Collections as a Service: A Research Library’s Perspective

Julie Linden, Sarah Tudesco, and Daniel Dollar

ACRL’s “2016 Top Trends in Academic Libraries” noted a “shift to the incorporation and integration of more continuous, ongoing, flexible, and sustainable review of collections,” and “an increasing need” to align collection development with “institutional research and curricular requirements.” This article describes these trends playing out at a top-ranked research library. Given the library’s long history of collecting print books, the article examines the monographic acquisitions program from multiple angles, including circulation, expenditures, approval plans, and e-book usage. There are also suggestions for additional metrics to evaluate collections as a service, as well as questions posed to frame an ongoing research and assessment agenda. The models described in this case study and the proposed research and assessment agenda can be adapted by other academic libraries, both to fit local institutional needs and as collection development and “collections as a service” are increasingly realized within interinstitutional collaborative networks.

Introduction

Yale University Library is a major academic research library with a sizable collection and a large annual collection development budget. The institution had an aggressive collection-building focus in the mid to late twentieth century, resulting in an extensive print collection that continues to serve a diverse community of scholars.

And yet, the institution’s significant resources do not insulate it from the environmental factors that shape and constrain all research libraries: information abundance, the scholarly communication marketplace, the impact of digital technology on every facet of higher education, and budget realities. Yale Library does not and cannot collect comprehensively; it cannot and does not serve as a “library of record” or as a safety net for other libraries that do not collect as much as they once did or might like to.

The “Collection Assessment” section of ACRL’s “2016 Top Trends in Academic Libraries” describes both a current state and future directions:

"There has been a remarkable shift to the incorporation and integration of more continuous, ongoing, flexible, and sustainable review of collections rather than ad-
hoc project-based models….There is an increasing need to establish more holistic
and agile approaches (both qualitative and quantitative) to manage budgetary
constraints while ensuring that collections are “responsive” and committed to
institutional research and curricular requirements and needs.2

Yale Library is developing a multifaceted assessment program to approach collection
development challenges. This case study focuses on assessment of monographic
acquisitions. Many research libraries have documented their experiences with acquiring
e-books,3 circulation and e-book usage analysis,4 collaborative collection development,5
and data-informed approaches to collection development.6 This paper builds on that
literature from the perspective of a large research library with a long history of collecting
print monographs, a relatively recent history of programmatic collection assessment,
and a collection development philosophy that posits collections as a service “in sup-
port of the teaching, research and public missions of the university”7—not collections
as ends unto themselves.

The article proceeds in four sections: a sketch of the library’s current collection develop-
ment landscape, a description of the foundational elements of the collection assessment
program, a discussion of how collection assessment informs collection development
decisions, and a research and assessment agenda for the research library community.

Collection Development at Yale University Library

Yale’s “About the Library” page describes the library circa 2016:

… immersion in the Library’s collections and engagement with our staff and
services is what transforms the intellectual and academic development of our
students….service to the Yale community is our top priority….

Our collection of nearly 15 million print and electronic volumes is housed in 15
libraries, including Sterling Memorial, Beinecke, and Bass libraries, the Center
for Science and Social Science Information (CSSSI), as well as many other school
and departmental libraries…. collections range from ancient papyri and early
printed books to digital collections and electronic databases; and they transcend
a wide array of formats including text, photographs, video, audio, data, maps,
and ephemera.8

When carefully parsed, these sentences reveal much about the current state of col-
lection development.

First, the statement includes a lot of large numbers: “nearly 15 million print and
electronic volumes….housed in 15 libraries.” Yale Library, like other academic librar-
ies, is adapting to changing demands on library facilities and services.9 The physical
on-campus footprint has been reduced by approximately 25 percent since 2009 through
the closing and consolidation of campus library locations.10 Now and in the near term,
physical libraries hold potential for more study spaces to accommodate increasing
enrollments and additional consultation and instruction spaces for new and expand-
ing library and campus programs.

Like many other research libraries with large legacy print collections, Yale has a
high-density shelving facility; and, while there is projected space to accommodate
transfers into the early 2030s, the facility cannot be further expanded.11 Between the
transformation of library spaces on campus and the pressures on library space off
campus, there is increasing need to be judicious about the ongoing acquisitions of
physical materials for the general collections.
The next part of the “About the Library” statement addresses the breadth of collections: “…from ancient papyri and early printed books to digital collections and electronic databases….”

In 2014, $38.7 million were directed to collections expenditures; that placed the institution at number 2 on the ARL rankings. However, the number is a combined expenditure total that includes the University Library, the Law Library, and the Beinecke Rare Book and Manuscript Library—three units that have separate budgets, strategies, and reporting obligations. The impact of Beinecke on the combined number is significant; its collections expenditures account for nearly 40 percent of the total 2014 collections spend. The analysis described in this case study focuses on the general research collections; Beinecke and Law are not included in the data presented.

