
- Libraries have no legal rights to preserve e-book content because of clauses in licensing agreements

A.2. Survey of the External E-Book Landscape

A.2.1. E-Book Landscape in North America (Literature Review)

A literature review was conducted to examine the e-book landscape in the academic community and commercial market in order to examine trends that impact e-book acquisition, discovery, access, and preservation. This project establishes a theoretical and methodological foundation for the E-Book Program Development Study and contextualizes the findings within the existing tradition of scholarship in the library profession.

The Definition of an Electronic Book (E-Book)

In studies conducted by Levine-Clark (2006), Herson (2007), and Shelburne (2009) findings indicate that there is no clear definition of the term *e-book*, and a small but significant percentage of sample groups were not sure what an e-book was (Staiger, 2012). For instance, Levine-Clark posed several open-ended questions to respondents, and many “confused e-book with e-journal or e-reserve” (Staiger, 2012, p. 356). Herson also found that students do not distinguish between types of sources, but are only concerned with whether a source is available in print or electronic formats (Herson et al., 2007). Staiger (2012) stated that this “lack of knowledge has implications for the quality of users’ engagement with the contents of e-books” (p. 356). However, the ability to clearly define what an e-book means at a given institution is linked with the general acceptance of the format by the user community.

The Oxford Companion to the Book provides a definition of the term *e-book* that has been adopted by a number of academic institutions. It defines the tool as a book-length publication in digital form, consisting of text, images, or both, and produced on, published through, and readable on computers or other electronic devices (Gardiner & Musto, 2010, p. 164). Also, it can exist in born digital form without a print equivalent (Gardiner & Musto, 2010).

E-Book Life Cycle Management

In the past decade, the development of technologies like e-book readers, mobile devices, and tablets has created a demand for content in a variety of formats. This demand has led to significant growth in the number of e-books purchased by academic libraries. However, e-books are a research, teaching, and learning tool that have different management needs than print monographs or e-journals. Currently, libraries are struggling with “how to manage and provide access to all of these new resources that do not fit neatly into any pre-existing workflow” (Beisler & Kurt, 2012, p. 96).

In many cases, e-book challenges extend beyond libraries’ jurisdictions. For instance, the “multitude of different e-book readers, formats, access platforms, and licenses makes it difficult for libraries to establish set procedures for acquiring and managing e-books” (Beisler & Kurt, 2012, p. 96). Also, there are vast inconsistencies within the e-book publishing industry that place limits on how libraries are able to provide access (Beisler & Kurt, 2012). Due to these

complexities, it is essential for librarians to understand the general e-book landscape, and how their institution fits into that context, in order to properly inform workflows and collection management policies at a given institution.

In an article published in *Against the Grain*, Carolyn Morris states that the first step to creating usable workflows is to acknowledge that e-books are vastly different from their print counterparts. The issues surrounding them are more complex, publishers and vendors supply them in a different way, and it is unwise to minimize the differences simply to preserve existing workflows (Morris, 2008). As new formats emerge, libraries must adjust policies and procedures to reflect changes (Beisler & Kurt, 2012). For instance, e-book workflows can be informed by print book models but ultimately, “differences in format require a new stream for processing, and this requires the library to create new procedures for handling e-books, from evaluation to activation and most stops in between” (Morris & Sibert, 2011, p. 110).

Developing a new workflow from the ground up is a daunting process and to date, there has been little published about e-book workflows, strategies, or procedures. Based on this research gap, it is difficult to determine what work has taken place at various academic libraries, and whether or not experimentation has been successful. In the absence of an “agreed-upon overarching framework of the processes associated with the management of e-books in academic libraries, it is difficult to compare and contrast the findings from studies or develop clear guidelines for practice” (Vasileiou, Rowley & Hartley, 2012, p. 283).

To address this research gap, the University of Nevada, Reno Libraries created a cross-departmental task force and built an e-book workflow. Their goal was to create an efficient and effective workflow that provided users with seamless service (Beisler & Kurt, 2012). It included the point of inquiry, acquisition, access, and disposition. The decision was made to build a workflow from the ground up in order to tackle traditional departmental divisions. Findings indicated that communication between departments was the largest obstacle that affected success rates of e-book workflows. However, they also discovered that developing a workflow became an opportunity for “departments and individuals to work closely together toward a common and worthy goal” (Beisler & Kurt, 2012, p. 109). The success of the project was due to cross-departmental collaboration and the ability to adapt tools on hand to the needs of the e-book workflow. For instance, the task force used SharePoint and the libraries’ electronic resource management ILS module (Innovative Interfaces Inc.’s ERM module) to promote communication at each phase of the workflow (Beisler & Kurt, 2012). The results indicated that workflows are necessary to inform libraries about e-book models that are user-centric and most suited to the needs of a user community (Beisler & Kurt, 2012).

Selection and Acquisition

The selection of e-books is a complicated process that is driven by institutional requirements for the acquisition of e-books. To learn more about this process, Soules (2009) conducted an Ebrary

librarians' survey examining factors that informed e-book purchases. The findings revealed that integration with other resources, download capability, the ability to support multiple file types, integration with a content management system or the institutional repository, and PDF formats ranked as important in e-book acquisitions (Soules, 2009).

Other researchers have stated that because of the complicated e-book landscape, identifying factors that contribute to informed e-book purchases is not enough. Blummer and Kenton (2012) recommend that libraries select a team of individuals to direct all e-book acquisitions, purchase processes, and initiatives. This model was put into place at the University of Worcester, and their e-book project group is composed of subject librarians, collections specialists, the electronic resources librarian, and library assistants (Blummer & Kenton, 2012).

A similar committee was established at the Indira Ghandi National Open University and is tasked with creating operating guidelines, principles, and potential strategies (Tripathi & Jeevan, 2008). The group also negotiates trial access for teachers and researchers as a means to evaluate prospective titles, makes decisions regarding subscription models, examines the long-term relevance of the content, and evaluates selected vendors (Tripathi & Jeevan, 2008).

At the University of Dublin, a small working group investigated e-book purchases and worked with academic units in the selection process. Main criteria for selection included ease of use, off-site access, multiple simultaneous users, and print and/or download options. In addition, the group invited prospective vendors to the Library to view demonstrations of platforms and evaluate their overall value to the institution (Blummer & Kenton, 2012).

Based on the results of a literature review of collection management practices from 2005-2012, Blummer and Kenton (2012) developed guidelines for the acquisitions of e-books in academic institutions. Their nine recommendations are as follows:

- Identify e-book acquisition staff;
- Partner with academic departments and especially distance education faculty in selecting titles;
- Provide a trial access to evaluate platforms;
- Consider the value of e-reference titles;
- Highlight currency in e-book packages;
- Focus on platform features such as ease of use and availability of specific features including the index, highlighting text, viewing large images, pasting, printing, and a variety of downloading options;
- Recognize the need for access models that allow simultaneous access with multiple users;
- Create a spreadsheet to differentiate among packages in the evaluation process;
- Understand licensing terms. (p. 76)

Print and Electronic Formats

The Library Journal's e-book survey reported a 93 percent increase in e-book collections among academic libraries since 2012. The survey also found that libraries anticipate e-book spending to comprise 20 percent of their budgets within five years (Blummer & Kenton, 2012). However, there are divided opinions on the subject of print versus e-book formats. Currently, many academic libraries hold the opinion that e-books and e-textbooks should coexist with print textbooks rather than replace them (Armstrong & Lonsdale, 2009). In many cases, the e-version is still viewed as a supplement to print copies.

Print and electronic texts are two different tools used for different reasons, and MIT suggests that libraries should collect content in both formats whenever possible. However, prior to purchasing an electronic version, there should be confirmation that it contains the same content available in print editions (MIT, 2012). The E-Book Strategic Plan Task Force at Yale University Library also encourages the acquisition of monographs in both print and electronic formats. This is because print books fulfill the need to collect, organize, and preserve knowledge while e-books support research, teaching, and learning initiatives (Yale University Library, 2013, p. 7).

A study by JISC (2012) found that e-books are not currently replacing the demand for print books despite the fact that e-journals have replaced back copies of printed journals (JISC, 2012). Another study conducted by the E-Books Strategic Plan Task Force at Yale University Library (2013) found instances that the adoption of e-books across library systems is uneven. This is often related to the fact that print versions are usually issued several months to a year before electronic versions. In many cases, the library already has the print books and so is reluctant to duplicate the purchase (Yale University Library, 2013). Because of uneven adoption rates and the unique needs of user communities, a survey by Ashcroft (2011) indicated that "49 percent of respondents indicated that usage statistics are the most important driver in e-book purchasing decisions" (Ashcroft, 2011, p. 401).

After conducting a number of focus groups, the JISC National E-books Observatory Project found that in many cases, the printed book is still the preferred format. This preference was linked to the physicality of printed books, a belief that printed books facilitate greater concentration, a belief that it is easier to scan a printed book, and the expectation that a printed page is easier to annotate, highlight, and make notes from (JISC, 2012). The study concluded that in most cases, "these reasons arise as a result of people thinking that using e-books is about making a choice *not* to use a printed book" (JISC, 2012, p. 44).

However, it is important to note that usage trends and beliefs linked to e-books vary across disciplines. In the sciences, electronic materials are heavily used because of the convenience and speed of locating information. However, users do not often use materials that are more than three years old. In a case like this, librarians can create a customized e-book plan to best suit users'

needs (Schell, 2011). For instance, librarians could create subject based e-book lists updated annually to highlight current content (Schell, 2011).

Across the academic community, and even within the publishing industry, there is the general belief that print formats and e-books are not in an either-or competition. The two formats “already coexist with each answering to different purposes and learning style” (Staiger, 2012, p. 360). However, there is a constant increase in the number of born digital books and journals being published. Since these items do not have a print equivalent, libraries may not always have the option of selecting a format (JISC, 2012).

Purchases versus Subscription Licenses

When examining the issues of purchase versus subscription, there is no clear cut preference across the library profession. Both are seen to have advantages and disadvantages, and the decision to purchase or subscribe to content often comes down to institutional needs. However, there is widespread agreement that decisions come down to stipulations in licensing agreements such as ensuring there are provisions for multiple access (preferably unlimited) and flexibility (Armstrong & Lonsdale, 2009).

The most important factor to take into account during any contract negotiation is users’ needs. It is important to keep the e-book priorities of students and faculty at the heart of licensing decisions (Blummer & Kenton, 2012). For instance, at the University of Liverpool Library, e-books are purchased directly from the publisher to avoid restrictive content and excessive digital rights management issues (Blummer & Kenton, 2012).

One of the largest issues facing academic libraries is that it is difficult to determine which titles or packages were purchased and which are accessed through subscriptions. This lack of information creates significant challenges when librarians and staff try to determine how collections can be used. There need to be systems that allow for easy consultation and dissemination of licensing terms to ensure compliance and also understand how library resources can be used or shared (Armstrong & Lonsdale, 2009).

Bundles versus Title-by-Title Purchases

In 2009, High Wire Press conducted a survey of 138 academic libraries to examine preferences between bundle or title-by-title purchases. The findings indicated that while many prefer to select books on a title-by-title basis, the reality is that bundles offer better pricing models, save time in selection, acquisition, and processing, and offer titles that are not sold on an individual basis (Newman, 2009). Other studies have found that the cost-per-use rate for individually-selected titles is seventeen times higher than for titles purchased through aggregate packages (Staiger, 2012).

Although bundles are more attractive in terms of cost, librarians find that it is difficult to determine what titles are available in each package and to acquire appropriate metadata records (Blummer & Kenton, 2012). Because of the complexities involved, some academic institutions believe that e-book selection should be done by committees rather than individual selectors. For instance, at Yale University Library, e-book purchases are done using a tier system that dictates how decisions are made. In this system, the Director of Collection Development, the Assistant Director of Collection Development, the Collection Steering Committee (CSC), and the eBook Working Group organize the purchase of e-book content into the following three tiers:

1. Tier One: e-book packages that are negotiated and purchased with central funds;
2. Tier Two: e-book packages that are negotiated and coordinated centrally, but are funded through cross unit cost sharing;
3. Tier Three: e-book content that is purchased by individual selectors. (Yale University Library, 2013, p. 8)

This structure eliminates much of the confusion that occurs when individual selectors negotiate or select e-book packages on their own (Yale University Library, 2013). Also, it allows Yale University Library subject specialists to “negotiate directly with publishers for bits and pieces of package deals that could be purchased collectively with less effort and deeper discounting than an individual selector can achieve” (Yale University Library, 2013, p. 8). Essentially, the tier system allows the Library to leverage its collective buying power to “secure advantageous pricing, a more strategic and predictable internal workflow, and the reduction of duplication across electronic platforms” (Yale University Library, 2013, p. 9). Collective purchasing of e-books also allows librarians at Yale to document their approval or disapproval of certain products in the market place (Yale University Library, 2013).

At the end of the day, the acquisition of packages and individual titles should be done in accordance with users’ needs. The MIT Statement of Scholarly E-Book Principles reflects this sentiment and states that “pricing models [should] allow institutions to purchase packages tailored to the needs of their local communities, allow for the selection of individual titles, and that do not require minimum purchases” (MIT, 2012, p. 1).

Metadata Records

Across the board, academic libraries agree that high-quality catalogue records provide the most effective means of discovery and access. In many cases, e-book metadata records are supplied by vendors. Findings from the JISC National E-books Observatory Project indicate that there are two central concerns from libraries in regards to vendor generated metadata. The first is the poor quality of MARC records, and the second is inappropriate ISBNs (Armstrong & Lonsdale, 2009). A study by Mincic-Obradovic (2009) found the other challenges include missing URLs and not indicating how an e-book differs from its print counterpart (Mincic-Obradovic, 2009).

At Yale University Library, the E-Book Strategic Plan Task Force surveyed Cornell University, Duke University, Princeton University, Stanford University, and the University of Michigan to identify key metadata challenges. Findings indicated that obtaining a perfect MARC record is difficult. There is also differences of opinion regarding whether e-books should have MARC records equivalent in detail to their print counterparts, or whether a poor record is better than no record at all (Yale University Library, 2013).

One solution that has been presented within the academic community is to add a MARC 856 field to an equivalent print record (Blummer & Kenton, 2012). However, due to the growth of e-book holdings at most libraries, it is strongly recommended that a separate record is created for each e-book (Blummer & Kenton, 2012). For example, at the University of Worcester's Information and Learning Services, each e-book title is catalogued individually to improve user access to their e-book and e-textbook materials (Blummer & Kenton, 2012). Also, the University of Surrey Library creates separate records for e-books in an effort to recognize the resource as an independent [tool]...with different functionality than print formats (Blummer & Kenton, 2012).

At the J.N. Desmarais Library of Laurentian University, a study was done to assess the importance of metadata records in discovery and access. Findings indicated that creating a metadata record for each e-book increased usage rates, particularly among grad students and faculty (Lamothe, 2013). In some cases, a metadata record doubled usage rates. However, the amount of time required to catalogue e-books presented challenges, and was largely related to the number of e-books purchased at one time, as well as the availability of preexisting MARC records (Lamothe, 2013). For instance, e-books purchased individually could be immediately catalogued, but cataloguing bundled titles could take anywhere from one week to six months (Lamothe, 2013).

A partnership between the University of Illinois at Chicago's University Library and the Center for Library Initiatives (CLI) developed a consortial review process aimed to improve MARC records provided by Ingram for their Springer e-book collection (Marin and Mundle, 2010). The group identified three central challenges including access issues, load issues, and record quality issues (Marin and Mundle, 2010). To remedy these problems, the group used MarcEdit, an "open source MARC batch editing tool that permits manipulation of the data to promote the identification and correction of record errors" (Blummer & Kenton, 2012, p. 80). The results of the study indicated that joint efforts from the consortial review and the vendor remained the most productive way to generate usable bibliographic records (Marin and Mundle, 2010).

Based on the results of a literature review of collection management practices from 2005-2012, Blummer and Kenton (2012) developed nine best practices for cataloging e-books.

1. Catalog records in library's integrated library system to improve findability;
2. Create separate catalog records for e-book titles, rather than adding MARC 856 field to print record;

3. Use full MARC format and add URLs for e-book access;
4. Consider the popularity of vendor-supplied records;
5. Recognize the need to edit vendor records to ensure that they meet local cataloging standards;
6. Consider the capability of the ILS for bulk importing, indexing, and deleting;
7. Identify the tools available for editing vendor e-book records to support collection analysis and searching in next-generation library systems as well as discovery tools;
8. Encourage vendors adopt the e-monograph guidelines issued by the PCC Provider Neutral E-Monograph Record Task Force for vendor-supplied records;
9. Weigh the cost of upgrading vendor records rather than creating original records for e-books.

(Blummer & Kenton, 2012, p. 82)

Library Catalogue and Resources

In 2009, a focus group report by Christ Armstrong and Ray Lonsdale stated that “there is a bewildering variety of e-content, and proliferation of ways to get to it. Users don’t know how to get what they want. Libraries face a big challenge in providing clear access routes to e-content” (Armstrong & Lonsdale, 2009, 28). Their findings indicated that most students locate e-books through the OPAC, so it is useful for e-book collections to be integrated into the catalogue. This way, students can locate books and e-books on a single interface (Armstrong & Lonsdale, 2009). In addition, “adding respective links to the e-books within the catalogue will ensure that, once a specific e-book has been discovered, a learner can select the link and gain immediate access to the e-book within the collection” (Armstrong & Lonsdale, 2009, p. 39).

Studies by Newman (2009), Nariani (2009), and Staiger (2012) indicate that the most common way users discover e-books is through the library catalog. For instance, Newman observed that the “traditional sources of book discovery continue to be important for e-books as well” (2009, p. 5). Essentially, users discover e-books through the library catalog and Internet searches. Nariani also found that catalogued e-books were used more often than those that had been promoted by email. Staiger reported that “the library catalog was by a wide margin the primary place where every category of respondents came upon e-books. In the case of respondents from the humanities or social sciences, well over 50 percent learned of e-books either from the library catalog or homepage” (2012, p. 356).

Librarians at the J.N. Desmarais Library of Laurentian University conducted a quantitative and systematic study of online e-book usage and discovered that in addition to the library catalogue, students accessed e-book collections from links off the Library’s website (Lamothe, 2013). The findings indicate that “library websites are critical e-book access points, and for the majority of undergrads, the primary e-book discovery tool” (Lamothe, 2013, para. 3).

While the library is an obvious source for increasing students' awareness of e-book collections, findings from a literature review conducted by Blummer and Kenton (2012) stated that "faculty [are] a valuable but underused source for increasing students' awareness of e-books in library collections" (p. 88). The ability of faculty and librarians to integrate e-books into the curriculum impact usage rates in a positive way. Armstrong and Lonsdale (2009) also discovered that one of the most significant ways that faculty can promote these resources is by providing links to relevant sections of e-book collections from an instructional platform (Armstrong & Lonsdale, 2009).

Promotion should not stop with linking to e-books from instructional platforms. There also needs to be standardized instruction that teaches students how to use e-book collections. Blummer and Kenton (2012) found that over 65 percent of students who use libraries' e-books recall learning about them in library instructional sessions (p. 90). Findings from a literature review by Ashcroft (2011) also suggest that librarians play an important role in raising awareness of e-book holdings. In the first place, users "need to know that their library provides e-books, then [they must know] how to find them" (p. 399).

At the end of the day, "awareness is largely dependent on local circumstances, most prominently but not exclusively such as the degree to which e-books have been promoted at a given institution" (Staiger, 2012, p. 356). Libraries should develop innovative and creative strategies to market e-book collections to targeted user groups. For instance, at the University College of Dublin, librarians email academics usage statistics as well as new e-book titles (Blummer and Kenton, 2012). In addition, Ashcroft (2011) discovered that promotional methods include "social networking applications, subject specific bookmarks advertising e-books, putting stickers on hard copy to advertise electronic availability, and placing dummy e-books on the shelf as a prompt" (p. 400).

Based on the results of a literature review, Blummer and Kenton (2012) developed a number of strategies to promote e-books to targeted user groups. Their eight suggestions are as follows:

1. Market e-books on the library's website through listings with databases, LibGuides, and on subject pages: host an e-book forum; provide a definition of e-book; highlight new purchases and freely available collections;
2. Include e-books in the library's OPAC and have a limit function to search e-books;
3. Involve faculty in e-book promotional efforts;
4. Support faculty's use of e-books in teaching, especially for distance education;
5. Provide instruction in using e-books, such as navigating platforms accessing features;
6. Send target e-mails to specific user groups;
7. Utilize social networking tools such as Facebook and blogs;
8. Make e-book marketing ongoing with a formal strategy. (Blummer and Kenton, 2012, p. 91)

Usage Trends in Academic Environments

Determining how e-books are used for academic purposes is a complex issue. It is not enough to understand who uses these resources and how they are used; librarians must also consider why e-books are or are not used. Unfortunately, the latter has not been widely researched or discussed in the professional community.

Over the past several years, a number of studies were conducted to determine the benefits and challenges users associated with e-book collections. Results from Beisler and Kurt (2012), Ashcroft (2011), Armstrong and Lonsdale (2009), and the ARL SPEC Kit 313 (year) all suggest that the main benefits include twenty-four hour access to materials, remote access, and the ability of multiple users to use one resource at the same time. Again, the challenges listed in all four studies are similar and signify complex problems that are often linked to the policies and practices of publishers and vendors. They include DRM, platform design, and file format compatibility with various e-readers.

In terms of user groups, doctoral students typically exhibit the strongest relationship with e-book usage (Lamothe, 2013). As one graduate student explained, “the advantage of e-books is immediate access to chapters in edited research volumes. Unlike journal articles, these chapters are rarely available as PDFs from publishers or in databases” (Staiger, 2012, p. 359). Within the undergraduate population, e-book usage is low; however, overall faculty demonstrated the weakest relationship with e-book usage (Lamothe, 2013). Staiger (2012) described faculty’s usage of e-books as task oriented – they search for quick information or use it to find a print version for extended research (2012).

A literature review by Staiger (2012) compared the results of two dozen studies regarding e-book usage by members of the academic community. Findings suggested that “academic users typically search e-books for discrete bits of information, a behavior summed up by the formula ‘use rather than read’” (p. 355). In general, members of the academic community do not immerse themselves in e-books for extended periods of time to examine entire arguments. Instead, they view e-books as “convenient sources from which to extract information for their scholarly endeavors” (p. 357). Essentially, e-books provide a means for power browsing. They allow users to preview a book without leaving their work stations, and then locate the print copy if the information is relevant to their studies (p. 358). A literature review by Ashcroft (2011) uncovered similar trends. Statistics showed that on average, “53.5 percent of students and 58.6 teachers dipped in and out of several chapters, whereas very low percentages read the whole book – 5.5 percent of students and 7.1 percent of teachers” (p. 401).

To understand how e-books are used, the University of Liverpool Library partnered with Springer and conducted a series of online surveys and focus groups. Results indicated that there was an 88 percent increase in the number of e-book chapters downloaded between June 2009 and July 2010 (Bucknell, 2010). The study went on to compare e-book usage with e-journal article

usage and found that the use of Springer e-journals increased significantly between 2008 and 2009, and suggests that having access to e-books on the same platform as e-journals does have an inflationary effect on the usage of e-journals (Bucknell, 2010). The figures also show that the number of unused e-book titles diminished each year, with older titles continuing to attract significant usage (Bucknell, 2010).

It is important to note that evidence suggests academic users expect the same functionality from e-books that they experience with e-journals. For instance, they want to download PDFs and expect that an e-book allows for multiple users simultaneously. When faculty or students cannot access an e-book because the limit on users is reached, they become frustrated and are often unaware of licensing limits (Ashcroft, 2011). Although there are obvious limits to the number of print books a library would purchase, it seems that “because multiple ease of access to the Internet, limits to accessing e-books are not recognized” (Ashcroft, 2011, p. 402).

To help user communities navigate the complex e-book landscape, librarians (particularly those who work in reference departments) should become familiar with a variety of e-readers and tablets (Buckley & Johnson, 2013). In addition, providing clearly written guides on downloading processes and functionality are invaluable to students, faculty, and library staff (Buckley & Johnson, 2013).

Functionality

As digital technologies continue to provide a wide variety of options in terms of information access, particularly in the commercial market, patrons expect to find e-books in academic libraries that support research, teaching, and learning activities. In general, users expect to view e-books on a variety of hardware platforms including workstations, laptops, dedicated readers, and mobile phones (Ashcroft, 2011). Today, “users want to be able to access the same e-books but at their convenience on a variety of devices” (Ashcroft, 2011, p. 401).

The fact remains that it is difficult for libraries to lend e-books. This is due to the fact that none of the publishers or vendors involved are working together to find solutions (Bradford, 2013). At this time, “the e-reader makers, library lending software developers, and the publishers are all working at odds” (Bradford, 2013, para. 4). One of the major challenges facing libraries is that the e-book market has not reached maturity, and there are “many formats competing for prime time, including Adobe PDF, Microsoft Reader, eReader, Mobipocket Reader, EPUB, Kindle, and iPad” (pcmag.com, n.d., para. 3). Currently, library users prefer e-books in PDF format, but this may change as technology continues to evolve (Newman, 2009). In all likelihood, e-books would have to be “compatible with a gamut of devices, in other words rendered independent of particular platforms, before they would present libraries with a feasible channel for provisioning materials” (Staiger, 2012, p. 363).

Currently, many library users are not confident that e-books provide desired features required for research, teaching, and learning. For instance, navigating between sections or chapters is

perceived as awkward when compared with maneuvering through a print book (Staiger, 2012). Also, features such as printing, copying, or saving e-book sections are ranked by users as more important than searchability (Staiger, 2012). Undergraduate and graduate students also look for indexes, a table of contents, and the full text search tool available in e-books (Blummer and Kenton, 2012). Also, the ability to highlight and annotate texts or follow links to other sources were of value (Blummer and Kenton, 2012).

In most cases, “users expect the same kind of liquidity that they have come to largely enjoyed with articles from e-journals: the ability to download them on whatever device they choose and print as much as they want” (Staiger, 2012, 359). When they encounter obstacles in these areas, they are frustrated. The vast majority of these challenges are not inherent to e-books themselves. Rather, they are the result of restrictions imposed by publishers and vendors (Staiger, 2012). This situation leaves libraries between a rock and a hard place as they address concerns from users without having the ability to remedy the situation.

During the 2008/9 academic year, Penn State University Libraries partnered with Sony Electronics to study the utility of e-books in research library collections. In particular, they investigated “the effect of reading devices on teaching, learning, and reading; the utility of such reading devices for individuals needing adaptive technologies; and how licensed and locally created digital content could be repurposed for use on portable reading devices” (Behler, 2011, p. 89). Results indicated that users want portability, E-Ink grayscale technology, and uni-function devices that do not distract from the process of reading (Behler, 2011). Criticisms of e-books included slow refresh time when turning pages and a lack of features such as annotation and highlighting capabilities (Behler, 2011). Many users also indicated that it is important for them to use content in any way they want or need to (Behler, 2011).

