

Customer Focused Collection Services in the Age of Big Data

Ying Zhang

University of Central Florida Libraries, Orlando, USA

Email address:

ying.zhang@ucf.edu

To cite this article:

Ying Zhang. Customer Focused Collection Services in the Age of Big Data. *International Journal of Intelligent Information Systems*. Vol. 7, No. 1, 2018, pp. 5-8. doi: 10.11648/j.ijis.20180701.12

Received: September 19, 2017; **Accepted:** October 29, 2017; **Published:** April 4, 2018

Abstract: As part of library core functions, collection services had always focused on resources and processes in the print age. With the advent of big data with prevailing digital technologies in the recent decades, academic libraries in the U.S. have increasingly brought customer into the center of collection services. Big data empower these customer-focused services in various formats and scopes. What are some common practices? How effective are they in addressing the customer needs while fulfilling the conventional goals of collection services? This article starts with a historical overview on the evolutions of collection activities from the perspectives of academic libraries in the U.S. It then shares several key trends and common practices enabled by big data to build collection services centering on customers, including demand driven acquisitions models, digital collections development, collection access and discovery enhancements and systematic collection assessments. The article also discusses the multitudes of implications and impacts brought by these new customer-focused collection services on the library and information science (LIS) profession, in technologies, in philosophies, in personnel, in budgets and certainly in user experience.

Keywords: Collection Services, Big Data, Academic Libraries, Demand Driven Acquisitions, Assessment, Digital Collections, Discovery Services, Customer Focused, Library and Information Science

1. Introduction

Collections and customers make up the two cornerstones of modern libraries; and hence form the two core functions for libraries. Originally, functions for library collections consisted of selection, acquisitions, collection development and management, cataloging, and preservation. These functions were commonly categorized as “technical services” in the U.S., as opposed to “public services” that typically include reference, instruction, and circulation. Created in the age of print domination, the terminology “technical services” distinctively implies the focus on the technical aspects of processing resources, instead to on users and customers directly [3].

The exponential development in digital technologies in the new millennium stimulated the steadfast growth of content, and brought myriads of electronic books and documents to libraries and their customers. The age of big data has risen at the same time. “Big Data” is a term that describes the large volume of data, both structured and unstructured, that could

inundate an entity on a day-to-day basis. They bear the characteristics of 3 “v”s: volume, velocity and variety [12]. The precise origin of the term is difficult to trace, nevertheless, it clearly has gained momentum with the prevalence of social media and the digital technology in the 21st century.

In the age of big data, various tools enable libraries to collect both quantitative and qualitative data straight from the users upon usage. These data illustrate not only what users want, but also how, when and where the usage takes place. Through analyzing and mining, these data not only measure the collection performance in terms of usage, but also help librarians to understand the user pattern and preferences. Big data push excellent lagging indicators upstreaming to leading indicators [4]. As a result, libraries can make data-driven decision in collection services to meet customer’s needs more effectively and efficiently. The traditional terminologies for collection functions such as “collection development” and “collection management” morphed into “collection services”, a term favored by increasingly more academic libraries in the

U.S. for its updated and expanded focuses on customers.

2. Historical Overview

Historically library functions related to collections encompassed selecting, organizing and preserving materials for as long as the history of libraries [8]. “Collection development” came into wide use in the late 1960s replacing “selection” to reflect the thoughtful process of developing a library collection in response to institutional priorities and community or user needs and interests [6]. As libraries expanded in collections and locations, weeding/deselection, assessments and relocation became necessities, and hence the term “collection management” emerged in 1980s as an umbrella term over “collection development”.

For over a century, library collections in the world grew sparsely till the modern days when higher education became attainable. In the U.S., professional librarians started taking over the responsibilities of identifying and selecting materials in the late 19th century [6]. Building or managing a print collection was an art, relying primarily on the librarians’ understanding of the library’s mission, their knowledge of the publishing within the scope of the subjects and collection convention. Librarians were the “arbiter of quality” of the materials added into libraries [6]. When publishing cycles were much longer and cost more, libraries served as the few places for books and resources. Selecting quality materials for the perceived demand vs speculated value were contingent upon the individual librarian’s perceptions. Libraries in the 19th to 20th century emphasize far more on the quality of collections by hiring the right professionals, such as bibliographers, and on establishing policies that reflected their philosophies of just-in-case. It was difficult to validate their efforts or to change the course of collection foci.