At the University of Nevada, Reno, librarians connected with users by providing resources in requested formats, and also offered users (including library staff) the chance to experiment with different e-readers (Beisler & Kurt, 2012). A cross-departmental team designed an “E-reader Bar” and invited patrons to try a variety of devices loaded with e-book content (Beisler & Kurt, 2012). Feedback indicated that “staff had benefited from having the chance to try different e-book readers and it made sense to give users the same opportunity” (Beisler & Kurt, 2012, p. 109).

At the University of North Carolina (UNC) Libraries, a number of recommendations have been developed to accommodate tablets, e-readers, smartphones and other mobile devices. First, they select e-books in ePub, XHTML, and other XML-based formats over PDF because “the former are reflowable files developed for digital publishing that can adapt their presentation to the output device and therefore typically easily download to and accurately display on a wide range of mobile devices” (University of North Carolina Libraries, 2012, p. 2-3). In contrast, PDF files “are not easily reflowable, do not adapt well to various sized displays and mobile devices, and therefore are difficult if not impossible to view on small screens that come with some e-readers

and smartphones” (University of North Carolina Libraries, 2012, p. 2-3). In cases where only PDF files are available, UNC recommends text-based Adobe PDF formats because they allow for “easy highlighting (copy and paste), keyword searching, improved downloading, and better support for disability access” (University of North Carolina Libraries, 2012, p. 2-3).

It is important to note that companies are creating new technologies to combat the current access issues libraries face due to restrictions enforced by publishers and vendors. 3M, the company who invented Cloud E-Book lending systems for smartphones and tablets, has developed its own reader for libraries. It is “designed specifically for libraries to lend out to patrons with its easy system. Book lovers can choose the e-books they’d like to read, then get the 3M Reader from the librarian, scan their barcode, and be done” (Bradford, 2013, para. 11). The only hitch is that most libraries currently use Overdrive and have not adopted 3M’s system (Bradford, 2013).

Preservation

The introduction of e-book formats to library collections has caused dilemmas in terms of preservation and stewardship. For instance, the National Digital Stewardship Alliance is working to “identify content at risk of loss, develop and adopt digital preservation standards, share tools and services, support innovation of practice and research, and promote national outreach for digital preservation” (Billington, 2013, p. 71). While there are issues including software and hardware obsolescence and storage space, one of the central issues is the fact that libraries do not have the legal rights to preserve e-books. Essentially, licensing agreements provide temporary access to e-book collections and do not allow libraries to own a copy of each individual file. As Yale University Library stated,

Traditionally, the Library would procure a print book in support of activities of members of the university and then preserve that book for future users. We could do this because we owned the book, owned the device used to store the book (the bookshelf) and employed staff to ensure the maintenance of the book for future use. Now, when the Library procures an electronic book in support of such activity there is no mechanism for the Library to preserve that eBook for future users (Yale University Library, 2013, p. 5-6).

Digital Rights Management (DRM) restrictions often prevent libraries from downloading or printing copies of e-books for archival purposes (Yale University Library, 2013). Currently, the only way in which libraries could preserve e-books is if “publishers were prepared to sell the Library digital eBook files with which the Library could do whatever it wanted. In the current market, publishers are not prepared to sell digital eBook files with no strings attached” (Yale University Library, 2013, p. 6).

In terms of libraries themselves, even if publishers were prepared to sell e-books, the majority do not have adequate infrastructure to house them. At this time, most do not have a “robust

information technology infrastructure (institutional repository) in which to store eBook files, [or] have a plan in place to migrate eBook files (or any other kind of digital files) from the current generation technology platform to the next” (Yale University Library, 2013, p. 6).

In regards to preservation concerns, Cornell University, Duke University, Princeton University, Stanford University, and the University of Michigan face similar challenges. When surveyed by Yale University Library, they stated that preservation is addressed “in their license negotiations with vendors” (Yale University Library, 2013, p. 15). In addition, they rely on third party systems like Portico and LOCKSS, as well as local repositories such as the Stanford Digital Repository (Yale University Library, 2013). The institutions stated that they are comfortable with the lack of e-book preservation in cases where there is a print edition in the collection. However, there are growing concerns surrounding dynamic e-book content that has no print equivalent (Yale University Library, 2013).

In reality, there is no e-book solution that “simultaneously meets both the ‘current use’ and ‘future use’ requirements” (Yale University Library, 2013, p. 7). In some cases, it may make economic sense for libraries to purchase an electronic format without thinking about long-term access (Yale University Library, 2013). In other cases, it may be appropriate to purchase titles regardless of current user demand in the hopes of preserving the content (Yale University Library, 2013).

Evaluation of Management Practices

In order to properly evaluate a workflow, it is essential to ensure that information is communicated and gathered from all departments and staff involved in the process. Buckley and Johnson (2013) recommend storing all documentation for the workflow in a shared location and revising it as needed. The keys to success include planning, communication, storing backups, and revisiting workflows to identify areas that require adjustment (Buckley & Johnson, 2013).

Also, it is essential to review and fully understand how users access and discover electronic resources. At the end of the day, e-book collections are meant to support research, teaching, and learning activities at academic institutions. The results of a literature review by Staiger (2012) indicated that “libraries, publishers, and content aggregators should be more responsive to how students gather and use information to complete classroom assignments (p. 361). Having a working understanding of how users interact with e-books provides insight into how existing initiatives meet information needs. At the University of Nevada, Reno, an evaluation of the e-book workflow revealed that there should be a higher focus on discover and user experience (Beisler & Kurt, 2012). In response, “a number of existing staff have been shifted over to a new department called Design and Discovery. This department came from a need to make discovery of resources and the online user experience a priority at the UNR Libraries” (Beisler & Kurt, 2012).

Finally, in order to properly assess usage trends, libraries need accurate and usable statistics from publishers and vendors in order to assess e-book collections. The JISC National E-books Observatory Project found that statistics provided by publishers and aggregators vary in quality. In many cases,

it is difficult for librarians to collect meaningful statistics from collections and want publishers and aggregators to send this data to them. Librarians want more time to reflect on the process of collection management and often have no time to collect meaningful statistics. In addition, qualitative studies should supplement quantitative analysis to provide deeper understanding into the way collections are discovered and used (Armstrong & Lonsdale, 2009 page v).

E-book providers need to adopt “a standard metric for reporting data on searches, viewings, and downloads, so that libraries can have a clearer sense of how the resources in which they are investing their funds are being used to facilitate comparisons among different e-book packages” (Staiger, 2012, p. 361). The review of accurate statistics helps publishers and vendors test assumptions about what librarians and users want and need from e-books (Newman, 2009).

Collaborative E-Book Management Models

The current e-book landscape is complex and in a state of constant flux. Libraries face challenges negotiating costs and licensing agreements, working with vendor generated MARC records, and discussing preservation models. In the current environment, many academic libraries form consortiums to pool resources and find solutions to pressing issues.

A study by Stachokas (2012) found that the “greatest focuses on consortia in 2009 were renegotiating licenses for electronic resources and budget management” (p. 144). There is a general acknowledgement in the library community that

sharing e-books through consortial arrangements can be a highly cost-effective way to introduce them to a collection. Since the management of the contract and invoicing are typically handled by the lead faculty in the consortium, the burden of training local staff with new skill sets is reduced. Often, the downloading of MARC records to the OPAC is handled centrally as well, further relieving consortium members of added work. In addition to the obvious benefits of competitive pricing through consortia, group selection of title-by-title e-books can create a diverse and rich collection. (Stachokas, 2012, p. 144)

In the future, consortia will continue to grow in importance because of their ability to set up advantageous terms with vendors, provide training in the area of electronic resource management, and take on professional advocacy roles (Stachokas, 2012). However, libraries should not limit membership to other academic libraries, but should also look for opportunities to include publishers and vendors (Stachokas, 2012). E-book management is a complex problem

and solutions will depend on collaboration from all members of the equation. In many cases, “librarians feel unconsulted and believe that it is necessary for publishers and aggregators to work more closely with them” (Beisler & Kurt, 2012, p. 98). The perspective that librarians, publishers, and vendors bring to the table may create new solutions to communal discovery, access, and preservation challenges.

An example of effective collaborative working relationships is evident in the Triangle Research Library Network Consortium (TRLN), which is composed of Duke University, North Carolina Central University, North Carolina State University, and the University of North Carolina at Chapel Hill. The central mission is to “marshal the financial, human, and information resources of their research libraries through cooperative efforts in order to create a rich and unparalleled knowledge environment that furthers the universities’ teaching, research, and service missions” (Triangle Research Libraries Network, 2013, para. 1). The goal is to move TRLN libraries and partner publishers to a decidedly electronic environment for materials that improve support for instruction and research (Triangle Research Libraries Network, 2013). This goal is achieved by working with “innovative and flexible publishers to expand library collections cooperation from print to e-books within a win-win context” (TRLN, 2013, p. 1).

Future Trends

Currently, the e-book landscape does not have universal standards that promote discovery and accessibility. E-book library lending is an alienating process; there needs to be a streamlined process for every device and publishers need to understand the technical side of e-book lending to alleviate anxieties (Bradford, 2013). One of the reasons this is not happening is because publishers are “driven by a fear of piracy, just as the music industry was and the movie/TV industry is now” (Bradford, 2013, para. 24). At BookExpo America 2013, American Library Association President Maureen Sullivan said that the e-book dilemma is a “classic example of disruptive innovation. It causes a lot of misunderstanding, it brings fears to light. When we experience disruptive innovation, it’s much more effective to think not ‘either/or’ but ‘and’” (Bradford, 2013, para. 26).

Greco and Osman (2013) also describe e-books and e-readers as a disruptive technology. While margins are higher on a digital book than a print book, publishers also believe that every e-book purchased is a print book that was not purchased (2013). “While some analysts argue that e-books do not greatly affect print unit sales, our research indicates the opposite. Between 2008 and 2015, [we] project that education textbooks will decline by 69.7 percent” (Greco & Osman, 2013, p. 456).

However, there are others who argue that the availability of e-books in libraries can benefit publishers by adding a free marketing and promotional component. For instance, there is “evidence that during periods of technological, social and economic change, people use libraries

more. With many bricks-and-mortar bookstores closing, publishers need new ways to ‘showroom’ their titles” (Feldman, Russell & Wolven, 2013, p. 18). Library readers are also heavy book buyers. One service that libraries could offer is in “connecting readers with authors. Libraries might offer to provide access to a publisher’s entire catalog...as a way of connecting readers with additional offerings which they may buy or request the library to purchase” (Feldman, Russell & Wolven, 2013, p. 18). Also, libraries offer readers advisory, a service that “stimulate interest in books through...recommendations. By expanding this service to the e-realm, libraries will strengthen their role of connecting readers with authors and books they might otherwise miss” (Feldman, Russell & Wolven, 2013, p. 18).

There are other forces acting on the e-book landscape that will have an impact on creation and sales over the coming years. For instance, “the open access movement will not replace for-profit vendors, but it will help to ensure wider access to information and play at least a small role in keeping overall costs down” (Stachokas, 2012, p. 145). The rapid growth of self-publishing is also likely to provide new options in the way that libraries acquire books. As Feldman, Russell, and Wolven (2013) reported,

a small group of libraries have already cut out the middle man...and maintain their own e-book servers. The rapid growth of self-publishing is bound to have some impact on library collections. The perception that self-publishing is merely a vanity press under a different name is quickly eroding. New reader opportunities already are being developed by innovative entrepreneurs. By next year, we may be talking about the demise of the e-book – it having been replaced by some more-advanced technology that savvy readers will come to expect. Reading and technological advances associated with digital reading will move ahead at a breakneck pace. (Feldman, Russell & Wolven, 2013, p. 6)

An example of a revolutionary reading experience was launched in December 2012 by the New York Times. The project is entitled *Snowfall: The Avalanche at Tunnel Creek* and is described as a “beautiful reading experience through the use of a clean layout, interactive maps, inlaid videos and graphics that move as you scroll. The result is an online reading experience like no other” (Gardner, 2012, para. 2). Brantley (2013) stated that through this project, the New York Times has essentially reset the bar for interactive online narratives.

In addition, there are a number of trends on the horizon that may influence how patrons interact with libraries. For instance, last year Amazon launched its Kindle Lending Library, available to those customers who own a Kindle and have an Amazon Prime membership. The program allows Kindle owners to “choose from more than 350,000 books to borrow for free with no due dates, including over 100 current and former New York Times best sellers and all seven Harry Potter books” (Amazon, 2013, para. 1). It is

yet to be determined whether or not these developments make libraries more or less attractive to publishers and patrons.

While many believe that print books will not disappear in the coming decades, the growth of digital products will have a profound influence on the market and create a set of winners and losers (Greco & Osman, 2013). For instance, those at an advantage include publishers producing high-profit e-books, authors and agents who share in heightened royalties because their books are only available in digital form, retailers of e-books, and stockholders of publishing firms who own high-impact titles (Greco & Osman, 2013). The individuals at a significant disadvantage in the e-book market include shipping and transportation companies who ship books to distribution warehouses, distributors handling shipments and returns, surety bond companies writing policies for books imported to the United States, and companies in developing nations who print books sent to the United States (Greco & Osman, 2013).

References

- Amazon. (2013). *Kindle owners' lending library*. Retrieved from <http://www.amazon.com/gp/feature.html?ie=UTF8&docId=1000739811>
- Armstrong, C., & Lonsdale, R. (2009, November). *E-book collection management in UK university libraries: Focus groups report*. Information Automation Limited, Final Report. Retrieved from <http://observatory.jiscebooks.org/reports/e-book-collection-management-in-uk-university-libraries-focus-groups-report/>
- Ashcroft, L. (2011). Ebooks in libraries: an overview of the current situation. *Library Management*, 32(6), 398-407.
- Behler, A. (2011). The Penn State University Sony e-book reader project. In S. Polanka (Ed.), *No shelf required: E-books in libraries* (pp. 88-90). Chicago: American Library Association.
- Beisler, A., & Kurt, L. (2012). E-book workflow from inquiry to access: Facing the challenges to implementing e-book access at the University of Nevada, Reno. *Collaborative Librarianship*, 4(3), 96-116.
- Billington, J. (2013). Federal agency and federal library reports. In D. Bogart (Ed.), *Library and book trade almanac* (pp. 55 - 77). Medford: Information Today, Inc.
- Blummer, B., & Kenton, J. (2012). Best practices for integrating e-books in academic libraries: A literature review from 2005 to present. *Collection Management*, 37(2), 65-97.
- Bradford, K. T. (2013, June 15). Paper rules: Why borrowing an e-book from your library is so difficult. *Digital Trends*. Retrieved from <http://www.digitaltrends.com/mobile/e-book-library-lending-broken-difficult/>

- Brantley, P. (2013). Books going digital: Betwixt and between. In D. Bogart (Ed.), *Library and book trade almanac* (pp. 19-31). Medford: Information Today, Inc.
- Buckley, M., & Johnson, M. (2013). The why's and how's of integrating downloadable academic e-books. *Computers in Libraries*, 33(1), 9-32.
- Bucknell, T. (2010). The big-deal approach to acquiring e-books: A usage-based package. *Serials*, 23(2), 126-134.
- Columbia University Libraries and Information Services. (2010, January). *Strategic Plan, 2010 - 2013*. Retrieved from https://culis.columbia.edu/content/dam/staffweb/admin/strategicplanning/2010-2013/CULIS_Strategic_Plan_2010-20131.pdf
- Feldman, S., Russell, C., & Wolven, R. (2013). E-book business models for public libraries: A response to publisher-library differences. In D. Bogart (Ed.), *Library and Book Trade Almanac* (pp. 3-18). Medford: Information Today, Inc.
- Frederiksen, L., Cummings, J., Cummings, L., & Carroll, D. (2011). Ebooks and interlibrary loan: Licensed to fill? *Journal of Interlibrary Loan, Document Delivery & Electronic Reserves*, 21(3), 117-131.
- Gardiner, E., & Musto, R.G. (2010). The electronic book. In M. F. Suarez (Ed.), *The Oxford companion to the book* (pp. 164 – X). Oxford: Oxford University Press.
- Gardiner, E. (2012, December 20). The New York Times pushes the limits of online storytelling. *The Economist Group*. Retrieved from <http://www.economistgroup.com/leanback/leanback-reading/the-new-york-times-pushes-the-limits-of-online-storytelling/>
- Greco, A., & Osman, E. (2013). The competitive advantage of U.S. book publishers in the international market for books. In D. Bogart (Ed.), *Library and Book Trade Almanac* (pp. 451 – 470). Medford: Information Today, Inc.
- Hernon, P., Hopper, R., Leach, M., Saunders, L., Zhang, J. (2007). E-book use by students: Undergraduates in economics, literature, and nursing. *Journal of Academic Librarianship*, 33(1), 3-13.
- JISC. (2012). *Preparing for effective adoption and use of e-books in education*. Retrieved from <http://blog.observatory.jisc.ac.uk/techwatch-reports/ebooks-in-education/>
- Lamothe, A. (2013). Factors influencing the usage of an electronic book collection: Size of the e-book collection, the student population, and the faculty population. *College & Research Libraries*, 74(1), 39-59.

- Levine-Clark, M. (2007). Electronic books and the humanities: A survey at the University of Denver. *Collection Building*, 26(1), 11.
- Martin, K., & Mundle, K. (2010). Cataloging e-books and vendor records: A case study at the University of Illinois at Chicago. *Library Resources and Technical Services*, 54(4), 227-247.
- Mays, A. (2010, September). Workflows in paradise: E-books, acquisitions, and cataloging. *Against the Grain*, 22(4), 56-59.
- Mincic-Obradovic, K. (2009). Ten years on: E-books at the University of Auckland Library. *Serials*, 22(3), S23-S29.
- MIT. (2012, March 13). Draft Statement of Scholarly E-Book Principles. Report created XXX. Unpublished.
- Morris, C. (2008, September). Buying e-books: Does workflow work? Part I. *Against the Grain*, 20(4), p. 85.
- Morris, C. (2008, December). Buying e-books: Does workflow work? Part II. *Against the Grain*, 20(6), p. 77.
- Morris, C., & Sibert, L. (2011). Acquiring e-books. In S. Polanka (Ed.), *No shelf required* (pp. 95-124). Chicago: American Library Association.
- Nariani, R. (2009). E-books in the sciences: If we buy it will they use it? *Issues in Science and Technology Librarianship*. Retrieved from <http://www.istl.org/09-fall/article3.html>
- Newman, M. (2009). *HighWire Press 2009 Librarian E-book Survey*. Retrieved from <http://highwire.stanford.edu/PR/HighWireEBookSurvey2010.pdf>
- PCMAG.com. (n.d.). *Definition of an e-book*. Retrieved from <http://www.pcmag.com/encyclopedia/term/42214/e-book>
- Sawyer, S.K. (n.d.). Electronic books: Their definition, usage and role in libraries. Retrieved from <http://libres.curtin.edu.au/libres12n2/ebooks.htm>
- Schell, L. (2011). The academic library e-book. In S. Polanka (Ed.), *No shelf required: E-books in libraries* (pp. 75-87). Chicago: American Library Association.
- Shelburne, W. A. (2009). E-book usage in an academic library: User attitudes and behaviors. *Library Collections, Acquisitions, and Technical Services*, 33(2), 59-72.
- Simon, D., & P. Sturges. (2012). Libraries, e-lending and the future of public access to digital content. Paper given at IFLA, November 2012. Unpublished. (Brantley, 2013, p. 29)

- Soules, A. (2009). The shifting landscape of e-books. *New Library World*, 110 (1), 7-21.
- Staiger, J. (2012). How e-books are used. *Reference & User Services Quarterly*, 51(4), 355-365.
- Stachokas, G. (2012). Looking ahead from now to 2012. In R.O. Weir (Ed.), *Managing Electronic Resources* (pp. 137-148). Chicago: American Library Association.
- Triangle Research Libraries Network. (2013). *Home*. Retrieved from <http://www.trln.org/index.htm>
- Tripathi, M., & Jeevan, V. K. J. (2008). E-book subscriptions in a distance education institution: A case of Indira Gandhi National Open University, India. *Serials Review* 34(2). 104-114.
- University of North Carolina Library. (2013, May 17). TRLN consortial e-books acquisitions guiding principles. Report created XXX. Unpublished.
- University of North Carolina Library. (2012, November 14). UNC University Library e-books platform recommendations and e-books collections strategy. Report created XXX. Unpublished.
- Vasileiou, M., Rowley, J., & Hartley, R. (2012). The e-book management framework: The management of e-books in academic libraries and its challenges. *Library & Information Science Research*, 34(4), 282-291.
- Yale University Library. (2013, March 3). *The eBook Strategic Plan Task Force*. Retrieved from <http://www.library.yale.edu/departments/collection-development/Yale-ebook-task-force-rpt.pdf>

A.2.2. The International E-Book Landscape

Overview of Emerging E-Book Markets

Note: The excerpt below was published in the following report:

Goertzen, M., Wolven, R., & Carroll, J. (2015). Preservation pending: The future of e-book access in the digital age. In D. Bogart (Ed.), *Library and Book Trade Almanac* (pp. 23-40). Medford: Information Today, Inc.

The E-Book Program Development Study primarily focuses on trends and challenges faced by stakeholders in the United States. However, as markets become increasingly global in nature, it is important to understand e-book landscapes outside of North America as they will impact collection development strategies in the future.

In the English speaking world, the U.S. has the strongest e-book market - e-book sales account for approximately a third of book revenues, particularly in the segments of adult fiction and genre fiction (Rüdiger Wischenbart, 2014). Across Western Europe, the adoption of e-books by consumers varies. In Germany, which has one of the largest book markets in the world, e-book purchasing trends are beginning to follow English language markets, but are approximately three years behind (Rüdiger Wischenbart, 2014). Across France, Spain, and Italy, e-book “momentum seems to be fairly limited” (Rüdiger Wischenbart, 2014, p. 20). Most surprising are trends observed in Sweden, where a “domestic e-book market has hardly taken shape, due to a mix of high prices for e-books and the role taken up by libraries who are lending significantly more e-books to readers than retail has sold so far” (Rüdiger Wischenbart, 2014, p. 20).

In other parts of the globe, a number of large markets are emerging. Below are four markets that may impact the future of online educational resources.

1. China: The world’s most populous country is also the second largest e-book market in the world, with a 52 percent growth in 2012 (Kelbanoff, 2013). It is also estimated that there are 500 million Internet users and 1.22 billion mobile phone users in the country (Kelbanoff, 2013). Based on the market potential, competition between e-book providers is growing. In 2013, Amazon launched a dedicated Chinese platform, Apple’s iPhones were cleared to run on China’s mobile networks, and the domestic online platform, Dangdang, offered their catalog for free (Rüdiger Wischenbart, 2014). Currently, the fastest developing market for English-language books is focused on libraries, schools, and institutions (Kelbanoff, 2013).
2. Russia: During 2013, Russian e-book sales surpassed those in the United Kingdom and Brazil, making Russia the third largest e-book market in the world (Gerden, 2014). The

country has a “thriving reading culture in which writers and intellectuals occupy a prominent role in the public sphere and in which books stand at the center of the country’s cultural ambitions” (Rüdiger Wischenbart, 2014, p. 58). Since 2011, the market increased by nearly 200% and is expected to continue (Gerden, 2014). Although the sale of dedicated e-readers has declined, sales from personal computers and laptops continue to drive the market forward. Since 2012, Apple opened an iTunes store in Russia, and Google opened a Play store, and Kobo announced plans to offer services (Rüdiger Wischenbart, 2014).

3. Brazil: The Brazilian book market has shown strong growth for years with no signs of slowing down. For instance, consumers in 2013 spent 8.8 billion in books, a growth of 7% from the previous year (Utsumi, 2014). Currently, e-books are not widely used, but companies like Amazon, Apple, Google, and Kobo (all of whom set up ebookstores in the country within the past two years) believe there is large potential for development. In 2013, a number of Brazillian publishers attributed 4% of their market shares to e-books (Utsumi, 2014). Another factor that is expected to influence the market is the fact that customers do not pay VAT or sales taxes on books in Brazil (Rüdiger Wischenbart, 2014). “The Brazillian senate is discussing a bill that would officially make both e-books and dedicated e-readers taxfree. If the bill passes, the prices of E-Ink Kindles, Kobo devices, and Nooks would have to come down steadily” (Rüdiger Wischenbart, 2014, p. 65).
4. India: Currently, close to 24% of texts published in India are English-language titles. To date “all the major publishers in India, such as Penguin Books India, Hachette India, and Westland, have digitized their English-language backlists” (Vyas, 2014, para. 2). There is also great progress taking place in terms of regional e-books. With 25 regional languages spoken across the country, there is an opportunity for writers “outside of the country to make further inroads into the Indian market through getting their works translated” (Vyas, 2014, para. 4). This is especially true for self-published titles, as the legal process of forming partnerships with translation companies is less complex. On the technological side, dedicated e-readers have not made a significant market impact due to consumers’ questions about their overall value, but companies like Flipkart have successfully introduced e-reading apps that can be used on tablets and phones, allowing consumer to “read e-books on devices they are most familiar with” (Vyas, 2014, para. 9).

Emerging Markets: Japanese Case Study

Jonas Timson, a librarian from Waseda University in Shinjuku, Tokyo, Japan, came to CUL on a professional exchange program during the fall of 2013. At this time, I had the opportunity to work with Jonas on a literature review that examined e-book trends in Japan. The goal of the study was to experiment with a methodology to document segments of the international e-book

landscape in order to flag trends that may impact the way e-books are created, distributed, and preserved by libraries and cultural institutions around the globe.

History of the e-book market in Japan. In Japan, there had been several “eBook booms” in the past. However, those were preceded by sales of eBook reader lead by electricity industry and any of them didn’t last long, for the preparation on contents by publishers didn’t go together. Subsequently, domestic eBook market developed as a unique market that certain contents such as cell-phone comics and Keitai-shousetsu (cell phone novels) got advanced and became a considerable market scale.

In Japan, movement surrounding eBook is getting active these years, due to the activities of Amazon in America and their extend business in Japan. New eBook readers have been invented and sold. Major examples for eBooks platforms are Reader (SONY), Kobo (rakuten), BookLive (BookLive). In addition, extension on eBook publishing by not only book dealers and publishers but many companies of other industries are significant movement. Examples in publisher field, honto (DNP/Maruzen), Kinopy (Kinokuniya), Kindle store (Amazon); in distributor field, Rakuten kobo ebook store; and in electronics field, Reader@Store (SONY), BooksV (Fujitsu), etc.... Recently, National Federation of University Co-operative Associations joined the market in January, 2013.

The special feature of current eBook movement which is different from former eBook boom is that publishers and booksellers are actively preparing the contents, and also the eBook business is progressed by Digital Publishing Initiatives Japan Co., Ltd. which is established by investments from several publishers in April, 2012. Some publisher, especially centered by major publisher, started to prepare eBook contents actively for the merchandise assortment of eBooks, with the approach from eBook sellers. The number of eBook contents is different between each eBook seller, but the number of the contents provided by largest eBook seller is about 100 thousand as of March 2013. The breakdown is novels, practical guides, easy instructions, and still comics are in large number. However, not only the digitized version of existing book but the number of eBooks published simultaneously with paper version is increasing. The number of eBooks which are worth for collected by university libraries is still small, but number of contents which is published in eBook format almost the same time as paper version is arising (ex. Iwanami-shinsho, Kodansha-gendai-bunko, etc...). Some publishers such as university publisher and academic publisher (ex. Tokyo Denki University Press, University of Tokyo Press, and etc...) are appeared to be endeavoring to publish eBooks.

Digital Publishing Initiatives. Digital Publishing Initiatives Japan Co., Ltd. (pubridge) was established by the investment of major publisher/printer and public and private sector fund called “The Innovation Network Corporation of Japan (INCJ)” in April, 2012, and they progress an improvement of the infrastructure of eBooks business. Pubridge also takes charge of digitizing existing books in “emergence project of contents digitization” which is conducted by Ministry of Economy, Trade and Industry. Pubridge is already providing digitized contents to each eBook

sellers. (In addition, Pubridge's progress is only B to B progress and they do not do B to C.) To consider from these aspects, it can be said that current movement of eBook is totally different from former situation that the contents are prepared actively and it will surely expand in future. Still, eBooks being sold recently has its paper version, and it seems that no "eBook only" contents exist at this stage, except cell-phone novels called "keitai-shousetsu". In addition, in America, Amazon is likely to progress "eBook only" publishing close to self-publishing.

Market and economics. For those domestic eBooks for individuals does not have low price system as of Amazon in America, for example, pocket-size editions (Shin-sho) are same price as of paper version (ex. Iwanami-shinsho and Ko-dansha bluebacks), or 10 to 20 percent cheaper than paper version (ex. Ko-dansha Gendai shin-sho). Price gap between paper version and digitized version are not the same even if books were published by the same publisher (ex. price gap between paper version and digitized version of Pocket size editions and monographs published by ©Bungeishunju. Ltd. is different). To consider this, it can be said that price setting might be done based on several elements.

Cultural materials. As the revised National Diet Library Act is going to be operated from July 1st, National Diet Library will start Online Material Collecting System (e-depo). NDL obligate private company/person to deposit some of the e-books they published. The target will be free online materials (e-books, e-journals) with no DRM.

The "free and no DRM online materials" which have to be legally deposited will be materials with specific codes such as ISBN, ISSN, DOI, or materials prepared in specific format(PDF, EPUB, DAISY). Specifically, those will be annual reports, factbooks, bulletins, investigative reports, business reports, academic papers, academic proceedings, technical reports, handbooks, newsletters, novels, children's books, etc....

Conversely, materials which won't be collected are simple data such as blogs, tweets, trade catalogs, school (class) newsletters, files intended to be used as an application and permission, materials which are identified to be similar printing as paper versions such books and magazines, materials stored for permanent basis which won't be erased (ex. materials in institutional repositories), and materials published before the operation of revised National Diet Library Act.

Collected materials will be stored as cultural assets in the same manner of paper materials collected through legal deposit system. National Diet Library will launch online legal deposit page in their website from July 1st and start accepting electric resources. Peruse service for deposited materials will be provided inside National Diet Library. The service will start on October, 2013.

Academic materials. Recently, eBooks for individual readers became prevalent inside and outside the country, and approaching eBooks became direct issue for university libraries, therefore our committee (Committee on Scholarly Information) investigated the situation on eBooks and concluded the issues and made proposals for university libraries.

Commercial market: EBooks are now being prevalent rapidly in United States, for their convenience and price. In Japan, the number of eBooks published at the same time with paper version is increasing, and domestic eBook market for individual readers seems to develop in future. However, the issue is that eBooks for individual readers cannot be purchased directly by university libraries.

The results indicate that historically, the e-book market in Japan is geared towards materials that can easily be displayed on mobile devices (e.g. comics). However, since Amazon opened an online store in the country, there has been growth in e-book reader sales. Also, a number of companies like Digital Publishing Initiatives Japan Co., Ltd. and Pubridge are working to improve e-book platforms and digitize print books. Despite this progress, it appears that publishers are still searching for the most attractive price point for e-books. At this time, it does not appear that e-books are priced lower than print – a common practice in Amazon’s American store.

Across academic and cultural institutions, there is a growing interest in e-books, but acquisition barriers exist due to the fact that e-books are often licensed to individual users as opposed to institutions. A number of national initiatives are promoting wider access to e-books. For instance, the National Diet Library has started collecting online materials and mandates that authors deposit copies of all published e-books. Specifically, information pertaining to business reports, academic papers, technical reports, and cultural heritage are given top priority. The goal is to make these materials publically available without DRM restrictions. All material collected through the National Diet Library are classified and stored as cultural assets in the same manner as print collections.

At the end of the project, we realized that although there was documentation regarding the e-book landscape and publishing trends in Japan, we did not have an understanding of how Japanese materials were accessed and used at CUL. This gap in the methodology was addressed by a case study about Argentine e-books that took place in 2014.

Emerging Markets: Argentine Case Study

In Fall 2014, Sarah Goldberg conducted the E-Book Program Development Study: Argentine Case Study to fulfill the requirements for an internship at CUL within the Internships in Academic Administration program, a new Graduate School of Arts and Sciences. The initiative gives advanced PhD students the opportunity to gain experience on the administrative side of higher education. Pamela Graham, Director of Area Studies/Global Resource Division and Director of the Center for Human Rights Documentation and Research, and Melissa Goertzen, E-Book Program Development Librarian, supervised the project (Goldberg, 2014).

The goal of the study was to identify the primary issues and challenges related to the collection of Argentine e-books at university libraries in the United States. The project was completed in

several stages: a literature review that examined the current e-book landscape in Argentina, a review of the current Argentine e-book landscape at CUL, and a series of informal interviews with faculty and advanced graduate students in the Department of Latin American and Iberian Cultures (LAIC).

This case study documents the challenges the Argentine publishing industry has faced due to a difficult economic climate. The majority of small and mid-sized presses, including many publishers of scholarly works, do not yet publish electronic versions of their titles (2014). E-book sales represent approximately 2 percent of total book sales in Argentina, compared with 30 percent in the United States (2014). At the same time, a growing open-access movement in Argentina offers opportunities for university libraries in the United States to expand access to Argentine e-books among their patrons (2014).

The findings indicate that current availability of Argentine e-books at CUL is limited. Interviews with faculty and graduate students in LAIC suggest several avenues through which CUL may improve awareness of, access to, and user experience of e-books from Latin America and Iberia for members of the department. The report concludes with recommendations that might better inform e-book collection practices in LAIC and Area Studies more broadly.

Finally, the study supports the overarching goals of the E-book Program Development Study by providing a sustainable framework and methodology that can further research regarding non-English language publishing in world regions beyond the United States and Western Europe. Specifically, it points to questions that should be asked as CUL continues to develop areas studies collections. This includes an understanding of stakeholders, examination of copyright laws that impact the Libraries' ability to acquire, share, and preserve non-English language materials, and documentation of emerging markets that could change the way e-books are produced and distributed.

B.1. Review of E-book Business Models

To explore emerging trends that will impact e-book collection development activities at CUL, an ad hoc group came together in 2014 to evaluate a number of e-book platforms, services, and business models that may offer new acquisition and access options for users at Columbia University. This work subsequently resulted in the development of an evaluation framework for new potential offers from vendors and partnerships with partner institutions. The task force consisted of the following six members:

7. Bob Wolven, Associate University Librarian for Bibliographic Services and Collection Development;
8. Rob Cartolano, Associate Vice President for Digital Programs and Technology Services;
9. Jeff Carroll, Director of Collection Development;
10. Breck Witte, Director of Library Information Technology Office;
11. Susan Marcin, Head of Electronic Resources Management: Technologies and User Experience;
12. Melissa Goertzen, E-Book Program Development Librarian.

Representatives from 3M Cloud Library, BiblioBoard, and MyiLibrary were invited to give presentations on campus. After each meeting, the task force reviewed content offerings, proposed license agreements, the platform interface, and acquisition workflows. In addition trial access was set up for each platform and selectors were invited to preview content and submit feedback.

At the same time, the ad hoc group held meetings with Terry Kirchner, Executive Director of the Westchester Library System, and Denise Hibay, Head of Collection Development at the New York Public Library, to learn how each of the above mentioned platforms functions in a practical library environment. The conversations revealed that a significant percentage of e-book budgets are directed towards platform and maintenance fees, which significantly reduces the percentage of funds that are available to purchase or license titles.

After reviewing content, business models, acquisition workflows, and feedback from selectors, it was determined that at this time the above mentioned platforms do not have applications at CUL. As an alternative to licensing these products, CUL and NYPL discussed a library card pilot that would promote e-book content at both institutions to the user community at Columbia University.

Consortial PDA Trial: BiblioBoard Pilot with BorrowDirect

In 2014 BiblioBoard approached BorrowDirect with a pilot consortial access model. The group agreed to a trial period in order to evaluate available content and the proposed business model. A platform trial was set up at CUL and all library staff were provided with the opportunity to view

content, experiment with platform functionality, and submit feedback. The comments indicated uncertainty regarding the content's relevance to research, teaching, and learning activities on campus. The feedback from BorrowDirect partners pointed to similar concerns. At this time, the consortial PDA model has been declined.

One large question that surfaced during the pilot is the general sustainability of PDA purchasing models. While the benefits of this business model are attractive in the short-term (e.g. access to an entire catalog with minimal up-front financial commitments), it may create significant long-term barriers to the creation and distribution of scholarly monographs. The topic came up during a recent panel discussion at the 2014 Charleston Conference. Representatives from YBP, Stanford University Press, and Oxford University Press were asked to discuss how purchasing decisions made by libraries influences the way scholarship circulates and if/how it is produced. One of the take home points was the fact that PDA models put a great deal of pressure on publishers because no one is shouldering the economic risk with them. In many cases, it is difficult to break even on titles sold through PDA. Sales are also made difficult due to the fact that e-books do not have "brands" in the same way that journals do. Currently, publishers are searching for sustainable business models that promote the continued publication of scholarly work. Librarians will play a role in shaping these models as the e-book landscape continues to evolve.

Hybrid Purchasing Model: Taylor & Francis Pilot

During the 2014 Charleston Conference, I had a chance to meet with Taylor & Francis and discuss a new proposed business model specifically tailored for the needs of CUL. The proposal resulted from a discussion regarding concerns about perpetual access to scholarly monographs. Essentially, the model functions as a hybrid between e-book subscription and content ownership – by purchasing the 2015 front list, CUL would have complimentary access to the full backlist for the duration of the year.

After receiving the formal offer from Taylor & Francis, librarians in the Science and Engineering Library Division ran a cost analysis and examined content. At that point, it was deemed that the offer was overpriced. After a series of negotiations, it was determined that CUL cannot commit to this purchase in the current fiscal year. However, there is expressed interest in the purchasing model and some would like to the offer in a future fiscal year.

E-book Apps and Preservation: Yale University Press

During the summer, Bob Wolven, Associate University Librarian for Bibliographic Services and Collection Development, created the email alias *edesiderata* to collect specific examples of e-book acquisition, access, and preservation challenges from selectors. An excellent example was submitted by Carole Ann Fabian, Director of Avery Architectural and Fine Arts Library, and Paula Gabbard, Senior Bibliographer at the Avery Library.

In summary, Yale University Press released an *Interaction of Color* mobile app for iPads that provides a new way for students and scholars to interact with “one of the most influential books on color ever written...[It] offers users an entirely new way to experience Josef Albers’s original masterwork” (Yale University Press, 2014, para. 1). While this is a landmark title and one that the Avery Library was eager to acquire for users, there was no apparent way to bring the title into the collection.

The Collection Development Department explored a number of avenues, including searching for the title in Gobi (only the print version was available) and exploring institutional licensing options through the Apple iTunes Store. While an institutional license was available, it did not solve the central accessibility challenge – the Avery does not own a public iPad where users can access the app. Also, the fact that the app is licensed as opposed to purchased hinders CUL’s ability to capture and archive the content.

The challenge led to a fruitful discussion between Jeff Carroll (Director of Collection Development), Melissa Goertzen (E-Book Program Development Librarian), Daniel Dollar (Director of Collection Development at Yale University), and Sara Sapire (New Business & Product Manager at Yale University Press). Several points of discussion are listed below.

- **Library Access:** The Apple Volume Purchase Program is currently the only workable pathway for broader access to the app. YUP has worked with Yale to provide bulk purchases for classroom use.
- **Preservation:** The underlying images and text that compose the app could be candidates for preservation. YUP is working with CLOCKSS on the Stalin Archive, but that solution—which may not be a complete one—is the only thing they had done in this space. Sara was clearly looking for a potential partner on digital preservation. Sara sees this as a growing problem as YUP starts to release more e-books with enhanced content (embedded videos, etc.).
- **Possible Next Step:** A possible next step is the development of a web-based database. This would open the door to institutional sales, and maybe a preservation strategy of providing the content on a drive to purchasing libraries.

The conversation established contacts with a major university press and opened discussion regarding access and preservation challenges. In this way, this case study provided an opportunity for CUL to advocate for libraries and provide leadership in regards to continued discussion, and hopefully solutions, to these challenges.

See the Appendix C to view a blog post entitled “The Fine Art of App Preservation” that describes this particular situation in further detail. The post was published on the Art Libraries Society of North America (ARLIS/NA) Collection Development Blog.

Knowledge Unlatched: An Open Access Pilot

The Knowledge Unlatched (KU) pilot developed out of concerns regarding the current system of Open Access publishing for scholarly books, particularly those in the humanities and social sciences (Knowledge Unlatched, n.d.). The objective is to create a sustainable Open Access model for scholarly books and secure long-term cost savings for institutions through a system of sharing costs linked to a Creative Commons (CC) license (Knowledge Unlatched, n.d.).

Essentially, libraries from around the world share the cost of a Title Fee paid to a publisher in return for a CC license and PDF deposit in HathiTrust. Because the Title Fee is a fixed amount, as more libraries participate in the program, the costs to individual institutions decrease. During the pilot, KU secured pledges from over 200 libraries, CUL included, and secured CC licenses for 28 front list titles from 13 scholarly publishers (Knowledge Unlatched, n.d.).

At the end of the pilot project, CUL had a number of questions regarding the sustainability of the model, the funding framework, and the amount of overlap with current collection holdings. The model is currently undergoing evaluation and considerations for future participation.

B.2. E-Book Cost Analysis Project

B.2.1. Development of a Cost Analysis Method

One of the initial goals of the E-Book Program Development Study was to create an inventory of all e-book holdings at CUL. It was soon apparent that due to the volume and complexity of the data set, this project was outside the scope of the study.

At this point, I began looking at available data sets and decided to experiment with a new methodology that would compile a snapshot of e-book purchases (i.e. subscriptions, packages, and firm orders) made throughout the 2013 fiscal year (FY2013). The goal was to document how funds are allocated to purchase e-book content, and determine if acquired content is of value to the user community.

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 (e-book subscriptions) or EB (purchased content including e-book packages and firm orders) 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 I worked with Daisy Alarcon, Student Assistant, to run a Voyager query for all library funds ending in EO or EB. After running the

cumulative query, a base list was created for each of the following categories: subscriptions (EO), package purchases (EB packages) and individual purchases (EB firm orders).

Next, 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 (see Tables 3 & 4). Statistical analysis was conducted to determine the total, average, median, high, and low costs of each category (see Tables 5, 6, & 7).

EO Fund (Subscriptions)	# of Titles	Amount
Entire EO Budget	97	\$475,671.58
Bulk (~71%) of Budget	18	\$332,521.27
Tail (~29%) of Budget	79	\$143,150.31

Table 3. Budget breakdown of the EO fund (e-book subscriptions) in FY2013.

EB Fund	# of Titles	Amount
Entire EB Budget	736	\$1,100,342.03
EB Fund (Packages)	35	\$1,002,031.98
Bulk (~68%) of Package Budget	6	\$739,833.62
Tail (~32%) of Package Budget	29	\$262,198.36
EB Fund (Firm Orders)	701	\$98,310.05
Bulk (~70%) of Purchase Budget	195	\$57,248.45
Tail (~30%) of Purchase Budget	506	\$41,061.60

Table 4. Budget breakdown of the EB fund (e-book packages and firm orders) in FY2013.

EO Fund (Subscriptions)	
Total EO Cost	\$475,810.02
Number of subscriptions	97
Average Cost	\$4,905.26
Median	\$2,261.22
High	\$57,005.85
Low	\$138.44

Table 5. Statistical analysis of the EO fund (e-book subscriptions) in FY2013.

EB Fund (Packages)	
Total EB Cost	\$1,100,342.03
Package Cost	\$1,002,031.98
% of Total EB Cost	91.07%
Number of packages	35
Average Cost	\$28,629.49
Median	\$8,033.53
High	\$218,891.59
Low	\$1,200.00

Table 6. Statistical analysis of the EB fund (e-book packages) in FY2013.

EB Fund (Firm Orders)	
Total EB Cost	\$1,100,342.03
Purchase Cost	\$98,310.05
% of Total EB Cost	8.93%
Number of purchases	701
Average Cost	\$140.24
Median	\$99.00
High	\$1,949.00
Low	\$9.99

Table 7. Budget breakdown of the EB fund (e-book firm orders) in FY2013.

I decided to dig deeper into the data set and determine if the value CUL receives from e-book collections is closely aligned with the associated costs. To accomplish this task, I pulled usage data into the analysis to determine if CUL's funds are invested in heavily used e-book resources. Again, given the scope of this project, a decision was made to focus on the development of a cost analysis methodology and limit assessment to the top six e-book subscriptions and packages (ranked by cost) and the corresponding title lists were collected (Tables 8 & 9). At the same time, the corresponding BR2 COUNTER report (usage data) was pulled from the vendor/publisher website. Then, data was filtered to remove titles published prior to 2013.

Top EO Subscriptions	Price	% of Total Cost	
Subscription A	\$57,005.85	11.98%	
Subscription B	\$54,788.40	11.51%	
Subscription C	\$47,381.00	9.96%	Info not available
Subscription D	32,798.40	6.89%	Info not available
Subscription E	\$21,280.00	4.47%	
Subscription F	\$16,071.33	3.38%	

Table 8. Cost breakdown of the top six e-book subscriptions in FY2013.

Top Titles for EB Packages	Price	% of Total Cost
Package A	\$218,891.59	21.84%
Package B	\$216,888.00	21.64%
Package C	\$114,950.00	11.47%
Package D	\$66,907.63	6.68%
Package E	\$62,515.00	6.24%
Package F	\$59,681.40	5.96%

Table 9. Cost breakdown of the top six e-book packages in FY2013.

At this point, I encountered an unexpected challenge; 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, I 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, I filtered the data for a second time by matching the 2013 title lists with COUNTER report data using Excel. Next, I adapted a cost analysis framework used by the University of Western Australia to analyze DDA e-book models (Davies & Morgan, 2013, p. 172) and 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 (see Tables 10 & 11).

EO Fund Analysis	Subscription A	Subscription B	Subscription C	Subscription D
2013 Cost	\$54,788.40	\$21,280.00	\$16,071.33	\$57,005.85
No. of titles	89,529	6,631	116	14,339
No. of sections viewed	33,834	2,185	85	1,649
No. of loans	2,581,195	119,725	21,899	6,219
% of titles without use after purchase	62.21%	67.05%	26.72%	88.50%
Average cost of e-book	\$0.61	\$3.21	\$138.55	\$3.98
Cost per use	\$0.21	\$0.18	\$0.73	\$9.17

Table 10. Results of the cost and usage analysis for e-book subscriptions based on BR2 COUNTER data.

EB Fund Cost Analysis	Package A	Package B	Package C	Package D	Package E	Package F
2013 Cost	\$59,681.40	\$62,515.00	\$113,377.81	\$216,888.00	\$114,950.00	\$66,907.63
No. of titles	704	446	408	5,988	1,553	1,945
No. of titles loaned	153	79	111	4,509	294	686
No. of loans	2,937	486	2,143	119,085	8,620	11,360
% of titles without use after purchase	78.27%	82.29%	72.79%	24.70%	81.07%	64.73%
Average cost of e-book	\$84.77	\$140.17	\$277.88	\$36.22	\$74.02	\$34.40
Cost per use	\$20.32	\$123.63	\$52.91	\$1.82	\$13.34	\$5.89

Table 11. Results of the cost and usage analysis for e-book packages based on BR2 COUNTER data.

B.2.2. E-Book Subscription Cost Analysis Project

After analyzing the cost and usage data of the top e-book subscriptions (see Table 10), it was determined that the cost per use of Subscription D was high (\$9.17 per use) compared to Subscription A (\$0.21 per use), Subscription B (\$0.18 per use), and Subscription C (\$0.73 per use).

The results were presented to the Selectors' Group and the E-Resource Usage Data Working Group (ERUDWG) at CUL. The consensus was to conduct a second analysis of Subscription D based on the following criteria: evaluation of content, overlap analysis, and interface analysis (based on a report compiled by the Licensed Electronic Resource Interface Working Group (LERIWG)). The results indicated that Subscription D contained a large number of outdated technical manuals (96% published before 2011), a high number of titles available through other platforms, and incomplete multivolume sets.

Next, Amanda Bielskas, Head of Collection Development for the Science and Engineering Library (also a member of ERUDWG), examined the title list and identified 394 high use titles (more than 20 page views). Of this subset, she discovered that 196 titles (49.75%) are available through other platforms at CUL. Then, she searched GOBI for the remaining 198 high use titles and discovered that the vast majority are available for individual purchase. Based on this analysis, it was determined that Subscription D does not contain a significant amount of unique content. A decision was made to cancel the subscription.

When the decision was made to cancel Subscription D, CUL had the opportunity to speak with the Vice President and a team of sales representatives on three separate occasions. CUL received a reduced quote, and again, we conducted a cost analysis to examine its value (Table 12). Again, we determined that the quoted cost was not reflective of the value of content provided and the offer was declined.

At that time, the sales team requested feedback from CUL regarding how to improve the platform, and were provided with study findings. After a series of negotiations, Subscription D was renewed for one year at an 80% discount, and CUL received an annual savings of \$51,000 on this subscription. In preparation for the cancellation many of the high-use titles had been ordered as individual e-books, however the money spent on those did not approach the amount saved in the newly discounted subscription.

Subscription D	FY2013 Analysis	Quote Analysis
2013 Cost	\$57,005.85	\$30,360.00
No. of titles	14,339	14,339
No. of titles loaned	1,649	1,649
No. of sessions	6,219	6,219
No. of pages viewed	45,523	45,523
% of titles without use after purchase	88.50%	88.50%
Average cost of e-book	\$3.98	\$2.13
Cost per session	\$9.17	\$4.88
Cost per page view	\$1.25	\$0.66
Cost per title view	\$34.57	\$18.41

Table 12. Results of the cost and usage analysis for Subscription D.

After the negotiation was completed, the cost analysis methodology described in this chapter was adapted by a colleague in the Science and Engineering Library Division (SEL) to assess a series of eleven subscription packages. Again, findings pointed to high cost per use and title overlap rates. Following the established assessment protocol, results were presented to ERUDWG for further analysis. After viewing the results, the SEL division made a decision to cancel the two packages with the lowest usage rates and negotiated a flat price increase (on an annual basis) for the remaining nine packages. In total, this assessment resulted in a savings of \$10,000 on an annual basis. The findings indicate that this cost analysis methodology can be adapted and applied across the library system at CUL to yield results.

The findings of this study were presented at the 2014 Charleston Conference. To view the full paper scheduled for publication in the proceedings, please see Appendix G.

B.2.3. E-Book Front List Cost Analysis Project

At CUL, a large percentage of resources are directed towards front lists. When I analyzed the top five packages ranked by cost (see Table 11), the cost per use appeared high (averaging at \$36). A closer examination of the data revealed that many front list titles are not available to CUL users until the end of the year (largely due to publication dates). It seemed that evaluating the cost per use of 2013 front list titles based on 2013 COUNTER usage reports did not accurately reflect their value.

In order to develop a method to evaluate the cost per use of e-book front lists, the 2011 Package F front list was selected for evaluation because of the fact that it provided access to a rich data set (e.g. title lists, usage statistics) spanning several years. I located the Package F title list and matched it against BR2 COUNTER reports ranging in date from January 2011 to April 2014. Then, I experimented with a method to observe how usage and cost change over time (see Table 13).

Package F	2011	2012	2013	2014	Total
Cost	\$69,300.00				
No. of titles	1,091				
No. of titles loaned	365	824	566	203	942*
No. of loans	5,822	28,855	11,430	2,463	48,570
% of titles without use each year	723 (66.27%)	264 (24.20%)	522 (47.85%)	885 (81.11%)	Average: 503 (46.10%)**
% of titles without use after purchase					149 (13.66%)
Average cost of e-book (cost/titles)	\$63.52				
Cost per use (cost/use)	\$11.90	\$1.99	\$1.50	\$1.43	
* Number of titles from the frontlist that have circulated at least once.					
**The average was calculated based on the 2011-2013 data sets.					

Table 13. Package F assessment based on cost and usage data.

It was expected that the number of loans would also increase over time, but the results indicate a different trend. Between 2011 and 2012, loans increased by more than 80%. In the following year, the number of loans dropped by more than 50%. After considering collection content and usage trends, it seems that there are two possible explanations: 1) the titles were included in course reading lists and/or course reserves, and 2) users downloaded Package F titles in 2012 when they became available through CUL. During this analysis, an attempt was made to identify all 2011 Package F titles that were included in course reserves over the past three years. However, the time involved to extract this data is not conducive to the time frame for the E-Book Program Development Study. The topic has been flagged for a future study and may inform recommendations made at the end of the project, as findings would provide information about how patrons use the e-book collection.

Acknowledgements: I would like to thank Daisy Alarcon at Columbia University for her assistance with data collection from Voyager, Serial Solutions, and COUNTER usage reports.

B.2.4. Cost Analysis Paper Presentation, 2014 Charleston Conference

Preprint version. Scheduled for publication in the 2014 Charleston Conference Proceedings.

The Buck Stops Here: Assessing the Value of E-book Subscriptions at the Columbia University Libraries

Melissa J. Goertzen, Columbia University, mjg2227@columbia.edu

Krystie Klahn, Columbia University, kk2926@columbia.edu

Abstract:

Over the past four years Columbia University Libraries (CUL) has seen exponential growth in electronic book (e-book) purchasing. These purchases have not only increased the depth and breadth of the collection, but they have also created new opportunities for remote learning and instant information access. In turn, this new push for purchasing electronic has created new demands in assessment to understand the true benefit of these resources, most notably in regards to annual e-book subscriptions.

In 2013, a new position aimed at developing an e-book strategy for CUL was devised. Shortly thereafter, a position was created in the Science and Engineering Library Division (SEL) that focused on assessment. These two positions fall in line with CUL's mission to support research and learning through evidence based decision making.

This paper explains an assessment methodology used within the library system at CUL to evaluate the cost-benefit of e-book subscriptions. By appraising several databases, we were able to analyze cost and usage to determine the actual value of these resources. The findings yielded a savings of approximately \$60,000 for the 2015 fiscal year. This is an ongoing initiative that will help us document the e-book landscape and build data sets that will inform collection development decisions.