As print became more affordable and scalable in the mid-20th century, libraries around the world began expanding their print collections steadily and gradually outpaced the library space [6]. As early as in the 1990s, due to budgetary challenges and the technology revolutions, a few pioneer librarians called for a change in collection development philosophy. For example, Tyckoson and Atkinson each predicted that collection development would change from just-in-case towards just-in-time in order to directly respond to customer’s requests [5].

Their predictions proved to be ahead of their time. They didn’t come true until more than two decades later and after the exponential development in electronic resources and availability of big data. As McKinsey Global Institute identified in its 2011 report, big data can add value to many businesses in five ways: creating transparency, enabling experimentation to discover needs, expose variability, and improve performance; segmenting populations to customize actions; replacing/supporting human decision making with automated algorithms; and innovating new business models, products and services [9].

Relying mainly on big data and the various analytics tools,

academic libraries in the U.S. in the past decade have been able to rapidly transform the practice of collection development into “collections services”, such as building collection directly from customer’s demand unmediated in new acquisitions models, developing digital collections to an expanded customer base, enhancing access and enrich data for better discovery, and conducting systematic assessments to provide more robust collections. Common practices of such customer-focused collections services are discussed below, using examples from the University of Central Florida Libraries in the U.S.

3. Discussion

The University of Central Florida (UCF) Libraries in Orlando, FL has proactively expanded the collection focus to its customers that represents the 2nd largest university populations in the U.S. In the past five years, UCF has implemented demand driven acquisitions, expanded its digital collections, launched a discovery tool and systematically conducted collection analyses and assessments. All these initiatives are made possible by big data; and represent the national trends libraries have taken toward customer-focused collection services.

3.1. Demand-Driven-Acquisitions

Coming to the spotlight about seven years ago, Patron-driven-acquisitions (PDA) fundamentally transforms the conventional library acquisitions and collection development. Commonly set up for electronic books or streaming videos nowadays, the concept of PDA had been around for a long time at a limited scale and mediated in the format of interlibrary loan or materials suggestions in the print age. In this new PDA model, a library will first identify a pool of titles or a profile of titles for access. The vendors or publishers will turn on the access. New titles are added continuously as they become available. The authenticated customers of that library wouldn’t know the difference of owned titles vs those access-only. Once the usage reaches the predetermined threshold, a purchase or a loan is triggered. The library would therefore only pay for what its customers have used. What sets it apart from its print predecessors was the removal of mediation. Once set up, the acquisitions of books go unmediated between the patrons and the library collections. The owned titles, therefore, mirrors more closely to the customer’s needs [11].

Evidence-based-acquisitions (EBA) emerged a few years after PDA, primarily based on the publisher’s needs to recover their investments more quickly and with higher predictability. In EBA, libraries pay a publisher or a vendor an upfront access fee to bring a large number of electronic books available to the customers. After an agreed period of usage, detailed usage data are provided to the libraries to determine which titles to purchase using the upfront fee paid. The rest of titles either are withdrawn from the user view or remain for additional access fee. While PDA is typically offered by an aggregator vendor or those who don’t directly

publish the content, EBA is often provided by a publisher on its native platform. The biggest advantage of EBA over PDA is the elimination of digital rights management (DRM), which may include simultaneous use, print/download options, number of uses within a time period, etc. The DRM-free model is far more favored by users.

Unlike the conventional bibliographer or selector models, demand-driven-acquisitions models like PDA and EBA would acquire books only upon usage by customers. By providing wider access to the users, they enable libraries build collections on known rather than speculative demands. All access activities are captured and stored as big data in very fine granularity, allowing the system to automate the process for the librarians to discover the usage pattern and drive decisions directly focusing on customer's needs. PDA and EBA each has its pros and cons for libraries. Financially PDA uses a pay-as-you-go model, while EBA is more parallel to a deposit model. EBA offers the libraries the advantage of predictable financial outlay and stable sets of titles. On the other hand, PDA stands out in terms of offering newer titles because of the updated profiles. It may take a little longer to set up a PDA, but the maintenances are less involved than EBA [13]. Each library needs to determine which one works better based on its local situation and consideration.