Part 1: Introduction

The dawning of the Digital Age revolutionized the way information and knowledge are created, produced, and disseminated in the academic community. The rapid integration of technologies with research, teaching, and learning activities has changed both information and access needs of user communities. A new reliance on electronic content coupled with the pressure of reduced stack space for print collections are factors that contributed to a new focus on electronic book (e-books) acquisition initiatives in academic libraries. Librarians now face the challenge of assessing and evaluating this new format in regards to the value it offers to users.

Columbia University Libraries (CUL) is one of the top five academic research library systems in North America and serves a community of over 3,750 faculty members and 26,000 full-time students at the Morningside Campus and Medical Center. The collections are housed across twenty-one campus libraries and include over twelve million volumes, 160,000 current journals and serials, and an extensive collection of manuscripts, rare books, microforms, maps, and audiovisual materials. In 2004, CUL began purchasing e-books in an experimental capacity. Due to positive reception by faculty and students, CUL began expanding e-book collections to support research, teaching, and learning activities across campus. Currently, CUL provides

access to more than two million e-book titles and expenditures comprise 25 per cent of the total book budget.

As the e-book collection continues to grow, CUL is developing a unique strategy and vision for e-book collection development, programs, and initiatives. To achieve this goal, the Collection Development Department launched the E-Book Program Development Study in 2013. This assessment project centers on the collection of essential data to drive the development of policies related to e-book acquisition, discovery, and access. In 2014, the Science and Engineering Library Division (SEL) created the position of Collection Assessment and Analysis Librarian in order to promote regular and standardized assessment of existing collections.

In 2014, 96 active e-book subscriptions were identified at CUL. The objective of this paper is to describe an assessment methodology used to evaluate the cost-benefit of these e-book subscriptions across the entire library system, and then within SEL. By appraising several databases, we were able to analyze the cost, usage, and value of these resources. This work involved conducting cost analysis, examining usage trends, and conducting title overlap analysis. The study was guided by the following four principles outlined in the CUL/IS Strategic Plan 2010 – 2013:

1. User-focused design;
2. Data-driven decision making;
3. Continuous assessment of results;
4. Flexible and adaptive response to user needs (Columbia University Libraries, 2010, p. 8).

The assessment method described in this paper was applied on two separate occasions in the Collection Development Department and SEL between November 2013 - May 2014. In both cases, the method provided actionable results and contributed to the standardization of e-book assessment and negotiation strategies across campus. The findings yielded a savings of approximately \$60,000 for the 2015 fiscal year. Since savings are tied to subscription packages, they will carry forward into coming years. In addition, both studies resulted in discussions with vendors regarding price negotiation and platform performance. Finally, we gained a deeper understanding of how resources can be allocated to best support the needs of user communities. The ongoing evaluation of e-book subscriptions will allow for evidence based decision making and assist in the continued documentation of the e-book landscape at CUL.

Part 2: E-Book Program Development Study

The E-Book Program Development Study is an ambitious, two year 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 objective at the heart of the E-Book Program Development Study is to develop a strategy and vision for e-book programs and collections at CUL. Essentially, the set of recommendations

that result from study findings will create a bridge between the current landscape and CUL's vision for future e-book initiatives on campus. A large part of this work involves the development of methodologies that examine how e-book resources are allocated, evaluate current subscriptions and packages, examine usage trends, and build collaborative relationships with vendors. The data that was gathered while developing these methodologies will be used to inform recommendations and policy statements regarding e-book collection development and management on campus.

The reality that the e-book landscape is constantly evolving was factored into decisions regarding the overarching assessment framework guiding this study. The research design was created so that it 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.

As part of this study, a large-scale cost analysis project was completed between November 2013 – March 2014. The goal of the project was to collect quantitative data that will inform e-book collection development policies in regards to fund allocation and preferred business models. While the study examined both e-book subscriptions and purchases, for the purposes of this paper, the discussion will be limited in scope to e-book subscriptions.

After discussions with the Director of Collection Development, and the Head of Electronic Resources Management, it was determined that e-books are most often purchased on the EO fund code. To collect financial data for all e-book subscriptions, a Voyager (CUL's integrated library system) query was run for all library funds ending in EO. After running the cumulative query, we created a base list of all e-book subscriptions at CUL. Given the enormous size of this data set and the time restrictions placed in the study, data collection was limited to subscriptions that had fund activity during the 2013 fiscal year (FY2013). This limited the base list to 96 e-book subscriptions. Next, subscription fees in FY2013 were totaled, and calculations were made to identify the top 70 per cent (bulk) and bottom 30 per cent (tail) of purchases within the budget. Statistical analysis was also conducted to determine the total, average, median, high, and low costs.

To examine usage trends, the top four subscriptions (ranked by cost) were selected and corresponding title lists were collected from the Continuing & Electronic Resources Management (CERM) Division. At the same time, the corresponding BR2 COUNTER report was pulled from the vendor website. At this point, we encountered an unexpected challenge; in several cases, multiple collections from the same vendor are purchased as separate items on the EO fund code. However, there is no apparent way to filter COUNTER reports by collection. At this point, we considered analyzing the data by vendor 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. Next, we adapted a cost analysis framework used by the University of Western Australia to analyze DDA models (Davies & Morgan, 2013, p. 172). After examining our data set, 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 (see Table 1). Please note that the results were calculated using confidential data. For the purposes of this paper, the numbers were

changed and percentages are not exact, but they reflect the trends discovered in the actual study findings.

	Subscription A	Subscription B	Subscription C	Subscription D
2013 Cost	\$50,000.00	\$20,000.00	\$15,000.00	\$60,000.00
No. of titles	80,000	6,000	125	11,000
No. of titles loaned	34,000	2,100	90	1,600
No. of loans	2,500,00	11,900	22,00	6,500
% of titles without use after purchase	62%	65%	25%	85%
Average cost of e-book	\$0.60	\$3.00	\$140.00	\$5.00
Cost per use	\$0.20	\$0.20	\$0.75	\$9.00

Table 1. Recalculated results of the e-book subscription cost and usage analysis.

After analyzing the cost and usage data of the top e-book subscriptions, it was determined that the cost per use of Subscription D was high (\$9.00 per use) compared to Subscription A (\$0.20 per use), Subscription B (\$0.20 per use), and Subscription C (\$0.75 per use).

The results were presented to the E-Resource Usage Data Working Group (ERUDWG) at CUL. The consensus was to conduct a second analysis of Subscription D based on the following criteria: evaluation of content, overlap analysis, and interface analysis. The results indicated that Subscription D contained a large number of outdated technical manuals (96 per cent published before 2011), a high number of titles available through other platforms, and incomplete multivolume sets.

Next, a team of librarians from SEL examined the title list and identified 394 high use titles (more than 20 page views). Of this subset, she discovered that 196 titles (49.75 per cent) are available through other platforms at CUL. Then, she searched GOBI for the remaining 198 high use titles and discovered that the vast majority are available for individual purchase. Based on this analysis, it was determined that Subscription D does not contain a significant amount of unique content. A decision was made to cancel the subscription.

One unexpected outcome of this analysis was the opportunity to speak with the Vice President and a team of sales representatives managing Subscription D on three separate occasions. The company requested feedback from CUL regarding how to improve the platform interface and content, and were provided with study findings. After a series of negotiations Subscription D was renewed for one year at an 80 per cent discount, resulting in a cost savings of \$50,000. At the end of FY2014, Subscription D will be assessed using the same methodology to determine if it will be renewed the following year. At this time, we are also investigating a number of marketing initiatives to promote content and will monitor usage rates to determine if there are noticeable changes over the next twelve months.

Part 3: E-Books in the Science and Engineering Library

The Science and Engineering Library Division (SEL) at CUL was once comprised of eight different libraries. Over the past few years there have been a series of library closures, with the most recent being the closure of the Engineering Library in June of 2014, bringing the total to four libraries. The closures have physically reduced library space and have forced the consolidation of a majority of print books to off-site storage. In addition, these closures coupled with the strong acceptance of electronic content by faculty and students have made SEL the perfect subject for e-book development. There is now a preference policy to purchase electronically when available. This in turn has led to the desire to learn how cost effective electronic collections are and if new models of purchasing would be beneficial.

In recent years, the SEL has moved away from subject specialist roles to focus on a functional role structure. This structure will allow our positions to evolve in a seamless manner regardless of how the landscape at SEL changes in the coming years. While there are still purchasing and liaison responsibilities related to specific subjects, librarian roles focus on division wide functions. For example, in May of 2014 a new position was created to focus on collection assessment and analysis of materials for the entire science division. This new position allows for the designation of much needed time towards the evaluation of our large collection of electronic and print content.

With regards to electronic content, increased cost trends that are inconsistent with library budgets provided the catalyst for initiating the new role. There needed to be a systematic and concentrated effort on assessment to ensure collections warranted purchase and were seen as valuable to the collection. Thus this position allows for the justification of purchases as well as cancellations. Moreover, continued analysis is beneficial in maintaining quality curation of collections. It can highlight under-utilized content as well as reaffirm valuable resources. In-depth looks at electronic content can even point to discovery or access issues, thus addressing silent threats.

The first assessment project for this role analyzed several subscription based e-book packages from one vendor (Vendor A). Vendor A provides access to over forty e-book subject collections. In 2014, CUL subscribed to eleven of these collections. Over the past few years each collection has seen a four to seven per cent cost increase annually, and overall usage of these e-books dipped in 2013. It was also noted that Vendor A reserves the right to drop content from collections without adjusting the price. Instead, they add content which is not always comparable. Before renewing these subscriptions, we needed to find out if each collection was still worthwhile for our library. We decided to conduct a cost analysis to prove or disprove benefit.

We were able to acquire usage data from COUNTER compliant BR2 reports posted on Vendor A's admin site, which aligns with our preference mandate to collect standardized data sets. When we compared usage statistics against fund data gathered from Voyager, we found that the eleven collections seemed like a good deal with a low cost per use (\$2.24), which is in line with comparable electronic subscriptions at CUL. However, we also needed to account for the fact that CUL does not subscribe to the eleven collections from Vendor A in one lump sum, but rather subscribes to each on an individual basis.

During assessment, we also discovered a 27 per cent overlap of titles across the eleven collections we subscribe to from Vendor A. This per cent overlap increased between each package and some overlap reached as high as 66 per cent. Upon further investigation, we also found that this content overlapped with titles available through other e-book platforms at CUL. This discovery prompted a deeper analysis of content available through Vendor A, and we discovered a long tail of use with a handful of titles receiving heavy use and the majority receiving low (in many cases zero) uses. In total, our patrons used approximately 25 per cent of collection content. Although this is a common trend in e-book subscriptions, we flagged the collections for further investigation nonetheless.

We also noticed that the cost per use for each collection ranged from \$.38 to \$11.21. While some collections were well worth the purchase price, others seemed to be far too high a cost. After breaking down the purchases, there were three collections highlighted as questionable (see Table 2 and 3).

		Questionable Packages		
	Vendor A Subscription	Package F	Package J	Package K
# of Titles in Database	9161	717	713	238
# of Titles Overlapped	2444	471	241	121
% of Overlap	27%	66%	34%	51%

Table 2. Results of the overlap analysis for Vendor A (FY2013).

		Questionable Packages		
	Vendor A Subscription	Package F	Package J	Package K
Cost	\$ 50,000.00	\$ 6,000.00	\$ 3,000.00	\$ 2,500.00
Unique Usage	22324	841	351	223
% Used	25%	27%	11%	21%
Cost Per Use	\$ 2.24	\$ 7.13	\$ 8.55	\$ 11.21

Table 3. Recalculated results of the cost and usage analysis for Vendor A based on BR2 COUNTER data (FY2013).

These findings were presented to our colleagues in SEL as well as ERUDWG for discussion. Open conversations with these groups helped us to further analyze the subscriptions and compile feedback that was presented to Vendor A. As a group, we decided to cancel the two lowest used collections and we negotiated a flat percentage increase for two years with the remaining nine collections. With the flat price increase, we will see a savings of \$10,000 annually and believe these collections will be sustainable and cost effective going forward.

Part 4: Conclusion: Library as Advocate and Negotiator

Our findings demonstrate that standardized usage reporting in the form of COUNTER reports are beneficial to the assessment of collections. The reports allow “the usage of online information products and services to be measured in a credible, consistent and compatible way using vendor-generated data” (ProjectCOUNTER, 2012, p. 1). Information such as title, publication year, and unique identifiers like DOI help librarians understand the genuine usage patterns at their

institutions. When vendors are compliant with COUNTER it makes assessment more manageable and meaningful by standardizing what is considered use. Therefore, we are no longer comparing apples to oranges but rather apples to apples.

The decisions that were made using this assessment methodology will not only affect our budgets but also the content in our collections. We learned that it is important to present data and findings to library divisions as well as committees or working groups to weigh options and provide feedback. Different viewpoints can point out issues that may have been initially excluded from assessment plans. In addition, discussing findings can increase awareness of initiatives, such as marketing collections to promote discovery, which may result in greater buy-in from stakeholders, thus increasing the value of e-book collections.

The idea that the term “librarian” is not synonymous with “negotiator” must be changed. We need to be our own advocates, make informed decisions, and demand change if change is needed. Large price increases that are inconsistent with subscription use should not be tolerated. Assessment projects such as the one described in this paper, prove the true value of resources and can be effectively used to negotiate acceptable price tags based on use.

The reality that the e-book landscape is constantly evolving was factored into decisions regarding the overarching assessment framework. The research design was created so that it 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. This study proves that this methodology can be used on varying collection sizes and provides actionable results.

In closing, systematic and routine assessment of collections is important in determining the true value of resources for our user community. It is imperative to make informed decisions when negotiating renewals rather than passing the “buck.” We hope that the experiences and findings outlined through this paper will help others implement an assessment program at their institution.

References

Columbia University Libraries. (2010, January). *Strategic Plan, 2010 - 2013*. Retrieved from https://culis.columbia.edu/content/dam/staffweb/admin/strategicplanning/2010-2013/CULIS_Strategic_Plan_2010-20131.pdf

Davies, T., & Morgan, M. (2013). E-books down under. *Proceedings of the Charleston Library Conference*. <http://dx.doi.org/10.5703/1288284315253>

ProjectCOUNTER (2012, April). *The COUNTER Code of Practice for e-Resources: Release 4*. Retrieved from <http://www.projectcounter.org/r4/COPR4.pdf>

Acknowledgements: We would like to thank Daisy Alarcon at Columbia University for her assistance with data collection from Voyager, Serial Solutions, and COUNTER usage reports.

B.3. E-Book Comparison Chart

Platform Information				E-book Functionality							
E-book Platform	Website	Platform Type	E-book File Format	Usage Reports--COUNTER Compliant	One User per Book	Unlimited/Multiuser Access	Full Text Search	Copy/Paste	Printing	Offline Reading	Annotation Tools
EBL (Ebook Library)	www.eblib.com	E-book Lending Service	Adobe Digital Editions	Yes	No	Yes (unlimited, multiuser)	Yes	Yes (up to 5% of any book's content)	Yes (up to 20% of any book's content)	Yes	Yes
Infobase eBooks	http://infobasepublib.com/ebooks	Publisher	HTML, PDF, ePub (for some, not all titles)	Yes	No	Yes (for collections)	Yes	Yes	Yes	Yes	Yes
UPCC book Collections on Project MUSE	http://muse.jhu.edu	University Press Consortium	PDF	Yes	No	Yes (unlimited)	Yes	Yes	Yes	Yes	Yes
eBooks on EBSCOhost	http://ebshost.com/ebooks	Aggregator	Adobe Digital Editions, Adobe PDF plugin (for Mac)	Yes	Yes (for perpetual purchases)	Yes	Yes	Yes (up to 50% of titles)	Yes	Yes	Yes
Ebrary	www.ebrary.com	Aggregator	Adobe Digital Editions, Java, enabling of cookies & popups required	Yes	Yes (unlimited, multiuser)	Yes (unlimited, multiuser)	Yes	Yes	Yes	Yes	Yes
MyLibrary	www.mylibrary.com	Aggregator	PDF, ePub	Yes	Yes	Yes (multiuser, perpetual)	Yes	Yes	Yes	Yes	Yes
Taylor & Francis eBooks	www.tandfebooks.com	Publisher	PDF	Yes	Yes	Yes (unlimited, multiuser)	Yes	Yes	Yes	Yes	Yes
Sharpe Online Reference	N/A	N/A	HTML	Yes	No	Yes (unlimited)	Yes	Yes (content & images)	Yes	No	Yes
Oxford Handbooks Online	www.oxfordhandbooks.com	Publisher	XML	Yes	No	Yes (unlimited)	Yes	Yes (in PDF)	Yes	No	Forthcoming
Oxford Reference	www.oxfordreference.com	Publisher	XML	Yes	No	Yes	Yes	Yes (in PDF)	Yes	No	Forthcoming
University Press Scholarship Online	www.universitypressscholarship.com	University Press Consortium	XML (printable & downloadable in PDF at chapter level)	Yes	No	Yes (unlimited, multiuser)	Yes	Yes (in PDF)	Yes	No	Forthcoming
Gale Virtual Reference Library	N/A	N/A	HTML, PDF	Yes	No	Yes (unlimited, perpetual)	Yes	Yes	Yes	Yes	No
Literati by Credo	http://literati.credoreference.com	Aggregator	HTML (files may be saved in PDF)	Yes	No	Yes (for collections)	Yes	Yes	Yes (by saving in PDF)	Yes	No
Palgrave Connect	www.palgraveconnect.com	Publisher	PDF for all; ePub for books published in 2011-	Yes	No	Yes (perpetual)	Yes	Yes	Yes	Yes	No
SpringerLink	http://link.springer.com	Publisher	HTML, PDF	Yes	No	Yes (perpetual)	Yes	Yes	Yes	Yes	No
Books247	www.books247.com	Aggregator	HTML, PDF	No; COUNTER style reports used	No	No (but available as an option within the subscription package)	Yes	Yes	Yes	Yes	Yes
Cambridge Books Online	http://ebooks.cambridge.org	Publisher	PDF	Yes	No	No (but available upon request)	Yes	Yes	Yes	Yes	Yes
SAGE Knowledge	http://knowledge.sagepub.com	Publisher	HTML, PDF	Yes	No	No	Yes	Yes	Yes	Yes	Yes
OverDrive	www.overdrive.com	Distributor/E-book Lending Service	PDF, ePub, Kindle	Yes	Yes (simultaneous & metered access)	No	Yes (for OverDrive Read only)	Yes (if publisher allows)	Yes	Yes	Yes
Questia	www.questia.com	Aggregator	HTML	No	No	No	Yes	Yes (portions of the content)	Yes (one page at a time)	No	Yes
Safari Books Online	www.safaribooks.com	Aggregator	PDF	Yes	No	No	Yes	Yes	Yes	No for libraries; Yes for corporate & consumer accounts	Yes
PsychBOOKS	www.apa.org/pubs/databases/psychbooks	Publisher	PDF	Yes	No	No	Yes	Yes	Yes	Yes	No
University Publishing Online	http://universitypublishingonline.org	University Press Consortium	PDF	Yes	No	Not standard but available upon request (via perpetual access)	Yes	Yes	Yes	Yes	No
SpringerReference	http://springerreference.com	Publisher	HTML, PDF	Yes	No	No	Yes	Yes	Yes (in PDF)	No	No

C.1. Paper Presentation, 2014 CUNY Library Assessment Conference

Note: The following paper was published in the proceedings of the 2014 CUNY Library Assessment Conference (<http://www.cuny.edu/libraries/conference/proceedings.html>).

**Scholarly E-Book Use across Disciplines:
Content Analysis of Usage Reports and Search Terms**
Nisa Bakkalbasi (Assessment Coordinator)
Melissa Goertzen (E-Book Program Development Librarian)
Columbia University Libraries, New York, New York, USA

Abstract

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 Columbia University Libraries (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.

In this study, we sought an innovative research method to understand e-book usage. This 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.

Introduction

Over the past decade, electronic books (e-books) have become increasingly popular in the academic community. In response to this demand, Columbia University Libraries (CUL) provides access to over two million e-books that support research, teaching, and learning activities across campus and within the wider scholarly community. As the collection continues to grow, CUL is developing a unique strategy and vision for e-book programs and initiatives. To achieve this goal, the Collection Development Department launched the E-Book Program Development Study in 2013. This ambitious assessment project centers on the collection of essential data to drive the development of policies related to e-book acquisition, discovery, and access.

During the same year, data collected through COUNTER usage statistics and the LibQUAL+ service quality assessment survey indicated 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, we wanted to understand scholarly e-book usage in various disciplines.

The aim of this study is to better understand how scholarly e-books are used in various disciplines in teaching, learning, and scholarly pursuits through readily available data. This study seeks to gather data to drive the creation of best practices and policies to support the delivery of e-book collections and programs that facilitate research, teaching, and learning across campus and within the wider scholarly community.

Literature Review

Determining how e-books are used for scholarly purposes is a complex issue. The e-book landscape is evolving at a rapid pace and a wide range of factors, including business models, e-book formats, and platform functionality, impact how library clients discover and access e-books for research, teaching, and learning activities. It is more important than ever for librarians to understand when, how, and why clients use e-books in order to design services that meet existing needs.

Over the past several years, a number of studies were conducted to determine how e-book use differs across scholarly disciplines. Littman and Connaway (2004), Christianson (2006), Bailey (2006), and Kimball, Ives, and Jackson (2010) examined e-book use according to subject and all suggest that the highest usage rates were typically found in computers, technology, business, and the sciences. The lowest usage rates were most often discovered in the humanities and arts. This finding was consistent across academic institutions of various sizes, funding structure, and missions. Staiger (2012) discovered a trend that suggests a relationship between the currency of an e-book and its relevance to researchers, particularly in fields like business, computer science and technology. He attributed this finding to the fact that researchers in these disciplines have an acute need for current information.

A study by Levine-Clark (2007) suggests that there is no correlation between the awareness of e-book collections within disciplines and e-book usage rates. At the University of Denver, Levine-Clark conducted a survey that measured knowledge and usage of e-books in the humanities. In total, 2,067 faculty, students, incoming students, and alumni responded. The results indicated that 74.4 percent of humanists were aware of e-book collections available through the university. In all other disciplines, awareness ranged from 49 to 69 percent. However, humanists use e-books less often than scholars in other disciplines.

A number of studies have been conducted to understand how e-books are used for research, teaching, and learning activities. Shelburne (2009) conducted a large scale survey to learn about e-book usage patterns at the University of Illinois. In total, 1,547 responses were received. The results indicated that 78 percent of e-book use was intended for research purposes, 56 percent for study, 2 percent for teaching, and 2 percent for other purposes.

Levine-Clark (2007) found that library users typically “use rather than read” e-books. Typically, the format is viewed as a convenient source that provides quick reference for scholarly endeavors. Results from a survey of 2,067 faculty, students, incoming students, and alumni indicated that 56 percent of respondents use e-books to read a chapter or article within a book, and 36 percent typically read a single entry or several pages.

Noorhidawati and Gibb (2008) and Berg, Hoffman, and Dawson (2010) suggest that e-books are primarily used for quick reference, limited reading, and citation checks as opposed to extended reading and research. In other cases, e-books serve as a convenient means to preview a text; students and faculty members peruse the e-version to gain a sense of the information, biases, or arguments presented in a scholarly monograph. If it is useful for their research purpose, a print version is often requested for extended reading.

A literature review by Staiger (2012) compared the results of two dozen studies regarding e-book usage by members of the academic community. Findings suggested that “academic users typically search e-books for discrete bits of information, a behavior summed up by the formula ‘use rather than read’” (p. 355). In general, members of the academic community do not immerse themselves in e-books for extended periods of time to examine entire arguments. Instead, they view e-books as “convenient sources from which to extract information for their scholarly endeavors” (p. 357). Essentially, e-books provide a means for power browsing. They allow users to preview a book without leaving their work stations, and then locate the print copy if the information is relevant to their studies (2012). A literature review by Ashcroft (2011) uncovered similar trends. Statistics showed that on average, “53.5 percent of students and 58.6 percent of teachers dipped in and out of several chapters, whereas very low percentages read the whole book – 5.5 percent of students and 7.1 percent of teachers” (p. 401).

E-book Collection at CUL

CUL is one of the top five academic research library systems in North America and serves a community of over 3,750 faculty members and 26,000 full-time students at the Morningside Campus and Medical Center. The collections are housed across 21 campus libraries and include over 12 million volumes, 160,000 current journals and serials, and an extensive collection of manuscripts, rare books, microforms, maps, and audiovisual materials. In 2004, CUL began purchasing e-books in an experimental capacity. Due to the positive reception by faculty and students, the Library continued to grow e-book holdings to support research, teaching, and learning activities across campus. Currently, CUL provides access to over two million titles.

CUL offers e-books through subscriptions packages (e.g. Knovel, Ebrary, Safari) as well as individually purchased titles. The Library also licenses e-books through publishers' packages, including Cambridge University Press, Oxford University Press, Springer, and Wiley. Over the past several years, CUL has partnered with a number of academic and research institutions through consortial groups to investigate business models for shared e-book purchasing, including the Manhattan Research Library Initiative (MaRLI), 2CUL, Knowledge Unlatched (KU) and the North East Research Libraries (NERL) Consortium.

Methodology

Before discussing the methodology in detail, it is worth mentioning that our initial thought was to create a survey to gather information about e-book use across disciplines. However, two key factors influenced our assessment strategy and motivated us to tap into existing data sources rather than developing a survey instrument. First, during our initial consultations, it became apparent that using a low-overhead data collection technique that would allow us to systematically collect information over time would be most appropriate for this project. Due to our interest in continuously monitoring our user base in an ever-changing e-book landscape, reliance on readily available, continuous, and accurate data was an important factor in creating an effective and sustainable assessment plan.

Second, as survey participation rates have declined, survey research has experienced significant challenges that impact its use in library assessment plans. Participating in a survey to provide thoughtful and reflective feedback requires time and effort from respondents. The quality of the data begins to deteriorate when potential respondents do not make the effort to submit a completed survey or leave the survey incomplete. Surveys are of little, or no use, if the response rate is low or the data is inaccurate. Based on the low response rates from a recent survey, and in an attempt to avoid survey fatigue, we investigated alternative approaches of data collection.

In this study, we sought an innovative research method to understand e-book usage. This 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 sets present CUL with an accurate, continuous, and objective picture of e-book use.

The study covers the period from January 1, 2013 through December 31, 2013. It is worth noting that CLIO became the default discovery tool for the library at the beginning of June 2013. Thus, searches tracked by Google Analytics prior to June 2013 are limited. We included eight major e-book platforms in the study (i.e. Springer, Wiley, Oxford University Press, Elsevier, EBSCO, Ebrary, Cambridge University Press, and Safari Books Online) to ensure e-books were included from all three major disciplines, namely humanities, social sciences, and sciences.

For the indicated time period, we exported all search terms limited by format to e-books from our Google Analytics account. After data clean-up and formatting, requested e-book titles from COUNTER reports and e-book search terms from our Google Analytics account were loaded

into the qualitative analysis software, NVivo to identify frequently used words and explore recurring patterns. Then, we performed text analysis to generate word frequency tables and word clouds for each of the frequency sets to graphically display how each of the collections, at least in terms of the titles used, covers a different sector of the e-book platform universe.

Findings and discussion

The most frequently repeated search word was “history,” which was entered 526 times into the search field to search for e-books. It was followed by the word “theory” (entered 378 times). The most frequently requested e-book title word was “edition” (repeated 3,284 times), followed by the word “volume” (repeated 2,306 times). In the preliminary analysis, we refrained from adding words such as "edition," "volume," and “2nd” to a stop list, as we determined they might shed a special light on what was being searched and delivered in some instances.

Table 14 lists the top 20 most frequently repeated search words and requested title words. We found an overlap of 60% (15 words) in both lists, indicating a correlation between search and delivery of e-books. The words that are present in both lists are reported in italics (see Table 1).

Table 14. Most frequently repeated search and requested title words

	Search terms			Requested title words		
Rank	Word	Length	Count	Word	Length	Count
1	<i>history</i>	7	526	edition	7	3284
2	<i>theory</i>	6	378	volume	6	2306
3	<i>social</i>	6	368	<i>history</i>	7	1949
4	introduction	12	359	<i>theory</i>	6	1777
5	<i>new</i>	3	358	<i>new</i>	3	1730
6	<i>analysis</i>	8	326	<i>american</i>	8	1689
7	<i>american</i>	8	309	<i>analysis</i>	8	1651
8	<i>handbook</i>	8	303	advances	8	1577
9	human	5	281	systems	7	1558
10	<i>research</i>	8	281	culture	7	1552
11	<i>health</i>	6	265	studies	7	1532
12	<i>world</i>	5	227	<i>world</i>	5	1510
13	<i>modern</i>	6	223	<i>guide</i>	5	1502
14	<i>guide</i>	5	219	<i>social</i>	6	1479
15	law	3	211	<i>handbook</i>	8	1468
16	medicine	8	207	applications	12	1412
17	<i>management</i>	10	198	<i>politics</i>	8	1367
18	rights	6	193	<i>science</i>	7	1365
19	war	3	191	<i>modern</i>	6	1230
20	<i>development</i>	11	188	<i>research</i>	8	1198

It is challenging to deduce reader intent from word frequencies, as text data remain widely open for interpretation. However, responses to open-ended questions from the most recent LibQUAL+ survey are consistent with our findings that e-book collections are widely used across all major disciplines to support instruction and learning. User sentiments from the LibQUAL+ survey mirror a number of findings mentioned in the literature review, namely by Levine-Clark (2007) and Shelburne (2009), who suggest that e-books are used primarily to read chapters or articles for study purposes.

The ability to analyze word frequencies allows us to dig deeper and think about the many usage patterns that we wouldn't otherwise observe. While relying on a text analysis tool for these sorts of conclusions feels a bit nebulous, future work could clarify and extend present findings. Next, we plan to dig deeper into the text data by running *exact match* and *stemmed word* queries for those titles with 50 or more uses included in large platforms such as Springer, Ebrary, and EBSCO. Our preliminary analysis convinced us that words like "edition," "volume," and "2d" should be added to the stop list. They appear high in some e-book collections, and not at all in others, which may point to differences in the way databases formulate their titles as opposed to differences in the content of e-book collections. We will carry out formal statistical analysis to investigate the rank correlation and measure the relationship between search terms and e-book titles to assess the significance of the relationship between them.

References

- Ashcroft, L. (2011). E-books in libraries: An overview of the current situation. *Library Management*, 32(6), 398-407.
- Bailey, T. P. (2006, January). Electronic book usage at a master's level 1 university: A longitudinal study. *Journal of Academic Librarianship*, 32(1), 52 – 59.
- Berg, A. S., Hoffmann, K., & Dawson, D. (2010, November). Not on the same page: Undergraduates' information retrieval in electronic and print books. *The Journal of Academic Librarianship*, 36(6), 518-525. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0099133310002168>
- Christianson, M. (2005, December). Patterns of use of electronic books. *Library Collections, Acquisitions, & Technical Services*, 29(4), 351-363. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1464905506000042>
- Levine-Clark, M. (2007). Electronic books and the humanities: A survey at the University of Denver. *Collection Building*, 26(1), 7-14. Retrieved from <http://search.proquest.com/docview/219889113?accountid=10226>

- Littman, J., & Connaway, L. S. (2004). A circulation analysis of print books and e-books in an academic research library. *Library Resources and Technical Services*, 48, 256-262. Retrieved from <http://www.oclc.org/content/dam/research/publications/library/2004/littman-connaway-duke.pdf?urlm=162892>
- Mays, A. (2014, January). Biz of acq —PDA, e-books, print books usage and expenditures: Knowledge ecosystem remix. *Against the Grain*, 25(6), 61-64.
- Noorhidawati, A., & Gibb, F. (2008). How students use e-books – reading or referring? *Malaysian Journal of Library and Information Science*, 13(2), 1-14. Retrieved from <http://strathprints.strath.ac.uk/19896/>
- Shelburne, W.A. (2009, June). E-book usage in an academic library: User attitudes and behaviors. *Library Collections, Acquisitions, and Technical Services*, 33(2009), 59-72.
- Staiger, J. (2012, June). How e-books are used: A literature review of the e-book studies conducted from 2006 to 2011. *Reference & User Services Quarterly*, 51(4), 355 – 365. Retrieved from <http://rusa.metapress.com/content/u084729074582u77/fulltext.pdf>

C.2. Exploring E-Book Use through Text Analysis

The following paper was published in the peer reviewed journal *Performance Measurement and Metrics* in November 2015.

Citation: Bakkalbasi, N. and Goertzen, M. (2015, November). Exploring academic e-book use: Part I through text analysis. *Performance Measurement and Metrics*, 16(3), 252-262.

Exploring academic e-book use: Part I through text analysis

Nisa Bakkalbasi (Assessment Coordinator)

Melissa Goertzen (E-Book Program Development Librarian)

Columbia University Libraries, New York, New York, USA

Abstract

Purpose

Over the past decade, as the e-book collection continues to grow, Columbia University Libraries (CUL) has been gathering information to develop policies related to e-book acquisition, discovery, and access. This paper investigates users' e-book search behavior and information needs across different disciplines.

Methodology

The research method utilizes text data from two sources: users' e-book search queries that were entered into the libraries discovery tool called CLIO and e-book title words provided by the COUNTER usage reports. The analysis involves identifying and quantifying certain words from users' search queries with the purpose of examining the contexts within which these words were used.

Findings

The prominence of topical words such as "history," "social," and "politics" in the list was an interesting reflection on the kinds of works users were looking for, as were the terms "handbook," "guide," and "manual." The high frequency of these words imply that users were searching for broad topics, reference works, or other collections of instructions, all of which are intended to provide ready reference.

Originality and value

Running search queries and e-book title words through a text analysis tool revealed new ideas related to what types of materials users search for and use. Text analysis of search terms and title words provided insight into the nature of e-book use, including broad topic (e.g. history), academic level of use (e.g. introductory), and genre/type (e.g. reference). While it is challenging to deduce reader intent from word frequency analysis, as text data remain widely open for

interpretation, the methodology has significant strengths that drive us to continue to use in future studies.

Introduction

Over the past decade, electronic books (e-books) have become increasingly popular in the academic community. In response to this demand, Columbia University Libraries (CUL) provides access to over two million e-books that support research, teaching, and learning activities across campus and within the wider scholarly community. As the e-book collection continues to grow, CUL has been gathering information to develop policies related to e-book acquisition, discovery, and access. This paper reports on the findings of one of the multiple methods used in a yearlong study to understand different aspects of e-book related issues.

Data collected through COUNTER usage statistics and the LibQUAL+ service quality survey tells 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. In a preliminary study, Bakalbasi and Goertzen utilized text analysis to explore e-book usage across disciplines. They found that e-books have grown in popularity in all disciplines not just in the sciences and social sciences. Given the highly diverse research needs of the university community, the authors recommended a larger and more robust study to understand scholarly e-book usage in different disciplines.

Literature Search

Determining how e-books are used for scholarly purposes is a complex issue. The e-book landscape is evolving at a rapid pace and a wide range of factors impact how user communities discover and access e-books for research, teaching, and learning activities. It is more important than ever for librarians to understand when, how, and why clients use e-books in order to design services that meet existing needs.

A number of studies have been conducted to understand how e-books are used for research, teaching, and learning activities. Shelburne (2009) conducted a large scale survey to learn about e-book usage patterns at the University of Illinois. In total, 1,547 responses were received. The results indicated that 78 percent of e-book use was intended for research purposes, 56 percent for study, 2 percent for teaching, and 2 percent for other purposes.

Levine-Clark (2007) found that library users typically “use rather than read” e-books. Typically, the format is viewed as a convenient source that provides quick reference for scholarly endeavors. Results from a survey of 2,067 faculty, students, incoming students, and alumni indicated that 56 percent of respondents use e-books to read a chapter or article within a book, and 36 percent typically read a single entry or several pages.

Noorhidawati and Gibb (2008) and Berg, Hoffman, and Dawson (2010) suggest that e-books are primarily used for quick reference, limited reading, and citation checks as opposed to extended reading and research. In other cases, e-books serve as a convenient means to preview a text; students and faculty members peruse the e-version to gain a sense of the information, biases, or arguments presented in a scholarly monograph. If it is useful for their research purpose, a print version is often requested for extended reading.

In 2010, the University of Illinois at Urbana-Champaign (UIUC) Library took part in a global study of Elsevier e-books sponsored by Elsevier Publishing. The goals were to examine the e-book discovery process, cost and usage trends, and the perceived value of Elsevier e-books to the user community. Findings indicated that e-book usage more closely resembles online serials than print books (Chrzastowski, 2011). The author attributed this trend largely to the differences of accessibility in electronic versus print collections. Also, users' online behaviors when using e-book collections included fact checking, brief reviews of content, online reading, and the ability to download e-books in PDF format.

A literature review by Staiger (2012) compared the results of two dozen studies regarding e-book usage by members of the academic community. Findings suggested that "academic users typically search e-books for discrete bits of information, a behavior summed up by the formula 'use rather than read'" (p. 355). In general, members of the academic community do not immerse themselves in e-books for extended periods of time to examine entire arguments. Instead, they view e-books as "convenient sources from which to extract information for their scholarly endeavors" (p. 357). Essentially, e-books provide a means for power browsing. They allow users to preview a book without leaving their workstations, and then locate the print copy if the information is relevant to their studies (p. 358). A literature review by Ashcroft (2011) uncovered similar trends. Statistics showed that on average, "53.5 percent of students and 58.6 percent of teachers dipped in and out of several chapters, whereas very low percentages read the whole book – 5.5 percent of students and 7.1 percent of teachers" (p. 401).

Research Design

Before discussing the research design in detail, it is worth mentioning that our initial thought was to develop a campus-wide survey. However, three key factors influenced our assessment strategy and motivated us to tap into existing data sources. First, during our consultations, it became apparent that using a low-overhead data collection technique that allowed us to systematically collect information over time would be most appropriate for this project. Due to our interest in continuously monitoring our user base in an ever-changing e-book landscape, reliance on readily available, continuous, and accurate data was an important factor in creating a sustainable assessment plan.

Second, as survey participation rates have declined, survey research has experienced significant challenges that impact its use in library assessment plans. Participating in a survey and providing

thoughtful, reflective feedback requires time and effort from respondents. The quality of the data begins to deteriorate when potential respondents do not make the effort to submit a completed survey. Surveys are of little or no use, if the response rate is low or the data is inaccurate. Based on the low response rates from a recent survey, and in an attempt to avoid survey fatigue, we investigated alternative approaches of data collection.

Finally, we know firsthand that librarians constantly manage a large number of responsibilities and projects in a given academic year. With this in mind, we wanted to create an assessment method that is flexible time-wise and manageable as we juggle multiple priorities throughout the academic year.

In this study, we sought an innovative research method to understand users' e-book search behavior, preferences, and information needs across different disciplines. This method utilizes text data from two sources: users' e-book search queries that were entered into the libraries discovery tool called CLIO¹ and e-book title words provided by the COUNTER² usage reports. The study covers a one-year period from January 1, 2014 through December 31, 2014. For the indicated time period, we exported all Page Path 1 URLs, which include search queries limited to e-book format (e.g. , /e-books?q=cultural+psychology+2nd+edition+2011&commit=Search) from the Libraries Google Analytics account. For the same period, we exported COUNTER Book Report 2 (R4) Number of Successful Section Requests by Month and Title from six major e-book platforms, that is Cambridge Books Online, Ebrary, EBSCOhost, ScienceDirect, Oxford Scholarship Online, and SpringerLink, to ensure e-books were included from all three major disciplines, namely humanities, social sciences, and sciences.

We began analysis by identifying and quantifying certain words or content from users' search queries with the purpose of exploring usage, and completed our analysis by examining the contexts within which these words were used. First, using the mixed method analysis software tool, NVivo, we ran two word frequency queries using two different text match options; exact-match and stemmed-word match (grouping the words with the same stem), to generate lists of the most frequently occurring search terms and e-book title words. This quantification was an attempt not to infer meaning but, rather, to explore usage. Then, we evaluated details of the highly occurring words within text to understand the meaning, purpose, and significance of these words, and its implications for e-book search, discovery, and delivery process.

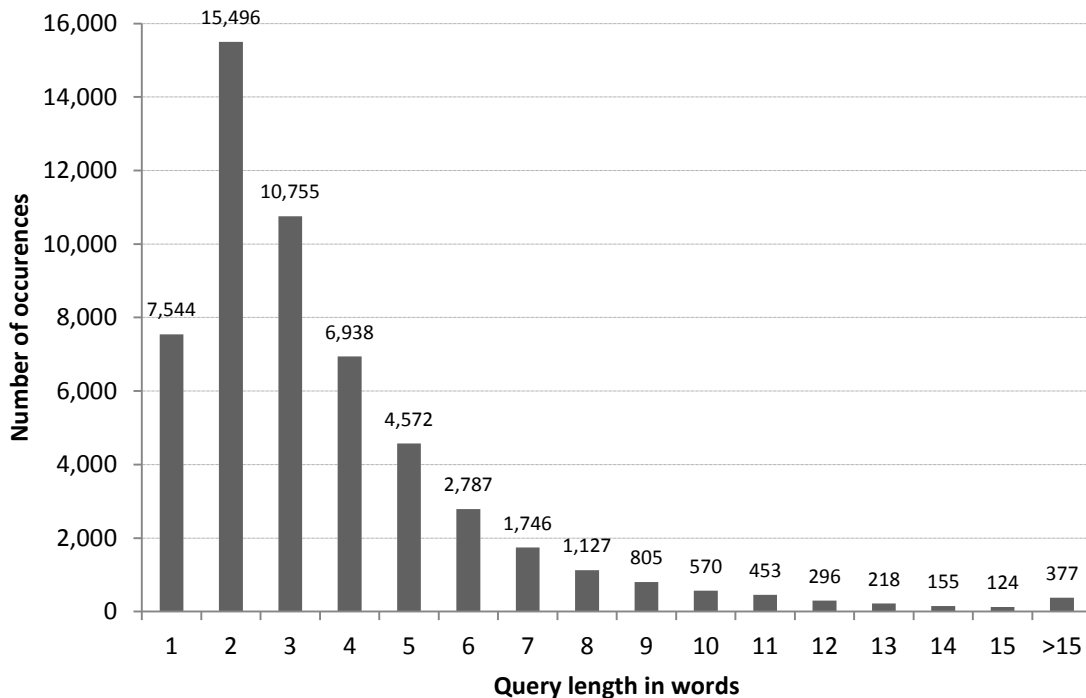
¹ The discovery tool CLIO covers the Library Catalog (including books, journals, video & audio recordings, and archival collections), articles in journals and newspapers, Academic Commons (digital repository of Columbia research), and Library Website. See, <https://clio.columbia.edu/>

² Launched in March 2002, COUNTER (**C**ounting **O**nline **U**sage of **N**etworked **E**lectronic **R**esources) is an international initiative serving librarians, publishers and intermediaries by setting standards that facilitate the recording and reporting of online usage statistics in a consistent, credible and compatible way. For more information, please see: <http://www.projectcounter.org/>

Results

In order to understand e-book search behavior, we began our analysis by calculating ‘*search query length*’ that counts the number of words in a search query, where a word is defined as a string of characters delimited either by a space or by the end of the query. We found that the average e-book search query length was 3.62 words, which was higher than the average search query length reported for major search engines (e.g. Google). The distribution of the query lengths was also rather different. In “Keyword and Search Engine Statistics” report published by KeywordDiscovery.com, a keyword research service provider, English users predominantly use 2 and 3 word search terms, with single keywords becoming less common searches. In CLIO e-book search queries, single word searches were much lower (7,544, 14%) than the single word searches reported for major search engines (29.53%) in the same report. The query length distribution is displayed in Figure 1 in detail.

Figure 1: Query length distribution represented in numbers



To investigate what longer, possibly more complex, queries look like, we ranked the search queries from the longest search query string to shortest search query string, and visually scanned for patterns. The visual scan revealed that a substantial number of search queries that included six or more search terms, were for known items. These search queries included highly specific information such as edition number, volume number, ISBN, and author names. The presence of punctuation marks, capitalizations, and initializations suggested that the users copied and pasted citations from documents and/or other resources into the search box (e.g. “Reamer, F.G. (2006).

Social work values and ethics“). Longer search strings may also indicate roughly 20% of our users are looking for known items. Similar conclusions would likely have been arrived at using manual coding techniques (i.e., content analysis), but this would have been more time consuming.

To gain insight into exactly what types of e-books users are searching for, we parsed all the search query strings (53,963) into individual query terms (152,637). We obtained a total number of 26,045 unique query terms, which excluded default English (US) stop words. Stop words are natural language words which have very little meaning, such as "and", "the", "a", "an", and similar words. After our preliminary analysis, we added words such as “volume,” “edition,” “2nd,” and “3rd” to the stop word list as it was not useful for analyzing informational content about the type of e-books that users were searching for.

We ran two word frequency queries using both exact-match and stemmed-word match options to generate a list of most frequently occurring e-book search terms. Quantitative text analysis provided insight into users e-book needs and preferences. The prominence of topical words such as “history,” “social,” and “politics” in the list was an interesting reflection on the kinds of works users were looking for, as were the terms “handbook,” “guide,” and “manual.” The high frequency of these words lead us to believe that users were searching for broad topics, reference works, or other collections of instructions, all of which are intended to provide ready reference.

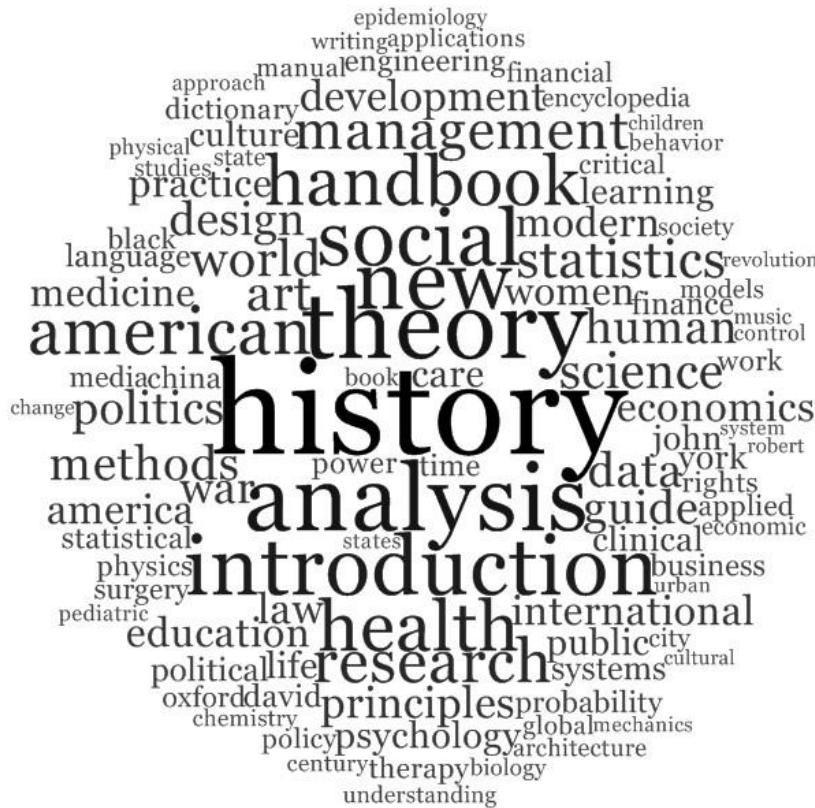
Rank	Exact-match word	Count	Stemmed-word	Count	Similar Words
1	history	1,062	history	1,096	histories, history
2	theory	787	theory	853	theorie, theories, theory
3	analysis	766	analysis	766	analysis
4	new	681	new	681	new
5	introduction	669	introductions	672	introduction, introductions
6	social	638	statistics	662	statistic, statistical, statistics
7	health	566	socialization	660	social, socialism, socialization, socializing
8	handbook	556	politics	620	polite, politeness, political, politics
9	american	539	americans	570	american, americanizing, americans
10	research	497	health	566	health
11	management	450	handbooks	561	handbook, handbooks
12	statistics	419	managment	529	managed, management, manager, managers, managing, managment
13	art	400	economics	514	econom, economic, economical, economics
14	world	399	research	510	research, researcher, researchers, researches, researching
15	science	397	practicing	490	practical, practice, practices, practicing
16	methods	383	sciences	481	scienc, science, sciences
17	data	381	cultures	474	cultural, culturally, culture, cultures, culturing
18	human	376	models	450	model, modeling, modell, modelling, models
19	edition	365	arts	441	art, arte, artful, arts
20	politics	365	methods	433	method, methode, methods
21	design	364	worlds	425	world, worlding, worldly, worlds
22	guide	364	internment	420	intern, internal, internalized, internally, internals, international, internationalism,

					internment
23	war	352	designs	416	design, designation, designer, designers, designing, designs
24	principles	341	systems"	416	system, systemic, systemically, systems, systems"
25	law	326	humans	414	human, humane, humanism, humanities, humanity, humans

Table 1: Top 25 frequently occurring e-book search terms ranked from highest to lowest.

Using NVivo, we also created word clouds to graphically display what types of e-books users were searching for, at least in terms of broad topic, level of academic use, and genre. A word cloud generated by using exact match option for the top one thousand most frequently occurring search terms implied that users across all three major disciplines, namely humanities, social sciences, and sciences, were searching for e-books.

Figure 2: Word cloud for e-book search terms generated by using top one thousand words



To determine how well our search term findings correlate with what items were actually being used, we conducted a similar analysis, using COUNTER Book Report 2 data. Based on the premise that the expected role of a book title is to provide a compact summary of the book, we ran word frequency queries to obtain a list of most frequently occurring words. The prominence of topical words such as “history” in both lists was an interesting reflection on the kinds of works being used, as were the terms “handbook,” “guide,” and “manual.”

Rank	Exact-match	Count	Stemmed-word	Count	Similar Words
1	history	2,508	cultures	3,047	cultur, cultural, culturalism, culturality, culturally, culture, culture', culture', cultured, cultures
2	studies	2,260	studying	2,758	studi, studie, studies, study, studying
3	american	2,179	history	2,648	historie, histories, history, history'
4	new	2,175	politics	2,563	politeness, politic, political, politically, politics, politics'
5	theory	2,119	theory	2,316	theorie, theories, theory, 'theory, theory'
6	culture	1,908	americans	2,294	american, americanism, americanisms, americanity, americanization, americanizing, americans, americans'
7	advances	1,747	systems'2014	2,245	system, system', systemic, systemism, systems, systems'2014
8	systems	1,738	advancing	2,186	advance, advanced, advancement, advancements, advances, advancing
9	social	1,737	new	2,178	new, 'new, 'new'
10	world	1,731	computing	2,070	computability, computable, computation, computational, computationally, computations, computed, computer, computers, computes, computing, computational
11	politics	1,705	moderns	1,999	modern, moderne, moderner, modernism, modernisms, modernities, modernity, modernization, modernizes, modernizing, moderns
12	analysis	1,669	worlds	1,862	world, worlde, worldly, worlds
13	handbook	1,555	science	1,840	science, science', sciences, sciences'
14	science	1,528	socialism	1,804	social, sociales, socialism, sociality, socialization, socializing, socially
15	modern	1,479	analysis	1,669	analysis
16	research	1,475	practicing	1,624	practical, practicalities, practically, practice, practiced, practices, practicing
17	international	1,460	handbook	1,599	handbook, handbooks
18	applications	1,372	developments	1,562	develop, developed, developer, developers, developers', developing, development, developments
19	series	1,253	internment	1,517	internal, internalization, internally, internals, international, internationalism, internment
20	management	1,234	research	1,512	research, researcher, researchers, researches, researching
21	development	1,228	managing	1,507	manage, managed, management, manager, managers, manages, managing
22	guide	1,207	applications	1501	applicable, application, applications
23	century	1,197	states'	1,374	state, state', stated, states, states'
24	health	1,191	globally	1,372	global, globalism, globalization, globalizations, globalized, globalizing, globally
25	america	1,180	humans	1,317	human, humane, humanism, humanities, humanity, humanity', humanization, humanized, humanizing, humans

Table 2: Top 25 frequently occurring e-book title words ranked from highest to lowest.

When we evaluated the word clouds for e-book title words, a similar trend emerged (see Figure 2). Words like “history,” “american,” “culture,” “politics,” “social” and “theory” are situated at the center of the clouds, meaning that they have high frequencies in both lists.

Financial Time Series (3rd Edition)			
Managing Organizational Behavior	Greenwood Publishing Group, Incorporated	EBRARY	8,621
Spatially Integrated Social Science	Oxford UPress	EBRARY	8,528
Principles of Turbomachinery	John Wiley & Sons	EBRARY	8,486
Japan to 1600: A Social and Economic History	U of Hawaii Press	EBRARY	8,454
Craft of Research (3rd Edition)	U of Chicago Press	EBRARY	7,853
PCR Cloning Protocols	Springer	SpringerLink	7,394
Wealth and Welfare States: Is America a Laggard or Leader?	Oxford U Press	EBRARY	7,292
Information Security and Privacy	Springer	SpringerLink	7,065
Effective Public Manager: Achieving Success in Government Organizations (5th Edition)	John Wiley & Sons	EBRARY	6,582
Business and Environmental Policy: Corporate Interests in the American Political System	MIT Press	EBRARY	6,340
Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics	MIT Press	EBRARY	6,269
Prenatal and Postnatal Care	John Wiley & Sons	EBRARY	6,264
Jossey-Bass Public Health : Health Communication : From Theory to Practice (2nd Edition)	John Wiley & Sons	EBRARY	6,181
Birth of the Universe and Fundamental Physics	Springer	SpringerLink	6,072
Human Behavior : A Cell to Society Approach	John Wiley & Sons	EBRARY	6,047
Gender Trouble : Tenth Anniversary Edition	Routledge	EBRARY	5,986
Current Therapy : Current Surgical Therapy (11th Edition)	Elsevier - Health Sciences Division	EBRARY	5,577
Handbook of Psychology, Volume12 : Industrial and Organizational Psychology (2nd Edition)	John Wiley & Sons	EBRARY	5,558

Table 3: Most frequently requested e-book titles from the COUNTER Book Report 2 (R4) – Number of Successful Section Requests by Month and Title Main Section Type: Book Chapter.