Demand driven acquisitions, like PDA and EBA, mark a total shift from the just-in-case to just-in-time. They embody all the five values to libraries identified by McKinsey: creating transparency, enabling experimentation to discover needs, expose variability, and improve performance; segmenting populations to customize actions; replacing/supporting human decision making with automated algorithms; and innovating new business models, products and services [9].

3.2. Digital Collections

In addition to demand driven acquisitions, many libraries like UCF Libraries are creating digital collections/repositories to supplement their collection services and reaching a greater community of customers. These digital collections may be born-digital such as new electronic theses and dissertations (ETD) where access remains restricted to the institutional customers, or locally digitized materials that are out of copyright or locally owned copyrights such as retrospective theses & dissertations (RTD), or linking digital repositories in open access where the worldwide customers can freely access. The priority of what forms these collections and the decision on how to present the collections are driven by the customer demands and usability. In return, these institutional digital collection services also contribute to the big data. Big data bestow the technological capability to these digital repositories for the huge amount of data storage and fast retrieval for the customers.

3.3. Discovery Services

As a type of big data, library resources are useful if they

can be easily discoverable and meaningful. In the age of big data, data discovery has been hailed as the next big trend [2]. Libraries increasingly are making efforts to enhance the discoverability and access to the growing resources including those acquired from vendors and publishers and their own home-grown contents. Various discovery tools have been developed and advanced to assist customers in navigating the myriads of print and electronic resources. Originally named "the next generation catalogs", discovery interfaces creating a layer to all the existing indexes, including online catalog, aim to facilitate the searching, retrieving, displaying and linking to the content of the resources [1]. Through its user interfaces, link-resolvers, APIs and interoperabilities with the traditional library catalogs, discovery interfaces offer customers a one-stop shop for library research information in a "google-like" experience for library research. UCF Libraries implemented a popular vendor discovery interface four years ago. Confirmed by the usage data, the UCF discovery interface has greatly increased customer's searches, result clicks and full-text retrievals. Even though not increasing content, discovery interfaces clearly add values in searchability, discoverability and simplified functions such as citation formatting and therefore enhance user experience for library collections.

3.4. Collection Assessments

The age of big data has brought the unprecedented amount of metrics for collections. The needs for continuously assessing the collection efficiency and effectiveness become critical in modern collection services. Effectiveness means collections are meeting the needs of the customers; and efficiency measures if these customer needs are met optimally in terms of access and quality or return on investments. As library budgets get compressed by inflationary costs, stagnate allocation and increased demands, assessments become paramount in collection services. UCF Libraries incorporates collection assessment into its routine collection services, and makes adjustment to collection strategies to align with the customer needs. Big data presented as statistics harvested from the content providers or integrated library systems (ILS) or web analytics tools from the Internet can depict extremely detailed usage. On the other hand, the sheer amount of these data are overwhelming. Without standardizations, they remain meaningless data points. In the U.S. National Information Standards Organization (NISO) spearheaded a Standardized Usage Statistics Harvesting Initiatives (SUSHI) to comply with the Counting Online User Networked Electronic Resources (COUNTER) reports [10]. These efforts establish guidelines and consistency in statistics creating and harvesting. Utilizing the big data analytics, usage data from various sources and platforms can be compared for meaningful assessments. Web analytics tools, such as Google Analytics, help identify what drives the web traffic to libraries' websites and how customers are reaching their collections online. All the assessments and analyses empowered by the various big data analytics, in return, help to build collection content and

provide access that are more versatile to customers.