Finally, we looked at the overlap between the e-book search terms and e-book title words. We found a 70% overlap between the exact-match search words and title words in the top one thousand most frequently occurring words, indicating a correlation between search and discovery of e-books

Rank	Search Term	Search Term Count	e- book title word	e-book title word Count
1	<i>history</i>	1,062	<i>history</i>	2,508
2	<i>theory</i>	787	studies	2,260
3	<i>analysis</i>	766	<i>american</i>	2,179
4	<i>new</i>	681	<i>new</i>	2,175
5	<i>introduction</i>	669	<i>theory</i>	2,119
6	<i>social</i>	638	<i>culture</i>	1,908
7	<i>health</i>	566	advances	1,747
8	<i>handbook</i>	556	<i>systems</i>	1,738
9	<i>american</i>	539	<i>social</i>	1,737
10	<i>research</i>	497	<i>world</i>	1,731
11	<i>management</i>	450	<i>politics</i>	1,705
12	statistics	419	<i>analysis</i>	1,669
13	<i>art</i>	400	<i>handbook</i>	1,555

14	world	399	<i>science</i>	1,528
15	<i>science</i>	397	<i>modern</i>	1,479
16	<i>methods</i>	383	<i>research</i>	1,475
17	data	381	<i>international</i>	1,460
18	human	376	applications	1,372
19	<i>politics</i>	365	series	1,253
20	<i>design</i>	364	<i>management</i>	1,234
21	<i>guide</i>	364	development	1,228
22	<i>war</i>	352	<i>guide</i>	1,207
23	principles	341	century	1,197
24	<i>law</i>	336	<i>health</i>	1,191
25	development	329	america	1,180
26	economics	326	human	1,130
27	<i>education</i>	324	<i>life</i>	1,110
28	<i>international</i>	322	<i>education</i>	1,106
29	america	320	women	1,105
30	<i>modern</i>	319	<i>law</i>	1,098
31	practice	313	<i>war</i>	1,079
32	medicine	312	global	1,073
33	women	309	<i>methods</i>	1,031
34	care	303	<i>introduction</i>	986
35	psychology	291	philosophy	965
36	public	287	practice	950
37	<i>life</i>	286	<i>engineering</i>	944
38	learning	279	language	925
39	<i>culture</i>	270	early	907
40	clinical	267	<i>design</i>	893
41	john	267	information	866
42	york	266	learning	857
43	<i>political</i>	251	<i>political</i>	840
44	<i>systems</i>	250	cultural	839
45	david	246	technology	827
46	finance	239	society	818
47	language	239	literature	814
48	power	238	<i>art</i>	770
49	statistical	234	public	763
50	<i>engineering</i>	233	china	760

Table 4: Overlap between the top 50 frequently occurring search term and e-book title word ranked from highest to lowest.

Conclusions and Discussion

The ability to analyze e-book search queries allowed us to discover many search patterns that we wouldn't otherwise observe. For instance, we discovered that the users' searching behavior when using a library catalog is different than users' searching behavior on the Web. The average search query length when searching CLIO (m=3.62 terms) compared to the average search engine query length (m=2.4 terms) may indicate that library users are more sophisticated in how they structure their queries when they are looking for very specific items or specific answers. In

a future study, it would be interesting to investigate the search terms without limiting it by format.

Running e-book search queries and requested e-book title words through a text analysis tool revealed new ideas and concepts related to what types of materials our users search for and use. Text analysis of search terms and requested title words provided insight into the nature of e-book use across disciplines, including broad topic (e.g. history), academic level of use (e.g. introductory), and genre/type (e.g. reference). It is challenging to deduce reader intent from word frequency analysis, as text data remain widely open for interpretation. However, responses to open-ended questions from the most recent LibQUAL+ survey are consistent with our findings that e-book collections are widely used across all major disciplines to support instruction and learning. User sentiments from the LibQUAL+ survey mirrored a number of findings mentioned in the literature review, namely by Levine-Clark (2007) and Shelburne (2009), who suggested that e-books were used primarily to read chapters or articles for study purposes.

While relying on a text analysis tool for these sorts of conclusions felt a bit nebulous, future work could clarify and extend present findings. Search phrases provided a good deal of information about what types of e-books users search for, but much less information about why searches are conducted or how satisfied users are with the discovery process as a whole. This knowledge gap must be taken into account in analyses and complemented by other techniques to provide a more complete understanding of search behaviors. Despite its limitations, there are two significant strengths of the methodology that drive us to utilize it in future studies: (1) real queries from real users with actual information needs and usage, and (2) the volume of the data analyzed and time-span it covered.

Notes

“Keyword Usage Statistics on the Average Number of Keywords per Search Phrase by Country”, available at: <http://www.keyworddiscovery.com/keyword-stats.html> (accessed 5 August, 2015).

References

- Ashcroft, L. (2011), “Ebooks in libraries: an overview of the current situation”, *Library Management*, Vol. 32 No. 6/7, pp. 398-407.
- Bakkalbasi N. and Goertzen M. (2014), “Scholarly E-Book Use Across Disciplines: Content Analysis of Usage Reports and Search Terms”, in Fiddler A. and Kendrick C, (Eds.), *Proceedings of the CUNY Library Assessment Conference, New York, NY*. Available at: <http://www.cuny.edu/libraries/conference/proceedings.html> (accessed 5 August, 2015).
- Berg, A. S., Hoffman, K. and Dawson, D. (2010), “Not on the same page: Undergraduates’ information retrieval in electronic and print books”, *The Journal of Academic Librarianship*, Vol. 36 No. 6, pp. 518-525.

- Chrzastowski, T. (2011), "Assessing the value of e-books to academic libraries and users. *Proceedings of the 9th Northumbria International Conference on Performance Measurement in Libraries and Information Services, UK*, 53-62. Available at: <https://www.york.ac.uk/about/departments/support-and-admin/information-directorate/northumbria-conference/9th-conference/proceedings/> (accessed 5 August 2015).
- Levine-Clark, M. (2007), "Electronic books and the humanities: A survey at the University of Denver", *Collection Building*, Vol. 26 No. 1, pp. 7-14.
- Noorhidawati, A. and Gibb, F. (2008), "How students use e-books – reading or referring?", *Malaysian Journal of Library and Information Science*, Vol. 13 No. 2, pp. 1-14.
- Shelburne, W.A. (2009), "E-book usage in an academic library: User attitudes and behaviors", *Library Collections, Acquisitions, and Technical Services*, Vol. 33 No. 2-3, pp. 59-72.
- Staiger, J. (2012), "How e-books are used: A literature review of the e-book studies conducted from 2006 to 2011", *Reference & User Services Quarterly*, Vol. 51 No. 4, pp. 355 – 365.

D.1. Faculty Interview and Student Focus Group Questions

1. Tell us about your experiences using e-books.
2. Thinking about the past academic year, how often have you used e-books?
Possible follow-up question: What are some of the advantages/disadvantages of using e-books?
3. When you use e-books, electronic articles, etc., what technologies or devices do you most often use? (e.g. PC, e-reader, smart phone, etc.)
4. When you want to use e-books for academic purposes, where do you search for/locate e-books? (e.g. through CLIO, Google, Amazon)
5. When you're using an e-book for academic work, what are three features that are most important to you? (e.g. ability to download chapters/entire book, copy and paste text, take notes, highlight)
6. Thinking about the past academic year, have you used an e-book from the university library?
Tell us about your experience.
7. Is there anything that Columbia University Libraries can do to improve e-book services or collections?
8. Have we missed anything?

D.2. Focus Group Questionnaire

Q1. Do you own any of the following items?

	Yes, I currently have one.	No, but I plan to purchase one within 12 months.	No, and I don't plan to purchase one within the next 12 months.
Android phone			
BlackBerry			
iPhone			
Other smart phone			
iPad			
Tablet (e.g. Nexus)			
Kindle			
Kobo			
Nook			
iOS			
Other e-reader			
Laptop computer			
Desktop computer			

Q2. Thinking about the past year, where did you search for e-books whether it was for academic or personal purposes?

	Several times a day 1	Once a day 2	A few times a week 3	Less often 4	Don't use 5
Columbia Library Catalog (CLIO)					
Search engine (e.g. Google, Yahoo)					
Google Book Search					
E-book platform (e.g. ebrary)					
Database (e.g. EBSCO)					
Publisher website (e.g. SpringerLink, Cambridge Books Online)					
Repository (e.g. HathiTrust, Project Gutenberg)					
Public library					
Online library (e.g. Oyster, Kindle Owners' Lending Library)					
E-book app (e.g. iBooks)					
E-book store (e.g. Amazon)					

Q3. Thinking about the past academic year, what materials have you used for academic work?

	Print format	Electronic format	Audio or video format	Not used
Books				
Reference Sources (e.g. dictionary, encyclopedia)				
Journals/Serials				
Dissertations				

Q4. Thinking of the past year, how many e-books appeared on your course reading lists?

- None
- 1 – 5
- 6 – 10
- 11 – 15
- 16 or more

Q5. What types of e-books would you like Columbia University Libraries to offer?

- Academic/peer reviewed titles
- Non-fiction titles
- Fiction titles
- Best sellers
- Other (please specify)

Thank you for participating in the E-Book Focus Group. Your feedback is very much appreciated.

D.3. Exploring E-Book Discovery and Use: Focus Groups & Interviews

Note: The following paper is scheduled for publication in the peer reviewed journal *Performance Measurement and Metrics*. Date: Winter 2016.

Exploring academic e-book use: Part II through focus groups and interviews

Melissa Goertzen (E-Book Program Development Librarian)

Nisa Bakkalbasi (Assessment Coordinator)

Columbia University Libraries, New York, New York, USA

Abstract

Purpose

In the paper entitled *Exploring Academic E-book Use: Part I through Text Analysis*, Bakkalbasi and Goertzen (2015) discussed a text analysis method utilized to explore e-book usage across disciplines at Columbia University. To verify the method, the authors conducted focus group and interviews sessions with faculty members and graduate students to understand *for what purpose* and *why* e-books are used in conjunction with scholarly activities.

Methodology

Participants answered eight pre-determined questions during one-on-one interviews and dual moderator focus group sessions. They were also invited to complete a questionnaire regarding e-book discovery, access, and use. All sessions were transcribed and the data was analyzed using grounded theory approach to examine emerging themes.

Findings

The findings suggest that faculty and graduate students use e-books for discontinuous reading and quick reference purposes. They value the ability to customize learning environments to suit immediate circumstances and needs. Frustration occurs when availability and accessibility are hindered by limitations imposed by platforms or licenses. Participants believe the library can advocate for users and work with vendors to develop business models that promote greater convenience and flexibility online.

Originally and value

The study complements and extends existing findings reported in the first part of this study (Bakkalbasi & Goertzen 2015) that utilized a text analysis method. The results indicate that text analysis is a reliable assessment method in the examination of usage trends across e-book collections. Also, the study brings a human sentiment to the discussion of e-book discovery, access, and use. It provided the user community with a voice and left the authors with a deeper understanding of existing information needs on campus.

Introduction

Since 2010, there has been marked growth in e-book holdings at Columbia University Libraries (CUL). At the present time, the Libraries provide access to over two million e-books and expenditures comprise 25 percent of the book budget. In response to this growth, CUL is developing a unique strategy and vision for e-book programs and initiatives across campus. This includes assessing how users discover, access, and use e-books to support scholarly activities.

The insight and feedback gained from observing and discussing information seeking behaviors will inform the future of e-book collection development activities and programs at CUL.

In the paper *Exploring Academic E-book Use: Part I through Text Analysis*, Bakkalbasi and Goertzen (2015) discussed a text analysis method utilized to explore e-book usage across disciplines. The method relied on data from two sources: users' e-book search queries that were entered into CUL's discovery tool called CLIO and e-book title words provided by COUNTER usage reports. The findings provided insight into the nature of e-book use and suggested that the format is used primarily for reference and instructional purposes. The authors determined that the strength of the text analysis methodology was the ability to examine information needs from real users based on real queries entered into CLIO over the course of one calendar year. However, they also determined that one knowledge gap remained: the ability to examine the intent behind observed information seeking behaviors.

In order to provide an additional layer of meaning to the results of the text analysis study, we set out to understand *for what purpose* and *why* readers use e-books in conjunction with research, teaching, and learning activities. In order to answer these questions, a series of focus group and interview sessions with faculty and graduate students based at Columbia University were conducted during the summer semester of 2014. The participant group represented three major divisions within the library system (i.e. humanities, social sciences, and sciences).

The focus group and interview sessions provided two distinct advantages that complement our previous work. First, the research described in this paper adds a human element to the investigation surrounding e-book discovery and access - it provided the user community with a voice and allowed us to uncover reactions or attitudes that are not apparent when relying solely on quantitative data. Second, the body of qualitative data collected through the focus group and interview sessions facilitated validation of previous results through cross-examination and verification using another method (Bakkalbasi & Goertzen 2015). Results indicate that the text analysis methodology is a reliable assessment method in the examination of usage trends across the CUL's e-book collections.

Literature Review

As the information environment evolves, it is more important than ever for librarians to connect with user communities and understand existing information needs. Opportunities to observe and discuss information seeking behaviors with users themselves are invaluable. Knowledge regarding their preferred means to discover, access, and use e-book can lead to the development of services that provide new opportunities for the integration of electronic content into research, teaching, and learning environments at academic institutions.

Over the past several years, a number of studies were conducted to determine the advantages and challenges users associate with e-book collections. Armstrong and Lonsdale (2009) and Beisler and Kurt (2012) suggest that the main benefits linked to e-book use include twenty-four hour availability of materials, remote access, and multiple user access options. In most cases, the challenges relate to issues that restrict access and functionality options including Digital Rights Management (DRM), platform design, and the incompatibility of file formats with various e-readers and technologies.

In a literature review that examined findings from e-book surveys, Ashcroft (2011) found that users' expectations regarding e-book use are also linked to accessibility and functionality. The seamless nature of information access available through the commercial market has not yet permeated the academic market. Users want to access e-books using a variety of technologies, including mobile devices, at their convenience (2011). However, licensing issues or DRM create barriers. For instance, some business models limit the number of users that can access individual titles simultaneously. Faculty and students are often unaware of limitations imposed by licenses because they are accustomed to the unrestricted and instant nature of information access through the Internet (2011). This expectation can result in frustration and confusion when users are asked to "wait in line" for e-book content.

To better understand how faculty use e-books in their research and teaching, Arizona State University Libraries held a focus group in the spring of 2007. The participants were asked to comment on their use of e-books as teaching and research tools, disciplinary differences in perceptions of e-books, and motivators for future use (Carlock & Perry 2008). The findings indicate that faculty had unsatisfactory experiences due to unreliability of access, lack of manipulability, and differences in functionality between various interfaces (2008). The authors suggest that academic libraries can provide training to faculty regarding different e-book platforms available through the library, provide better course support, and advocate for faculties' research and teaching needs during discussions with vendors (2008).

In order to examine e-book use and its value to academic users, the University of Illinois at Urbana-Champaign (UIUC) Library was invited by Elsevier Publishing to participate in a global study of e-book use and perceived value (Chrzastowski 2010). The study included user interviews, surveys, and logbook questionnaires to capture users' experiences searching, retrieving, and using e-books on the Elsevier platform. The results indicated that access, the ability to search and navigate through texts, and the ability to download content to laptops ranked as the top advantages associated with e-book use (2010). Participants' online behavior also suggested that users value the ability to briefly look at information, read from the screen, and download PDFs (2010).

Chrzastowski's findings are mirrored in a number of papers that examined how e-books support academic work. In a literature review that examined e-book acquisition and maintenance in academic libraries, Blummer and Kenton (2012) found that users value the ability to navigate between chapters, print sections of chapters, and search content in an efficient manner. For instance, undergraduate and graduate students value the use of indexes, table of contents, and full text search tools when using e-books for research and learning purposes (2012). This behavior could relate to the fact that members of the academic community do not immerse themselves in e-books for extended periods of time to examine entire arguments.

At Miami University, librarians conducted oral interviews with a group of faculty, graduate students, and undergraduates to examine user attitudes and beliefs about e-books. The results indicate that format preference is linked to intent of use. For instance, in cases where a text is intended for leisure reading, participants prefer print books. However, when users require only a portion of a text, they believe that e-books provide advantages in the form of portability and search functionalities (Shrimplin, Revelle, Hurst & Messner 2011).

Findings published in the Ithaka S+R US Faculty Survey 2012 stated that the majority of faculty respondents indicated that “electronic versions of scholarly monographs play a very important role in research and teaching” (Ithaka 2013, p. 31). When asked to consider a variety of activities and indicate how much easier or harder it is to perform each activity in print or digital format, respondents indicated that exploring references and searching for particular topics is much/somewhat easier in digital form than in print. They said that reading in depth and reading cover to cover is much/somewhat easier in print form than digital (2013).

Staiger (2012) discovered that academics “use rather than read” e-books. In many cases, users view e-books as a type of online environment that allows them to search and extract information (2012). Staiger said that academic users have years of experience working with e-journals. In many cases, users expect the same kind of flexibility that they enjoy with articles: the ability to download PDFs, view content on whatever device they choose, and print content without restriction (2012). When they encounter obstacles in these areas, users often become frustrated. The vast majority of challenges related to functionality are not inherent to e-books themselves. Rather, they are the result of restrictions (e.g. DRM) imposed by content owners (2012). This situation places libraries between a rock and a hard place as they address concerns from users without having a direct ability to remedy the situation. However, knowledge of the current e-book landscape and users’ information needs can lead to opportunities for advocacy, and ideally, the development of solutions that are beneficial to both the library and publishing community.

At the University of Oklahoma, librarians invited faculty members working in the pure and applied sciences to complete an online survey following individual interviews featuring e-book demonstrations. The conversations focused on current and potential e-book use for research and teaching purposes (Bierman, Ortega, & Rupp-Serrano 2010). The findings suggest that just as e-journals found acceptance in the academic community, e-books will become more commonplace in the near future due to factors like convenience and ease of navigation. Essentially, users want the ease of use they encounter on the Web to be a regular part of their e-book experiences (2010). Finally, librarians should advocate for the needs of user communities by sharing feedback and concerns with publishers. The development of e-book platforms benefit from observations librarians can provide about the online behaviors of users (2010).

Research Design

Two central goals guided the research design of this investigation. The first was to collect experiential qualitative data sets that provided insight into *for what purpose* and *why* the user community utilizes e-books for scholarly activities, namely research, teaching, and learning. This investigation included uncovering general themes regarding e-book use on campus and across disciplines. The second goal was to empower the user community and give them a voice as decisions are made regarding the future of e-book collection development at CUL.

The time frame we were given to complete this portion of our study coincided with the 2014 summer semester at Columbia University. We realized that this could present challenges in terms of participant recruitment - this is a time of year when faculty and students are frequently away from campus. It became apparent that it would be particularly difficult to coordinate faculty focus group sessions around summer sabbaticals and conference schedules. After careful

consideration, we determined that scheduling one-on-one interviews with faculty members would be the most appropriate method to discuss information needs and general e-book usage. However, we also decided that our sample would contain enough faculty participants to equal in size to at least one full focus group.

After obtaining IRB approval, we began work on the recruitment process. We worked with library administrators to generate a list of faculty members who represented three major disciplines across campus (i.e. humanities, social science, and science). Then, we distributed personalized email invitations until we had reached our target sample size. In total, we met with nine faculty members at their offices and conducted 45 minute interview sessions. Each session was recorded for transcription purposes.

In terms of the student focus groups sessions, we decided to restrict participation to graduate students for two reasons: Columbia University is a research institution and students studying in graduate and professional programs account for approximately 70 percent of the student population, and also graduate students are engaged in research, teaching, and learning activities across campus. To recruit students, we sent email invitations through listservs and connected with student clubs across campus. Then, potential participants were asked to complete a brief screening survey to determine their eligibility for the study (e.g. must be at least 18 years of age, must be enrolled as a full-time or part-time student, etc.). Participants were also offered small incentive prizes. Experience using e-books was not part of the selection criteria – we felt it was equally important to understand why students choose to use or not use e-books to support academic activities. In total, we recruited 11 graduate students for two, 90 minute dual moderator focus group sessions. Each session was conducted in a conference room located in the Butler Library and again, each was recorded for transcription purposes.

We created a list of eight pre-determined questions for use in both the focus group and interview sessions (see Appendix A). They examined popular discovery tools, the frequency and purpose of e-book use, desired functionalities, and experiences using e-books for academic purposes. Following each session, participants were also provided with an opportunity to complete a brief 5 question questionnaire regarding popular technologies, platforms, and discovery tools used to access e-books.

We discovered that a point of saturation was hit fairly quickly - since we were most interested in discovering general themes regarding e-book discover, access, and use across campus, we noticed that the same benefits, challenges, and limitations were discussed in each session. After speaking with the 20 participants (nine faculty members and 11 students), we felt that we had collected sufficient data to make broad statements and verify findings from the text analysis study (Bakkalbasi & Goertzen 2015).

After each focus group and interview session was transcribed, the data was analyzed using grounded theory approach, in which responses are categorized and coded to examine emerging themes. Each response was treated as unit of data and was examined in the context of the following categories: general e-book use, academic use, discovery, functionality, and technology. Word counts were also conducted to determine how many times specific words or phrases (e.g. names of file formats, e-book platforms, etc.) were mentioned. Upon the

completion to data analysis, all recordings and transcriptions were stored on a secure server shared by the two authors to protect the anonymity of study participants.

Results

Reading Technique and E-book Use

One of the central themes that we discovered in both the focus group and interview sessions was that the decision to use e-books for academic purposes was largely dictated by the reading technique most suited to the reading task at hand. These results support findings reported by Shrimplin, Revelle, Hurst & Messner (2011), Blummer and Kenton (2012), and Staiger (2012). At CUL, participants said that they prefer print formats for continuous reading (e.g. prolonged reading without interruption for research or learning purposes, making annotations, etc.) and electronic formats for discontinuous reading and quick reference (e.g. confirming citations, referring to quotes, etc.). Faculty and students indicated that they use e-books at least once a week for latter tasks. When asked to provide examples, a faculty member working in the social sciences said that e-books are used “most often for checking quotes, looking at a chapter, or determining if it’s a reading I want to find out more about so that I can assign it. It’s a preliminary approach, but rarely the final way I read.” A graduate student working in the humanities also stated that “if a source needs to be studied in depth, I need the print. If I’m just skimming something or need to get the general gist of a chapter, I’m happy to use an e-book.”

Above findings corroborate the broad conclusions reported in the first half of our study involving text analysis of search terms and retrieved titles (Bakkalbasi & Goertzen 2015). The results provided insight into the nature of e-books used across disciplines including broad topics (e.g. history), academic level of use (e.g. introductory) and genre/type (e.g. reference). At that time, we also discovered that a substantial number of search queries containing six or more words were for known items; they contained highly specific information such as edition number, volume number, and ISBN.

When we completed the first half of the study (Bakkalbasi & Goertzen 2015), we found it challenging to glean reader intent as the text data remained wildly open to interpretation. However, the above findings are in agreement with results of text analysis (2015). The themes that emerged during the focus group and interview sessions indicate that users do indeed dip in and out of known e-book titles for the purpose of quick reference, introduction to texts or arguments presented in texts, and so on.

The Definition of Convenience

Through discussions with faculty and students, we discovered that the decision to use e-books for discontinuous reading and quick reference is related to a large advantage provided by the format: convenience. These findings support user behaviors noted by Bierman, Ortega, & Rupp-Serrano (2010) and Ashcroft (2011). Based on our discussions with faculty and students at CUL, we determined that convenience is composed of three criteria: content availability (e.g. not having to “wait in line” for an e-book), accessibility (e.g. remote access, availability of multiple file formats, etc.), and usability (e.g. search, print, and download capabilities).

When we pursued the topic of convenience further, we discovered that faculty and students do not have a preference for specific technologies or platforms through which to access e-books.

Rather, they desire a variety of discovery, access, and functionality options that can be tailored to particular circumstances and working environments. Essentially, they desire the ability to customize e-book use to suit a specific set of needs in a given moment. As an example, a faculty member working in the sciences said, “sometimes I want to project content in a classroom...but then at other times, what I really want is portability...to have snatches of time to read, and if I’m doing that, then it’s nice to have a portable device I can take with me.” A graduate student who deals with a lengthy commute commented on the ability to customize e-book use around his travel schedule: “it’s hard to bring physical books with me. If [the library] has a link to an e-book, I generally prefer to use that. The big difference is how the publisher presents the e-book...the clunky platforms don’t let you download chapters. The good platforms let you download a chapter and then you can bring it to class, print it, or search it.”

Perpetual Access to Content and Annotations

When we analyzed the qualitative data collected from interview sessions, we discovered a theme that was entirely unique to faculty members. This is the idea of perpetual access to content and information, not only in regards to e-books themselves, but also in terms of annotations and notes. The faculty members we spoke with do not feel confident that electronic annotations will be preserved for years to come. They said that in many cases, they refer to annotations over the course of their entire career. A professor working in the humanities said, “if I lose annotations down the road, I could potentially lose years of my work. It could vanish and that’s it. That would be devastating.” A faculty member in the social sciences said, “I’ll flirt with the e-book when I need something in the moment, but I’m not going to make a commitment to e-books that may disappear. Imagine working on a book for 12 years, and then CUL changes its subscription service or the publisher goes out of business...you lose your notes and it’s horrible.” If an e-book is pulled from a platform or a publisher ceases to exist, faculty are at risk of losing work that is irreplaceable. In many cases, we found that faculty prefer to annotate print formats because they know the text, along with their work, will remain safe on their shelves for years to come.

Creating Desired E-book Environments for Academic Work

As we conducted the focus group sessions, we also discovered one theme that was unique to graduate students: they are willing and capable of creating e-book experiences that suit their needs, even if this means working around the library. Students indicated that through word of mouth, they often find online learning environments that are suited to their research and learning needs. These can include subscription, foreign, and in some cases, illegal databases of electronic content. For example, a graduate student working in the humanities said,

I’ve downloaded from publishers or Internet sources, like large Russian websites. It’s not just me, a lot of my colleagues and grad students use these sites...for reasons mentioned, like the fact that a lot of ‘real’ e-books are unusable. If someone really wants to use it, they will go through the trouble of unlocking a PDF or removing copyright restrictions, and then they will share it. A lot of people download e-books and share e-books on these e-book databases. A lot of people in the academic community do this.

A second graduate student working in the humanities mentioned that many students pay subscription fees to e-book databases that provide desired functionality and access. When describing one such database, he said,

[The database] is 90 dollars a year, but it gives you access to the full text and they are all digitized. You can search for a topic and it will pull up a list...of academic titles. You can do full text searches, make annotations, and then you can get different citation [styles] so you can drop it into your paper with a quote. You can save the e-books that you're using to your account. I like it because it makes e-books useable. I would like it if you could print the e-books as well, but still, it's pretty good.