4. Conclusion

Regardless the benefits, none of these collection services initiatives were implemented without debates or challenges at UCF Libraries. Each initiative requires funding and personnel resources. They may not fit all libraries. And sometimes the result may not be as immediate as expected. The EBA pilot implementation at UCF, for example, proved quite ambiguous in terms of the return on investment. Digital repositories are costly to implement, but also are still searching for an ideal system that works well for libraries. Like any big data discovery tools, although excel in each of use, the library discovery tools suffer from limited depth exploration [2]. Discovery tools add substantial expenditures to the library budget. Libraries often have to rely on external expertise in implementing the tools effectively. And they require periodical maintenance and updates. Collection assessments could be very time-consuming. Aligning data to ensure quality cross-sectional or time-series analyses are not always readily easy and quick. Two years ago the major revision on COUNTER reports threw a huge curveball to all libraries conducting collection analyses. Although big data provide the potential for tremendous value-adding concepts, libraries have to balance the budgetary, personnel and technological resources to implement these collection services. And finally, despite the unprecedented potential of predictive analytics big data can promise, libraries should be aware of the possible disadvantages when employing big data both as a leading and a lagging factor to avoid self-fulfilling prophecy. Nevertheless, these challenges cannot overshadow the benefits these customer-focused collection services bring. UCF's annual LibQual library survey has consistently indicated an upward trend in customer satisfaction, especially in recent years after these new collection services.

In the age of big data, an enormous amount of various data are being captured every second at an astounding speed. According to the latest IDC report, the amount of big data worldwide will easily mushroom to 180 zettabytes by 2025, from the 4.4 zettabytes in 2013 to give a perspective of growth [7]. The exponential increase in big data will continue bringing profound changes and challenges to the library and information science (LIS) profession. Big data enable LIS professionals to stay in tune with the customer needs and in return to serve them accordingly. Harnessing the ability to adapt and learn new technologies and maintaining the openness to changes in philosophies and proficiency in data analysis are skills and aptitudes extremely high in demand. Not only LIS professionals have to achieve high standards in customer services in the public services, but also we need to take advantage of the big data to build and improve customer-focused collection services in order to stay

competitive in the broad information services.

References

- [1] Breeding, M. (2015). The Future of Library Resource Discovery: A white paper commissioned by the NISO Discovery to Delivery (D2D) Topic Committee. Baltimore, MD: NISO.
- [2] Elliott, T. (2015, March 23). Big Data Discovery Is The Next Big Trend In Analytics. Retrieved from ZDnet: <http://www.zdnet.com/article/big-data-discovery-is-the-next-big-trend-in-analytics/>
- [3] Evans, G. E., Intner, S. S., & Weihs, J. (2002). Introduction to Technical Services. Greenwood Village, CO: Libraries Unlimited.
- [4] Goodwin, G. (2014, February 18). Leading Vs. Lagging Indicators and the Coming LoT-Big Data Explosion. Retrieved from LNS Research: <http://blog.lnsresearch.com/blog/bid/194257/Leading-Vs-Lagging-Indicators-and-the-Coming-LoT-Big-Data-Explosion>
- [5] Gorman, G. E., & Miller, R. H. (1997). Collection management for the 21st century: a handbook for librarians. Westwood, CT: Greenwood Press.
- [6] Johnson, P. (2009). Fundamentals of collection development and management. Chicago, IL: American Library Association.
- [7] Kanellos, M. (2016, March 3). 152,000 Smart Devices Every Minute In 2025: IDC Outlines The Future of Smart Things. Retrieved from Forbes: <http://www.forbes.com/sites/michaelkanellos/2016/03/03/152000-smart-devices-every-minute-in-2025-idc-outlines-the-future-of-smart-things/#3439e47669a7>
- [8] Magrill, R. M., & Corbin, J. (1984). Acquisitions management and collection development in libraries. Chicago, IL: American Library Association.
- [9] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition. Washington, DC: McKinsey Global Institute.
- [10] SUSHI FAQ: General Questions. (2010, July 8). Retrieved from NISO: <http://www.niso.org/workrooms/sushi/faq/general>
- [11] Swords, D. A. (2011). Patron-driven acquisitions: history and best practices. Boston: De Gruyter Saur.
- [12] What is Big Data? (2016). Retrieved from SAS: http://www.sas.com/en_us/insights/big-data/what-is-big-data.html
- [13] Zhang, Y. (2014, December 9). Evidence Based Acquisitions: does the evidence support this hybrid model? Retrieved from Information Today: <http://www.infotoday.eu/Articles/Editorial/Featured-Articles/Evidence-Based-Acquisitions-does-the-evidence-support-this-hybrid-model-101019.aspx>