The results suggest that student place high value on user experience, access, and ultimately convenience.

The findings suggest that faculty and graduate students use e-books on a regular basis for discontinuous reading and quick reference, and value convenience and the ability to customize research and learning environments to suit immediate needs. They said that it is frustrating when availability, accessibility, and usability are hindered by the limited functionality of academic interfaces and platforms. Participants readily acknowledged that these challenges are not linked directly to the library. However, they said the library can advocate for the user community and work with publishers to develop flexible business models that promote greater convenience and flexibility online.

The combined findings of both the text analysis (Bakkalbasi & Goertzen 2015) and interview and focus group sessions provide us with greater insight into how faculty and students intend to use e-books in conjunction with scholarly activities. For instance, use often supports teaching and learning activities that involve discontinuous reading and quick reference. Based on these finds, we plan to continue discussions with publishers and aggregators in order to advocate for platforms that are DRM free, preserve annotations, and provide access to multiple users simultaneously, particularly in cases when e-book titles are purchased for course reserves. In this way, we hope to build an e-book collection that supports learning activities and removes frustrations associated with access and functionality barriers.

As we continue to build strong relationships with faculty members and students across campus, we plan to solicit feedback as we evaluate existing e-book platforms to determine which provide online environments that best support teaching and learning activities. Now that we have developed a method to examine search and discovery trends (Bakkalbasi & Goertzen 2015) and have established strong channels of communication with the user community, we are in a position to explore evolving information needs on campus. We will continue our explorations by evaluating e-book platforms accessible through CUL to determine which provide online environments and general functionality required by the user community. This will assist in collection development decisions as we consider the future of the e-book collection at CUL and invest resources in materials and services that best support observed information needs.

Conclusions and Discussion

This paper discusses the results from the second half of a year-long study that examined e-book discovery, access, and usage behaviors in our user community. The qualitative data collected through faculty interviews and student focus group sessions provided a greater understanding of the intent behind information seeking behaviors and e-book usage observed in a previous text analysis study (Bakkalbasi & Goertzen 2015). The themes that emerged during the focus group and interview sessions suggest that the text analysis method is a reliable instrument to examine search and discovery trends across e-book collections. More importantly, this study brings a human sentiment to the discussion of e-book discovery, access, and use. It provided our user community with a voice and allowed them to share experiences so that CUL develops a deeper understanding of existing information needs. In this way, librarians can incorporate feedback into service development and proceed with confidence as collection development decisions are made.

Through our discussions with faculty members and students, we were able to identify broad themes regarding e-book discovery, access, and usage trends. In the future, we hope to continue examining these trends in a more granular fashion. For instance, we would like to conduct additional focus groups with undergraduate students in order to document discovery and usage trends. Also, we hope to further experiment with our original data set in order to determine if the text analysis method can be used to compare and contrast search and discovery patterns related to all formats collected at CUL. Through these activities, we will continue to facilitate strong channels of communication with our user community, develop a deeper understanding of information needs across campus, solicit feedback regarding library services, and understand how collection materials support scholarly activities.

References

- Armstrong, C. and Lonsdale, R 2009, 'E-book collection management in UK university libraries: focus groups report', Available from: <http://observatory.jiscebooks.org/reports/e-book-collection-management-in-uk-university-libraries-focus-groups-report/> [2 September 2015].
- Ashcroft, L 2011, 'Ebooks in libraries: an overview of the current situation', *Library Management*, vol. 32, no. 6/7, pp. 398-407.
- Bakkalbasi N. and Goertzen M 2014, 'Scholarly E-Book Use Across Disciplines: Content Analysis of Usage Reports and Search Terms', *Proceedings of the CUNY Library Assessment Conference, New York, NY*. Available from: <http://www.cuny.edu/libraries/conference/proceedings.html> [5 August, 2015).
- Bakkalbasi N. and Goertzen M 2015, 'Exploring academic e-book use: part I through text analysis', *Performance Measurement and Metrics*, vol. 16, no. 3, pp.252-262.
- Bierman, J., Ortega, L., & Rupp-Serrano, K 2010. 'E-book usage in Pure and Applied Sciences'. *Science & Technology Libraries*, vol. 29, no. 1/2, pp. 69-91.

- Beisler, A. and Kurt, L 2012, 'E-book workflow from inquiry to access: facing the challenges to implementing e-book access at the University of Nevada, Reno', *Collaborative Librarianship*, vol. 4, no. 3, pp. 96-116.
- Blummer, B. and Kenton, J 2012, 'Best practices for integrating e-books in academic libraries: a literature review from 2005 to present', *Collection Management*, vol. 37, no. 2, pp. 65-97.
- Carlock, D. M., & Perry, A. M 2008. 'Exploring faculty experiences with e-books: a focus group'. *Library Hi Tech*, vol. 26, no. 2, pp. 244-254.
- Chrzastowski, T 2011, "Assessing the value of e-books to academic libraries and users", *Proceedings of the 9th Northumbria International Conference on Performance Measurement in Libraries and Information Services, UK*. Available from: <https://www.york.ac.uk/about/departments/support-and-admin/information-directorate/northumbria-conference/9th-conference/proceedings/> [15 August 2015].
- Schonfeld, R. C., & Housewright, R 2013, *US faculty survey 2012*. Available from: <http://www.sr.ithaka.org/publications/us-faculty-survey-2012/> [15 December 2015].
- Shrimplin, A. K., Revelle, A., Hurst, S., & Messner, K 2011, 'Contradictions and consensus - Clusters of opinions on e-books'. *College & Research Libraries*, vol. 72, no. 2, pp. 181-190.
- Staiger, J 2012, 'How e-books are used', *Reference & User Services Quarterly*, vol. 51, no. 4, pp. 355-365.

Appendix A: Sample Focus Group and Interview Questions

1. Tell us about your experiences using e-books.
2. Thinking about the past academic year, how often have you used e-books? Tell us about your experience.
3. What are some of the advantages/disadvantages of using e-books for academic purposes?
4. What technologies or devices do you most often use when accessing e-books for academic purposes?
5. When you use e-books for academic purposes, where do you most often search for and/or find e-books?
6. When you use an e-book for academic work, what three features are most important to you?
7. Thinking about the past academic year, have you used an e-book from the university library? Tell us about your experience.
8. Is there anything that Columbia University Libraries can do to improve e-book services or collections?

D.4. The Question of Intent: Building E-Book Collection in the Academic Community

Citation: Goertzen, M. (2015, May 12). [The question of intent: Building e-book collections in the academic community](#) [Web log message]. Written for the Art Libraries Society of North America (ARLIS/NA) Collection Development Blog.

Two years ago I arrived at Columbia University Libraries (CUL) and was tasked with completing the E-Book Program Development Study. The central goal was to document the e-book landscape on campus and within the academic community in order to create collection development recommendations and strategies that will shape future e-book services at CUL. When I read the project description, the opportunity seemed both exciting and daunting – the opportunities for experimentation were enormous, but I knew that the complexities involved with e-book collection development, not to mention the speed at which formats and technologies are evolving, would present significant challenges. In the beginning, I felt like I had to provide solutions that would solve the collection development challenges we face on a daily basis. Now, I believe that discovering the right questions to ask in this evolving landscape is the work that will guarantee the long-term success of collection development policies and strategies.

Our questions set us on a journey of discovery and open our minds to opportunities for innovation. When we challenge our assumptions, look at issues from multiple perspectives, and test what we believe against evidence collected along the way, we begin the process of pulling back the layers of a problem to uncover the golden thread – the core issue that ties together all of the seemingly disconnected elements of an investigation. This process provides a context for research findings and can be the starting point for strategic planning and collaborative relationships that define how collections and services will be delivered in the future.

Several months into the study, I realized that I was operating on the assumption that users prefer electronic content for research, teaching, and learning activities. Throughout society, we rely so heavily on the Internet, mobile technologies, and social media to gather and disseminate information, that I assumed users would have a preference for e-books for all scholarly activities. However, as I started to collect usage statistics, examine discovery and access trends, and speak with faculty, students, and library staff through interviews and focus groups, I realized that my initial impressions of content use were far too simplistic and did not tell the full story. As I dug deeper into data sets, I started to ask more and more questions about when, how, and why users gravitate towards certain formats (e.g. print, electronic, archival materials) to support scholarly activities and build knowledge around specific subject areas.

At first, I worried that the discovery of questions rather than solutions would result in a lengthy list of proposals for future studies. But, the opposite turned out to be true. The inquiry process provided a focus and pulled everything I had observed into one overarching question: what is the intended use of e-book content? For instance, are individual titles requested by the user community for inclusion in course reading lists, research pursuits, general reference, or archival purposes? As an example, I'll refer back to a previous blog post that discussed CUL's strategy to acquire the [Interaction of Color](#) app. After considering the text's value to the academic

community and the unique learning environment provided by the electronic format, CUL wanted to acquire an archival copy to guarantee long-term access. Once the intent was determined, a decision was made not to pursue a license through the iTunes Store (which would only guarantee short-term access) but instead, discuss e-book preservation with the publisher of the app. CUL discovered that academic publishers are also grappling with the issue of preservation and we are now exploring business models that allow for long-term access to enhanced e-book content.

I believe that every e-book collection development decision we must make, ranging from business models, license agreements, acquisition workflows, the marketing of collections, and preservation hinges on the issue of intent of use. Once we have answered this question, we can enter into meaningful conversations with stakeholders, advocate for the needs of users, and develop collections that aim to meet immediate and long-term information needs.

As the E-Book Program Development Study comes to a close, I look forward to setting out on a new journey of exploration and continue advocating for the needs of our users.

E.1. Blog Post: The Fine Art of App Preservation

Citation: Goertzen, M. (2014, December 1). The fine art of app preservation. [Web log message]. Written for the Art Libraries Society of North America (ARLIS/NA) Collection Development Blog

In 1963 Josef Albers published the *Interaction of Color*, which revolutionized the way scholars and students study relationships between colors. Albers developed the text to serve as a hands-on kit that provides an interactive environment to conduct silkscreen color studies. However, there are limits to interactivity, as the book was originally published in print format. Now, fifty years after its original publication the text has been released in a digital format that brings Albers' original vision to life. Yale University Press released the *Interaction of Color* app, which includes over 125 color plates, 60 interactive studies, and commentary from Albers. This is a landmark publication and won the George Wittenborn Memorial Book Award for excellence in art publishing from the Art Libraries Society of North America earlier this year.

Librarians are facing challenges because the app is sold exclusively through the iTunes Store and is only compatible with the latest iPad. Currently, Columbia University Libraries does not provide access to public tablets and there is not an existing framework regarding how to “lend” these devices to users.

A second challenge involves preservation of the app. Because its value to the scholarly community is tied to its interactive learning environment, librarians searched for a means to purchase a preservation copy that guarantees long-term access to both content and functionality. They discovered that because the app can only be licensed through iTunes (currently, there is not a business model that supports a flat out purchase of the app), there is no way to capture and archive content. Even if it was, there are no benchmarks that estimate the costs of long-term app maintenance and storage.

During the Fall semester, several of us at Columbia Libraries had an opportunity to discuss access and preservation challenges with the New Business & Product Manager at Yale University Press. We learned that academic publishers are also grappling with these issues, and there are opportunities to work together and discuss strategies related to the preservation of enhanced e-books. One idea is to begin with the preservation of underlying images and text. While this excludes functionality, it provides a starting point that may better acquaint information professionals and publishers with preservation challenges at hand.

It will be interesting to see how preservation discussions develop in the coming months and years. For now, partnerships are developing with the collective goals of providing access and preserving scholarly content for generations of researchers and students to come.

E.2. E-book Preservation Special Report

Note: The following report was published in the 2015 *Library and Book Trade Almanac*.

Citation: Goertzen, M., Wolven, R., & Carroll, J. (2015). Preservation pending: The future of e-book access in the digital age. In D. Bogart (Ed.), *Library and Book Trade Almanac* (pp. 23-40). Medford: Information Today, Inc.

Preservation Pending: The Future of E-book Access in the Digital Age

Melissa Goertzen
Robert Wolven
Jeffrey D. Carroll³

Abstract

For a very long time human knowledge has been recorded on print media and passed down, or preserved, for future generations. Since the latter half of the Twentieth Century however this human knowledge has been increasingly recorded in digital media. Whether we will be as successful as our predecessors at preserving human knowledge will depend on the steps we take now to identify and address the risks and threats when knowledge is stored in digital form. This report will focus specifically on e-books as a subset of all digital media. In some respects we are lucky in that the conversation surrounding the risks and threats to journal literature as it passed from print to digital has been going on for more than two decades, and some promising models and ideas have emerged. On the other hand e-books present problems that differ from those on the journal side and which will need to be addressed from scratch, as it were, with no prior models on which to build. Nor will this report provide ready answers. Rather, by bringing the conversation here we are throwing down the gauntlet and challenging all members of the information ecosystem to think seriously about the issues and to take the steps necessary to work toward a solution that is of mutual benefit to us all.

Introduction

The term preservation in libraries refers to the overall system of efforts aimed at maximizing the longevity of our cultural heritage and of the outputs of scholarly inquiry. Research is a

³ Melissa Goertzen is the E-Book Program Development librarian at Columbia University Libraries. Her research focuses on the documentation of the e-book landscape and the development of strategic plans related to e-book collection development programs. Robert Wolven is associate university librarian for bibliographic services and collection development at Columbia University Libraries. He is a former co-chair of the American Library Association's Digital Content and Libraries Working Group (DCWG) and a current member of the DCWG subgroup on digital preservation. Jeffrey D. Carroll is director of Collection Development at Columbia University Libraries. His work most recently has focused on the opportunities and challenges of cooperative collection development and vendor negotiations across multiple research institutions.

cumulative activity; scholars of today build on the findings and hypotheses of scholars that have gone before. As scholarly interests and societal values change, works that may have seemed to have limited, ephemeral interest when first published may take on new importance. Libraries have been the primary means of ensuring the survival of the scholarly and cultural record. For printed materials, they have accomplished this through *conservation*—the treating of single items in an effort to mitigate the effects of deterioration— by storing materials in controlled environments, and by converting the content to new formats when necessary. Each of these actions is but one element of a preservation strategy. Preservation practices evolve in the delicate balance between art and science. With each new generation, volumes of intellectual and cultural output undergo rigorous evaluations to determine which artifacts merit preservation efforts. Conservation efforts are made to preserve both the item itself as well as its content. Then, efforts are made to place items into a historic context that forms the basis for a common cultural heritage. However, as society becomes increasingly reliant on digital technologies to produce and disseminate information, the development of strategies to preserve this information has lagged behind. This is due in large part to the rate of technological changes taking place in the sharing of information and the complexity involved in constructing an environment that accounts for the quick pace at which this rapidly evolving technology becomes obsolete. The preservation of digital content becomes a daunting task facing information professionals.

Today, much discussion surrounds the development of e-book preservation practices. In many cases, proposed strategies to tackle this complex challenge are still theoretical and abstract. However, last year librarians at Columbia University Libraries (CUL) received an acquisition request for a unique born digital item that rooted our preservation challenges in an entirely practical context.

In 1963 Josef Albers published the *Interaction of Color*, which is hailed as the canonical work regarding the study of and relationships between colors. Albers developed the text to serve as a hands-on kit that provides an interactive environment to conduct silkscreen color studies. However, there are limits to interactivity as the book was originally published in print format. Now, fifty years after its original publication the text has been released in a digital format that brings Albers' original vision to life. Yale University Press released the *Interaction of Color* app, which includes over 125 color plates, 60 interactive studies, and commentary from Albers. This is a landmark publication and won the George Wittenborn Memorial Book Award for excellence in art publishing from the Art Libraries Society of North America in the past year (YaleNews, 2014).

Librarians at CUL faced a number of central challenges. Because its value to the scholarly community is tied to its interactive learning environment, we searched for a means to purchase a preservation copy that guarantees long-term access to both content and functionality. We discovered that because the app can only be licensed through iTunes (currently, there is not a business model that supports a flat out purchase of the app), there is no way to capture and

archive content. Even if there were such a method, there are no benchmarks that estimate the costs of long-term app maintenance and storage.

While there is debate regarding whether this item should be considered an e-book, it still raises valid questions about e-book preservation practices and strategies. Until recently, e-versions of texts have been viewed as supplementary to print copies (Armstrong and Lonsdale, 2009). Since they were acquired to “back up” print copies, preservation of the e-version was not considered urgent or even necessary. Today, we stand at a tipping point and discover an increasing number of e-books that have no print equivalent. Like the *Interaction of Color* app, many of these items have significant cultural and intellectual value that may impact the way knowledge is disseminated within the research community and general public in the coming years.

In today’s information society, securing the funds and server space to preserve text-only files can be a challenge. However, this process seems straightforward compared to what e-books are becoming: interactive pieces of media that are hybrids of databases, websites, and games (Robertson, 2014). In order to develop effective preservation strategies and frameworks, the library community first must understand on an intellectual, technical, legal, and financial level what preservation means for items in the digital world (Robertson, 2014). For instance, does digital preservation apply to content, functionality, or both? Is it possible to secure legal rights to preserve content that is licensed? How do we guarantee perpetual access to content that is tied to hardware, software, and file formats that will become obsolete within the next decade?

While the answers to the above questions will shape future preservation policies, a larger challenge looms on the horizon – the sheer size of the digital universe. With an exponential growth in born digital content, there is a massive body of electronic information at our fingertips – portions of it merit preservation while others do not. In this environment, how do librarians decide what items should be captured, stored, and managed over time? The answer will impact how our intellectual and cultural heritage is shaped for future generations of scholars, students, and readers.

The Landscape: E-books vs. E-journals

By 2005, the information profession faced a significant digital preservation challenge with the growth of e-journal publication and many libraries beginning to collect some journals solely in digital form. At that time, the academic community recognized a pressing need to take action to preserve scholarly electronic journals to guarantee perpetual access to information. The need for a stable and reliable preservation strategy was outlined in a report entitled *Urgent Action Needed to Preserve Scholarly Electronic Journals*, which was authored by library leaders representing research institutions in the United States and received the support of the Association of Research Libraries (ARL) (Kirchhoff, 2011). The document defined the initial strategy as a “way of managing risk, first, against the permanent loss of electronic journals and second, against having journal access disrupted for a protracted period following a publisher failure” (ARL, 2005). In

the following years, a task force under the guidance of OCLC's Research Libraries Group (RLG) and the National Archives and Records Administration (NARA) developed standardized tools and metrics. The most notable was the Trusted Repository Audit Checklist (TRAC) based upon the Open Archival Information Systems (OAIS) Reference Model. Today, digital repositories that have passed TRAC certification have the highest level of protection and security of e-journal content. Examples of these repositories include CLOCKSS (Controlled Lots of Copies Keeps Stuff Safe), Portico, and HathiTrust (Center for Research Libraries, n.d.).

Given the fact that standardized metrics now exist for e-journal preservation (and to date, have proven successful), one may ask why TRAC has not been more widely applied to e-book preservation. After all, both formats share similarities in their general makeup – they are both electronic resources that consist of two parts: files and metadata. However, when the surface is scratched, e-book preservation presents unique challenges that have not been encountered in the past. These include, but are not limited to, the following six issues:

1. E-book versions – journal articles, once published, tend to be stable; by contrast e-books – particularly trade books – may continue to be revised and enhanced;
2. Digital rights management (DRM) – relates to the common practice of embedding technologies in e-book files or devices that restrict access to content based on licensing terms;
3. Metadata – relates to managing the hierarchy of metadata at many levels of the publication (e.g. chapters, volumes, and series);
4. Legal issues – refers to the rights to use, share, and preserve e-books;
5. Format issues – refers to the coupling of e-book content with hardware, software, or distribution platforms;
6. Business models – refers to the sustainability of e-book licensing models and a fractured business landscape driven by proprietary platforms.
(Kirchhoff & Morrissey, 2014), (Kelley, 2014), (Kirchhoff, 2011).

With this complex e-book landscape, libraries must develop preservation practices and strategies that promote perpetual accessibility, usability, authenticity, and discoverability of information. Together, these four qualities ensure “fitness of use,” a primary principle that guides and informs ongoing preservation decisions, actions, and financial investments (Smith, 2004).

Global Trends

This report primarily focuses on preservation trends and challenges faced by stakeholders in the United States. However, as markets become increasingly global in nature, it is important to understand e-book landscapes outside of North America as they will impact preservation strategies in the future – particularly in terms of legal rights and collection development priorities.

In the English-speaking world, the U.S. has the strongest e-book market – e-book sales account for approximately a third of book revenues, particularly in the segments of adult fiction and genre fiction (Rüdiger Wischenbart, 2014). Across Western Europe, the adoption of e-books by consumers varies. In Germany, which has one of the largest book markets in the world, e-book purchasing trends are beginning to follow English language markets, but are approximately three years behind (Rüdiger Wischenbart, 2014). Across France, Spain, and Italy, e-book “momentum seems to be fairly limited” (Rüdiger Wischenbart, 2014, p. 20). Most surprising are trends observed in Sweden, where a “domestic e-book market has hardly taken shape, due to a mix of high prices for e-books and the role taken up by libraries who are lending significantly more e-books to readers than retail has sold so far” (Rüdiger Wischenbart, 2014, p. 20).

In other parts of the globe, a number of large markets are emerging. Below are four markets that may impact the future of online educational resources.

1. China: The world’s most populous country is also the second largest e-book market in the world, with 52 percent growth in 2012 (Kelbanoff, 2013). It is also estimated that there are 500 million Internet users and 1.22 billion mobile phone users in the country (Kelbanoff, 2013). Based on the market potential, competition among e-book providers is growing. In 2013, Amazon launched a dedicated Chinese platform, Apple’s iPhones were cleared to run on China’s mobile networks, and the domestic online platform, Dangdang, offered its catalog for free (Rüdiger Wischenbart, 2014). Currently, the fastest developing market for English-language books is focused on libraries, schools, and other institutions (Kelbanoff, 2013).
2. Russia: During 2013, Russian e-book sales surpassed those in the United Kingdom and Brazil, making Russia the third largest e-book market in the world (Gerden, 2014). The country has a “thriving reading culture in which writers and intellectuals occupy a prominent role in the public sphere and in which books stand at the center of the country’s cultural ambitions” (Rüdiger Wischenbart, 2014, p. 58). Since 2011, the market increased by nearly 200% and is expected to continue (Gerden, 2014). Although the sale of dedicated e-readers has declined, sales from personal computers and laptops continue to drive the market forward. Since 2012, Apple opened an iTunes store in Russia, and Google opened a Play store, and Kobo announced plans to offer services (Rüdiger Wischenbart, 2014).
3. Brazil: The Brazilian book market has shown strong growth for years with no signs of slowing down. For instance, consumers in 2013 spent 8.8 billion Brazilian reais (over \$4 billion) in books, a growth of 7% from the previous year (Utsumi, 2014). Currently, e-books are not widely used, but companies like Amazon, Apple, Google, and Kobo (all of whom set up e-book stores in the country within the past two years) believe there is large

potential for development. In 2013, a number of Brazilian publishers attributed 4% of their market shares to e-books (Utsumi, 2014). Another factor that is expected to influence the market is the fact that customers do not pay value-added taxes or sales taxes on books in Brazil (Rüdiger Wischenbart, 2014). “The Brazilian senate is discussing a bill that would officially make both e-books and dedicated e-readers tax-free. If the bill passes, the prices of E-Ink Kindles, Kobo devices, and Nooks would have to come down steadily” (Rüdiger Wischenbart, 2014, p. 65).

4. India: Currently, close to 24% of texts published in India are English-language titles. To date “all the major publishers in India, such as Penguin Books India, Hachette India, and Westland, have digitized their English-language backlists” (Vyas, 2014, para. 2). There is also great progress taking place in terms of regional e-books. With 25 regional languages spoken across the country, there is an opportunity for writers “outside of the country to make further inroads into the Indian market through getting their works translated” (Vyas, 2014, para. 4). This is especially true for self-published titles, as the legal process of forming partnerships with translation companies is less complex. On the technological side, dedicated e-readers have not made a significant market impact due to consumers’ questions about their overall value, but companies like Flipkart have successfully introduced e-reading apps that can be used on tablets and phones, allowing consumers to “read e-books on devices they are most familiar with” (Vyas, 2014, para. 9).

These global trends are likely to prove of increasing importance to libraries in the U.S. as well. Many public libraries acquire books from other countries to serve foreign-speaking populations. Research libraries acquire a range of publications from all world regions to serve both current and future scholarship. Just as e-book markets are developing in different ways and at different rates, though, the legal regimes, licensing terms, and business models that affect the ability of libraries to preserve this content present wide differences, and thus complicate the challenge.

Current Preservation Trends

Library Journal's e-book survey discovered a 93 percent increase in e-book collections among academic libraries since 2012. The survey also found that libraries anticipate e-book spending to comprise 20 percent of their budgets within five years (Blummer & Kenton, 2012), a figure that could range anywhere from a half million dollars for a medium-sized academic library to over two million for a large one. Despite the increase in e-book acquisition, it is not clear if or how e-book content will remain accessible in years to come.

Historically, the first-sale doctrine has provided a crucial underpinning for preservation work: when libraries purchase a print book, they are free to rent it out, resell it, or keep it forever (Smith, 2004). In this environment, libraries own their copy of the physical book, own the storage unit (often a bookshelf), and maintain the book for future use. Digital objects have disrupted this model. The idea of purchasing a “copy” of an e-book or “lending” it to someone

else is no longer relevant. In the digital world, the first-sale doctrine is no longer applicable in a legal context. In 2010, an appeals court ruling in the case *Vernor v. Autodesk* established that individuals cannot “buy” a piece of software, only acquire a permanent and non-transferable license to it (Smith, 2004). For libraries, this means that the legal ownership of individual titles, the storage unit (often a piece of hardware or software), and the ability to maintain files for future use are tied to the content provider – often a publisher or software developer.

Libraries license e-books and usually these licenses do not include the legal right to capture or preserve content (Billington, 2013). In the past, the inability to preserve e-book content has not been a concern as long as a print version is available in the collection. However, as born digital content becomes increasingly common, the information profession faces a renewed urgency to preserve content that does not have a print equivalent. This is particularly problematic when DRM restrictions prevent libraries from downloading or printing copies of e-books for archival purposes. In the current market, publishers are often not prepared to sell digital e-book files to libraries due to fears of piracy, decreased revenue, or an absence of business models that support such transactions.

On the flip side, even if publishers were prepared to sell e-books files, the majority of libraries do not have adequate infrastructure to house them. Though there are some notable exceptions, at this time, most do not have an institutional repository capable of storing digital files for generations to come (Yale University Library, 2013). Also, many do not have the financial resources, server space, or ability to perform continuous file maintenance and migration activities that are components of digital preservation programs.

Given the current landscape, trade-offs between ease of access and long-term preservation are necessary. In many cases, data dependences on hardware, software, or publisher’s and distributor’s servers create barriers. Information in the digital world is immaterial and must be created each time it is used (Smith, 2004). Obstacles surrounding the ability to provide both ease of access and ensure long-term preservation are linked to data dependencies on hardware, software, or publisher’s (and distributor’s) servers (Smith, 2004, Kirchhoff, 2011). This means that even if a library owns an e-book, it is often impossible to migrate those bytes from one platform to another. In order to facilitate both information accessibility and preservation, “the intellectual content of the book must be unpacked from its reliance on a particular hardware and software and then that content must be securely stowed away and maintained by one or more preservation agencies” (Kirchhoff, 2011, p. 34). Due to the complex relationship between access and preservation, there is currently no e-book solution that “simultaneously meets both the ‘current use’ and ‘future use’ requirements” (Yale University Library, 2013, p. 7). In some cases, it may make economic sense for libraries to purchase an electronic format without thinking about long-term access. Technical manuals, for instance, are in strong demand when current but quickly become obsolete. . On the other hand, academic libraries will often purchase titles (e.g., new foreign literature) not because of current user demand but to preserve the content in anticipation of future interest. (Yale University Library, 2013).

In the research community, both ease of access and preservation are vital to the continued development of disciplines, as many “subjects of inquiry and methodologies rely heavily on retrospective [and] current resources” (Smith, 2004, para. 1). For the many stakeholders involved in preservation, access to information removes knowledge barriers and promotes the widespread sharing and vetting of ideas.

Preservation Stakeholders

Today, e-book collections reside in the cloud, meaning that content is stored on server networks that run applications and provide online access to digital resources (GCF Global, 2015). Access to content is gained by purchasing a license, such as a subscription. In the current e-book landscape, there is no guarantee that content will be available long-term – or even next week for that matter. Events like mergers between service providers, evolving licensing models, and the obsolescence of hardware and software can result in the discontinuation of access to countless e-book titles. The instability of e-book collections and packages is a significant barrier as preservation relies on perpetual access to content (and some suggest perpetual functionality as well).

The challenges related to e-book preservation involve a range of stakeholders with very diverse interests, needs, and limitations. They include libraries (both public and academic), publishers, distributors, third-party storage services, and authors. The development of stable and sustainable e-book preservation programs will involve a deep understanding of the perspectives, responsibilities, and roles that each stakeholder brings to the table.

In the library profession, preservation challenges are felt in different ways by public and academic libraries. On the public side, many librarians are tasked with developing collections that meet the present information needs of diverse user communities. Much of this work involves licensing fiction and non-fiction titles, often through services like OverDrive or 3M, released by mass-market publishers who have commercial interests. Traditionally, facilitating long-term preservation has not been a necessary role for public libraries. In many cases, they do not have the legal rights, funds, or data infrastructures required to preserve this content (Ciabattari, 2013) (Kelley, 2012). Some public libraries have begun to challenge this model by negotiating deals with small and non-traditional publishers that allow for e-book files to be mounted on local library servers. While their primary impetus is to gain more control over the content rather than preservation per se, in the current e-book landscape, these libraries may be in the best position to preserve e-books as well.

When discussing the relationship between libraries and for-profit publishers, it is important to keep in mind that “the time horizons of the preservation community and of the commercial sectors are radically different” (Smith, 2004). In the private sector, data retention is often limited to anywhere between five and ten years (Smith, 2004). The time horizon of the preservation community “must include many generations of inquiring humans, not just the next two

generations of hardware or software upgrades” (Smith, 2004, para. 35). Some have cited this fundamental difference as the reason why for-profit companies have not played a strong role in the development of preservation systems, even though they provide leadership in areas like information technologies and digital asset management (Smith, 2004). The fundamental difference in data retention schedules is an area where creativity and experimentation will be required to develop preservation programs that promote buy-in from all sides and provide value to all stakeholders involved.

In academia, the ability to guarantee perpetual access to scholarly content is a pressing issue. In the research community, both ease of access and preservation are vital to the continued development of disciplines because inquiry and discovery rely on the utilization of current and retrospective resources (Smith, 2004). In academia, one way that librarians promote preservation is by “brokering deals with publishers and third-party storage services like Portico,⁴ which hold scholarly literature in a kind of escrow” (Robertson, 2014, para. 6). Librarians provide funding to such ventures through access fees and are guaranteed access to content when “trigger” events occur – most notably when titles go out of circulation or publishers cease operations (Robertson, 2014).

Due to the complexities and costs of digital preservation repositories, it is unlikely that individual institutions will serve as preservation centers (Smith, 2004). In the current networked environment, “one does not need access to a physical object to have access to information, [and] the relationship between ownership (and physical custody) of information and access to it will be transformed” (Smith, 2004, para. 36). It is likely that a small number of actors, such as Portico, will serve the academic community to preserve the scholarly record (Portico, 2015). In the future, third-party storage services will be an essential part of the “public good information economy that research and teaching have traditionally relied upon for core services such as preservation and collection building” (Smith, 2004, para. 36).

While preservation discussions between libraries, publishers, and third-party storage services will continue for some time, there is a growing volume of titles that are springing up outside of traditional publishing frameworks – they fall into the category of self-published e-books. Last year alone, editors at The Washington Post Book World received 150 self-published books *each day* (Charles, 2014). Established authors like Stephen King, along with those embarking on a writing career, are beginning to turn to services offered through Amazon and Kickstarter to fund or promote their work. This is not a trend that will disappear in the near future. In some cases, corporate publishers are beginning to step into the self-publishing arena as well. In 2012, Simon & Schuster partnered with the self-publishing company Author Solutions, Inc. to form Archway Publishing, a service that assists authors to reach audiences, achieve publishing goals, and make the leap to traditional publishing (Author Solutions, 2012) (Ciabattari, 2013). The preservation challenge surrounding this type of content (besides determining which titles merit preservation

⁴ See <http://www.portico.org/digital-preservation/>.

efforts) lies in the fact that e-books are currently exempt from mandatory legal deposit. Further, self-published e-books do not require an International Standard Book Number (ISBN), and there is not a simple way to contact or negotiate with independent authors regarding preservation issues (Kelley, 2014). This is a particularly harrowing issue when self-published works have scholarly merit. Currently, there exists a large opportunity for preservationists and computer scientists to work with scholars to develop and document standards for resources that are self-published for research and teaching purposes (Smith, 2004).

Market Segments

If e-books are to be preserved for long-term access, the impetus for action is most likely to come from those institutions that purchase – or license – them. An understanding of market sectors, and in particular of the ways different categories of e-books are acquired by libraries, is thus crucial to understanding the issues, barriers, and potential solutions for e-book preservation.

Broadly speaking, the library market for e-books mirrors that for print, with two major sectors: trade books aimed at the general reading public are primarily marketed to public libraries; scholarly works are aimed at the college and university population and acquired by academic libraries. While there is some overlap, and occasional “crossover” titles such as Thomas Piketty’s *Capital in the Twenty-First Century*, most publishers tend to focus on one segment or the other. This division has, thus far, resulted in significant differences in business models and in the terms under which e-books are sold – differences that have profound implications for long-term preservation and access.

While the academic e-book market is still evolving, certain core characteristics seem likely to persist. E-books are sold both as individual titles and in large groups or packages, both directly by publishers and through intermediary distributors, or aggregators, who group titles from many publishers in a common platform. Libraries will license access to some collections on a subscription basis, paying for access year by year, and losing access if payment stops. Many titles, however, are licensed for “perpetual access” with a higher one-time payment. Perpetual access licenses very often include provisions for archiving, either by defining terms under which the library may acquire and use the digital files, or by depositing those files with an agency such as CLOCKSS (www.clockss.org), LOCKSS (www.lockss.org), or Portico. All three of these agencies were founded to preserve e-journals, but CLOCKSS and Portico have since expanded to embrace e-books, with CLOCKSS holding nearly 25,000 e-books (CLOCKSS, 2012) and Portico almost half a million (Portico 2015). These arrangements reflect research libraries’ common interest in long-term preservation of the scholarly record.

By contrast, most trade e-books have been available to public libraries for only the past few years, and the markets and business terms are still diverse and changing rapidly. Again, though, certain common features appear to be emerging. Most trade publishers do not sell e-books directly to libraries, but work through a small number of distributors such as OverDrive, 3M, and

Baker & Taylor. While some publishers make titles available for purchase (or, more accurately, perpetual access), it is far more common for e-books to be licensed for a limited term of one or two years, or for a limited number of circulations. Even for those titles available for “purchase,” long-term access depends on the distributors, with no provision for external archiving. Again, these models reflect the practical concerns of public libraries, which primarily collect to meet current reading interests, with a high turnover in collections.

In addition to these two broad patterns, there are of course many variations and experiments. Some public libraries, notably the Douglas County (Colo.) Public Library,⁵ working primarily with smaller and non-traditional publishers have developed models for purchasing and mounting the e-book files themselves, thus securing at least the potential to preserve these books. Some academic libraries have become publishers of scholarly e-books, either on their own or in partnership with university presses. These and similar experiments offer potential alternative models for preservation that may in time prove more broadly influential.

For a large and growing number of e-books, however, library markets are virtually non-existent. New types of publication and new methods of publishing are emerging that have not, as yet, found their way into the standard channels by which libraries acquire content. Four main categories can serve as illustrations:

1. Non-profit publishers: Academic libraries have an interest in many works that are not produced by commercial publishers but by non-governmental organizations, government agencies, and research institutes. Many of these works are now published online, and while freely available on the web, they are not sold to libraries. Although not often thought of as e-books, they have formed an important part of research library collections.
2. Open access publishing: Although the scholarly open access movement has thus far focused largely on journal articles, it is starting to become a force in e-book publishing as well. These works are, by design, freely available, but for that very reason don't fit existing library methods of acquisition and preservation.
3. Non-traditional publishers: The relative ease of distribution over the web and sale through online retailers has brought a proliferation of small publishers and digital offerings from non-publishers (such as advertising agencies, art galleries) offering a mere handful of titles, often available only for individual purchase and download to personal e-readers.
4. Self-publishing: The ease of self-publishing on the web has led to an explosion of hundreds of thousands of titles. Most of these works are again only available for individual purchase and download to e-readers, or as part of e-book subscription services.

For many of these publications the role libraries should play in providing access and the importance of long-term preservation remain uncertain. What does seem clear is that the need for action to secure future access will not wait for that clarity to develop.

⁵ douglascountylibraries.org

Issues and Barriers that Require Solutions

Lack of clarity is pervasive when it comes to e-book preservation. Who should bear the responsibility? How can the costs be supported? What are the technical requirements? What legal rights are involved? Those with the strongest interest in outcomes are not at present in a position to be able to take direct action. Yet, a clearer understanding of the issues and barriers may suggest potential solutions.

Responsibility and Cost

Primary responsibility for preserving print books has long rested with research libraries: university libraries, the largest public libraries such as New York Public Library, and (in the United States) the Library of Congress. Publishers are supportive of preservation in principle, but are perforce driven by financial considerations, and likely to preserve their output only so long as potential sales exceed the cost. Public libraries are primarily responsible for serving the current needs and interests of their communities, and are not funded to sustain access to works once demand for them has faded.

Given the opportunity, scholarly publishers and academic libraries have both invested in e-book preservation, primarily through participation in CLOCKSS and Portico. Costs are shared by contributing publishers and member libraries. Though the economics are not transparent, it seems probable that at least some of the publishers' costs are passed on to libraries and other e-book purchasers in the form of higher prices. Though the long-term costs and success of this model remain to be seen, it offers a reasonable prospect that most mainstream scholarly monographs will be preserved.

The situation with regard to trade books is far murkier. Many popular books of the past – best sellers, dime novels, picture books, mysteries, science fiction and other genre fiction – are now preserved in academic library collections. With few exceptions however, these libraries didn't acquire such works as they were published, but long after, through donations from private collectors and secondary markets such as used book dealers. This deferred acquisition – and subsequent preservation – is made possible by the first-sale doctrine, under which owners of print books have the right to dispose of them as they choose, lending, donating or selling their copies to others. Many more popular books – from cookbooks to carpentry, self-help to devotional literature – are acquired by the Library of Congress through legal deposit (requiring copies of each new print publication to be deposited with the Copyright Office.)

Currently, however, as noted above, the first-sale doctrine does not apply to e-books. Instead, the rights of owners, even those who have purchased “perpetual access” are limited by licenses, and as noted earlier, few licenses for trade books include any provisions for archiving. Unless this situation changes, research libraries will no longer be able to rely on deferred acquisition to collect and preserve these works. While copyright deposit is mandated for electronic works with no print equivalent, this requirement has thus far only been applied to e-journals, and the

potential of the program to provide preservation and long-term access is imperfectly understood. If trade e-books are to be preserved, a new understanding of roles, responsibilities and financing will need to develop, in all likelihood one that actively engages publishers and public libraries, as well as research libraries.

E-books are subject to both legal and technical constraints that limit what libraries can do to ensure their preservation.

Rights and Obligations

While e-books are subject to the provisions of copyright law, as has been noted most e-book purchases are also governed by licenses that specify what actions the purchaser may take. However, in many cases the license a library signs when purchasing an e-book is only part of the picture. Several parties are involved in the chain of creating, producing, and distributing an e-book, and each may retain certain rights. Here, as elsewhere, the situation is somewhat different for scholarly and trade publications.

Authors of scholarly monographs are generally interested in broad dissemination of their work. The contracts they sign with publishers – often university presses – may assign copyright to the publisher or leave it with the author, and may allow the author certain rights of limited distribution (such as deposit in an institutional repository) but rarely place limits on the publisher's right to distribute the work. Library purchases are often made directly from the publisher, providing a relatively straightforward process for negotiating some form of preservation rights. By contrast, both authors and publishers of trade books depend on sales for their livelihood and survival. They are thus deeply concerned about unauthorized and uncontrolled distribution. The markets, pricing, and royalty structures for e-books are volatile, and all parties involved are cautious about giving away too many rights and suspicious of potential threats.

In both spheres, scholarly and trade, rights issues are further complicated because many library e-book purchases are made not directly from the publisher, but through a distributor such as OverDrive, 3M, ebrary, and many more. The library's rights to the content are controlled by its license with the distributor, while the distributor's rights are in turn limited by its contracts with many publishers, whose own rights are determined by agreements with hundreds or thousands of authors. In such an environment it is no wonder that the final link, between library and distributor, is generally quite limiting and cautious.

Complicated as this may seem, e-book licenses for major publishers both trade and academic at least offer opportunities for negotiation with high-volume impact. For the output of small independent publishers, self-published authors, and works on the open web, the situation is much worse. Many small publishers' and authors' works are for sale only through major on-line retailers – Amazon, Apple, and Barnes & Noble. As we've seen, they are marketed to individual consumers, whose rights are governed by those retailers' standard contracts, leaving little-to-no

room for library negotiation of preservation, or indeed any other, terms. Any attempts to negotiate directly with authors and publishers will necessarily be scattered and diffuse.

We might then turn with a sigh of relief to a class of e-books that are freely available for download from the web. While many of these are not what we typically think of as e-books – research reports from scientific institutes, policy briefs from think tanks and other non-governmental organizations – they are of considerable interest to academic libraries. Even here, though, rights issues are not absent. Most of these works are subject to copyright and, unless they are made available under some form of Creative Commons license, the ability to download a copy for personal use does not automatically convey the right to make further copies for preservation and access.

Technical Requirements

Finally, securing the right to preserve an e-book does not necessarily include the ability to do so. Publishing formats for e-books are diverse, and attempts to standardize on a single format such as EPUB3 have thus far met with limited success. Many are produced in multiple formats for use on different e-readers; preserving a single format may or may not suffice for future use. Far more than with journal articles, e-books are aesthetic objects, where style and appearance are considered important aspects of the reading experience; preservation of the words and images is necessary, but not always sufficient. Moreover, the e-books that are most at risk – those without print equivalents – are often the most challenging to preserve, incorporating multiple media and experimental features. DRM software is often part of the standard distribution package, requiring a separate format for preservation purposes.

Finally, some e-books, such as the *Interaction of Color*, are beginning to be produced as apps, with the application software an integral part of the work. From a technical standpoint, these objects might better be viewed as a software preservation challenge, rather than as e-books per se. Conceptually and intellectually, however, they seem more related to books than other types of software.

A relatively small number of libraries are likely to have the wherewithal, in technical expertise, storage capacity and back-up storage to preserve and render back e-books from diverse sources on a large scale. Just as with e-journal preservation, concentration of effort and collaboration around agencies such as Portico and CLOCKSS are likely to be the norm.

Recommendations for Action

For years to come, attempts to preserve e-books are likely to prove messy, uncoordinated, inconclusive, and incomplete, yet that is still better than waiting for things to become simpler and clearer. Failed attempts and partial successes can help narrow the range of possible solutions. Small-scale experiments may address specific technical, legal, or financial challenges, and then build towards more systematic solutions. Engaging multiple stakeholders in different

ways can help foster a sense of collective ownership of the issues. The recommendations that follow are offered, not as resolutions, but as explorations.

Focus on Outcomes

Ultimately, preservation is about ensuring the survival of e-books so that they may continue to be available for future use. However, different types of content may suggest different types of use, by different groups, at different times. For any group of e-books, libraries will need to consider:

4. Who is likely to want access in future? For what purposes? When?
5. Is preservation of the content (text and images) sufficient? How important are other features?
6. Are there classes for which a semi-dark archive, with limited or embargoed access, is sufficient?

Focusing on one standard, ideal model of preservation for all types of e-books is likely to impede incremental progress that may ultimately get us closer to the ideal.

Build Consensus on Priorities

Numerically and financially, a high percentage of the e-books purchased by libraries are still also produced in print and preserved in that form. At the other extreme, many works produced only as e-books are not purchased by libraries at all. With limited resources to devote to e-book preservation, should libraries focus on value (protecting investment, broad interest) or risk (ephemeral interest, entrepreneurial publishing)? Should they seek quick results (low-hanging fruit), or concentrate on difficult questions with potential for broad impact? Different stakeholders may well answer these questions differently. Finding like-minded partners may help organize action on multiple fronts.

Clarify Roles

There is broad consensus within the library community, and at least in principle among publishers, that preservation is important, but little discussion of who can and should take responsibility for different aspects – negotiating rights; storing content; maintaining format integrity; migrating access; supporting costs. Here again, different stakeholders are likely to have different perspectives, and roles may vary for different types of e-books. Two types of action are needed:

3. Sectors within the library community organized by type (public, academic), size (large, small) and function (preservation, collection development) need to organize discussions and build consensus around willingness to act and effective locus of activity.

4. Libraries, individually and through organizational channels, need to propose a variety of potential models to publishers and providers of e-books to test their willingness and ability to support specific preservation actions.

Experiment

A major barrier to e-book preservation is that it has been difficult to see how anyone can take action without massive changes in the way e-books are licensed, distributed, and stored. While systemic change is still difficult, small scale experimentation is possible.

6. As noted earlier, CLOCKSS, LOCKSS, and Portico are already preserving tens of thousands of e-books. Individually and collectively, libraries can take a more active role in encouraging new publishers to participate in these programs.
7. New open-access ventures such as Knowledge Unlatched have secured agreement from HathiTrust, Portico, and others to preserve and provide access to their e-book publications, a model that might be extended to additional content (HathiTrust, 2014).
8. Some of the e-books most at risk are the output of innovative, start-up ventures (some of which have already suspended activity). These firms might be amenable to a “rescue” operation in which one or more libraries would secure the right to preserve (and possibly provide access to) their e-book content.
9. Research libraries might explore a model used for some print archives, in which a particular publisher’s e-book output would be deposited with the library for preservation, with access embargoed for an interval, and/or restricted to researchers on the archive’s premises.
10. Many self-published books may ultimately be of more interest as a corpus than for reading as individual titles. Libraries might seek agreement from providers of self-publishing platforms to secure and preserve the content for non-consumptive use.

These examples are cited merely to suggest a range of possibilities. Ultimately, none may prove viable or extensible, but as with e-books themselves, innovation will require a willingness to fail sometimes.

Reward Success

If innovation requires a willingness to fail, it is also true that limited successes, however imperfect, must be encouraged. Publishers and libraries who demonstrate a willingness to take risks in order to preserve e-books ought to be recognized, with encouragement to others to follow and build on their efforts. Rewards may take many forms – favorable notice in library forums and media; formal awards and citations; and, most tangibly, increased business for those most committed to preservation. This is not to suggest that all initiatives should be received uncritically, but that acknowledgment of good-faith effort should accompany critical analysis, and that the best form of criticism is to point the way to something better.

Conclusion

At a time when the library profession is in transition due to the proliferation of technology in society, one may ask why time and resources should be invested in a complex challenge that addresses the future state of library collections. Are we borrowing trouble when we should be focused on today? In reality, the policies that guide e-book acquisition decisions can be coupled with preservation strategies. The digital universe is slowly erasing the distinction between collecting content “just in case” (i.e. an item proves later to be valuable) and “just in time” (i.e. for use today). As technologies evolve, digital items need to be optimized for preservation at the time of their creation, the time of their deposit into repositories, and managed over time (Smith, 2004) to guarantee long-term access in an environment where hardware and software obsolescence is guaranteed.

The good news is that we are in the early days of e-book collection development. There is time to develop strategies and frameworks that support the preservation of e-book content and functionality. At the same time, the knowledge that digital collections are not permanent signal that we must take action and begin experimentation. In the future, there will be “no uniqueness, no scarcity, [and] no category of ‘rare’” (Smith, 2004, para. 45) when dealing with digital information. These resources are part of our national cultural heritage, and the responsibility for ensuring their survival must be shared. Publishers, foundations, and the federal government must all play a part. But, action will only begin if librarians take a more active role in the stewardship of collections and build effective relationships with external stakeholders to collectively evaluate information and determine what items best represent the cultural heritage of the early 21st century. The answers will not be simple, but as evidenced through advances in e-journal preservation, they are not impossible.

As librarians at Columbia University Libraries grappled with the preservation questions surrounding the *Interaction of Color* app, we had an opportunity to discuss preservation challenges with Yale University Press. We learned that publishers are also grappling with these issues, and there are opportunities to work together, experiment, and discuss frameworks related to the preservation of enhanced e-books. For instance, one strategy may be to begin with the preservation of underlying images and text. While this excludes functionality, it provides a starting point that may better acquaint information professionals and publishers with the challenges at hand. As we learn from each other, we move towards sustainable, long-term preservation solutions one byte at a time.

References

- ARL. (2005, October). *Urgent Action Needed to Preserve Scholarly Electronic Journals*. Retrieved from <http://www.arl.org/storage/documents/publications/ejournal-preservation-15oct05.pdf>

- Armstrong, C., & Lonsdale, R. (2009, November). *E-book collection management in UK university libraries: Focus groups report*. Information Automation Limited, Final Report. Retrieved from <http://observatory.jiscebooks.org/reports/e-book-collection-management-in-uk-university-libraries-focus-groups-report/>
- Author Solutions. (2012, November 27). Simon & Schuster and Author Solutions launch Archway Publishing. Retrieved from <http://www.authorsolutions.com/News-Events/News-Releases/2012/Simon-and-Schuster-and-Authors-Solutions-Launch-Archway-Publishing/>
- Billington, J. (2013). Federal agency and federal library reports. In D. Bogart (Ed.), *Library and book trade almanac* (pp. 55 - 77). Medford: Information Today, Inc.
- Blummer, B., & Kenton, J. (2012). Best practices for integrating e-books in academic libraries: A literature review from 2005 to present. *Collection Management*, 37(2), 65-97.
- Center for Research Libraries. TRAC and TDR Checklists. Retrieved from <http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying-0>
- Charles, R. (2014, October 1). No, I don't want to read your self-published book. *The Washington Post*. Retrieved from <http://www.washingtonpost.com/blogs/style-blog/wp/2014/10/01/no-i-dont-want-to-read-your-self-published-book/>
- Ciabattari, J. (2013, September 3). Now there are 5. *Library Journal*. Retrieved from http://lj.libraryjournal.com/2013/09/publishing/now-there-are-5/#_
- CLOCKSS. (2012). *Home*. Retrieved from <https://www.clockss.org/clockss/Home>
- GCF Global. (2015). *What is the cloud?*. Retrieved from <http://www.gcflearnfree.org/what-is-the-cloud>
- Gerden, E. (2014, January 9). Russian e-book market matures, shows more potential for growth. *Publishing Perspectives*. Retrieved from <http://publishingperspectives.com/2014/01/russia-is-now-the-worlds-third-largest-ebook-market/>
- HathiTrust. (2014, April). HathiTrust rights and access working group charge [Web log message]. Retrieved from http://www.hathitrust.org/rights_and_access_charge
- Kelbanoff, A. (2013, December 16). The tantalizing English language eBook opportunity in China. *Digital Book World*. Retrieved from <http://www.digitalbookworld.com/2013/the-tantalizing-english-language-ebook-opportunity-in-china/>

- Kelley, M. (2012, October 18). Random House says libraries own their ebooks. *Library Journal*. Retrieved from <http://lj.libraryjournal.com/2012/10/opinion/random-house-says-libraries-own-their-ebooks-lj-insider/>
- Kelley, Michael. (2014, October 3). How libraries preserve e-books. *Publishers Weekly*. Retrieved from <http://www.publishersweekly.com/pw/by-topic/industry-news/libraries/article/64271-check-it-out-with-michael-kelley-how-libraries-preserve-e-books.html>
- Kirchhoff, A. (2011) E-books: The preservation challenge. *Against the Grain*, 23(4). Retrieved from <http://docs.lib.purdue.edu/atg/vol23/iss4/10/>
- Kirchhoff, A. & Morrissey, S (2014, June 1). *Technology Watch Report 14-01: Preserving e-books*. Retrieved from <http://www.dpconline.org/advice/technology-watch-reports>
- Portico. (2015). *Our Organization*. Retrieved from <http://www.portico.org/digital-preservation/about-us/our-organization>
- Robertson, A. (2014, May 9). The fight to save endangered books. *The Verge*. Retrieved from <http://www.theverge.com/2014/5/9/5688146/the-fight-to-save-endangered-ebooks>
- Rüdiger Wischenbart Content & Consulting. (Spring 2014). Global eBook: A Report on Market Trends and Developments. Retrieved from www.wischenbart.com/upload/1234000000358_04042014_final.pdf
- Smith, A. (2004). Preservation. In S. Schreibman, R. Siemens & J. Unsworth (Eds.), *A Companion to Digital Humanities*. Oxford: Blackwell.
- Utsumi, I. (2014, March 10). *The Brazilian book market*. The Brazil Business. Retrieved from <http://thebrazilbusiness.com/article/the-brazilian-book-market>
- Vyas, H. (2014, September 11). India's ebook industry shows great potential. *Digital Book World*. Retrieved from <http://www.digitalbookworld.com/2014/indias-ebook-industry-shows-great-potential/>
- YaleNews. (2014, May 14). 'Interaction of Color' app wins George Wittenborn Memorial Book Award [Web log message]. Retrieved from <http://news.yale.edu/interaction-color-app-wins-george-wittenborn-memorial-book-award>
- Yale University Library. (2013, March 3). *The eBook Strategic Plan Task Force*. Retrieved from <http://www.library.yale.edu/departments/collection-development/Yale-ebook-task-force-rpt.pdf>