Finally, the “About the Library” description declares that the collections are a service: “immersion in the Library’s collections and engagement with our staff and services is what transforms the intellectual and academic development of our students…. service to the Yale community is our top priority.” The library’s Collection Development Philosophy, adopted in 2013, reflects this service imperative. The philosophy provides the foundation for subject-specific statements that outline the scope of collecting in support of the university’s current departments and programs. The subject-level statements are reviewed annually as part of the collection development budget allocation process; as academic programs change and develop, the library adjusts its collecting priorities and allocations, as appropriate.

Aligning collecting priorities and budgets with the stated priorities of the parent institution may seem obvious, but major research libraries have long been saddled with expectations about what they will and should collect for worldwide scholarly communities and for posterity. Although a 2013 report on “The Global Dimensions of Scholarship and Research Libraries” declared that “older models of ‘comprehensive collecting’ are dead,” the perceived need to make this statement well into the twenty-first century illustrates how deeply ingrained in library culture is the aspiration toward comprehensiveness. Mark Sandler describes print collecting at large research libraries in the latter half of the twentieth century:

In the world of bigger is better, dozens of library bibliographers at a given institution, or in a given discipline nationally, had every incentive to try to outdo each other in bibliographic gluttony….Research libraries are now burdened with the obligation to manage an accumulated glut of legacy print collections that demand staff attention and occupy valuable space.

The way the library now frames its collecting philosophy is a deliberate move away from the competitive quest of acquiring for the sake of the collection itself. However, as Sandler notes, simply stating that the library collects in support of the university’s teaching, research, and public missions does not ensure success:

It does not do much good to say that the goal of collection building is to support the research and instructional needs of the campus, if (1) we do not know much about those needs; (2) we do not understand the relationship between library resources and the fulfillment of acknowledged needs; and (3) we have not considered whether there might be more efficient or effective ways of satisfying campus needs than by funding the development of a large and locally managed library collection.

A crucial element of the institution’s collection development landscape addresses Sandler’s third point above: participation in the “BorrowDirect” service, an expedited
resource-sharing network among the Ivy Plus libraries. BorrowDirect has proven to be an efficient and effective way of meeting campus needs for borrowing print monographs. The Ivy Plus libraries are expanding beyond resource sharing to encompass collaborative collection development. These collaborations have thus far been smaller-scale projects focused on distinctive materials in specific subject areas, although plans are being laid for more ambitious efforts. These efforts align with directions outlined by Dempsey, Malpas, and Lavoie that see libraries “becoming increasingly embedded in networks of collaboration, cooperation and consolidation.”

Collection Assessment

Yale Library has, over many years, taken several different approaches to collection assessment. Some activities are routinized and ongoing, such as the annual gathering and reporting of collection metrics to the Association of Research Libraries, using circulation statistics to select materials for the off-site Library Shelving Facility (LSF), and using COUNTER data to guide e-resource renewal and cancellation decisions. Other collection assessments have been more project-focused or location-specific. For example, the Cushing/Whitney Medical Library gathered and analyzed data to inform print cancellation decisions; the Social Science Library and Kline Science Library conducted surveys to help determine (among other things) the place of print collections in the new Center for Science and Social Science Information. Until recently, these collection assessments have been isolated projects.

With the appointment of Yale Library’s first Assessment Librarian in 2013, the library’s Collection Development Department seized the opportunity to develop an assessment program that could operate at a higher and broader level, beyond management of specific location-based print collections or decentralized e-resource renewal and cancellation decisions. Further, this enabled the organization to develop assessment models that could be regularly refreshed with current data as well as aggregate formerly disparate data sources.

The data illustrated here are fundamental elements of the collection assessment program. The long time series allow for analysis of trends over more than a decade.

![FIGURE 1](image_url)

**FIGURE 1**

Collection Expenditures for Top 5 Ivy Plus Libraries

Note: In mid-2014, the Beinecke Rare Book and Manuscript Library suspended its acquisitions for one year, in preparation for its building renovation. The steep decline in Yale’s expenditures from 2014 to 2015 illustrates that Beinecke’s collection expenditures represent a significant percentage of the overall total.
Expenditure Trends
To provide some context for Yale’s expenditure trends, figure 1 shows fourteen years of ARL collection expenditure data for Yale and four other Ivy Plus libraries. The steep decline in Yale’s collection spending from 2014 to 2015 illustrates the magnitude of the Beinecke Rare Book and Manuscript Library’s collections budget. When Beinecke suspended its acquisitions program for one year, to prepare for its building renovation,26 Yale’s total collection spend came closer to other Ivy Plus peers and much further away from top spender Harvard. The remaining charts in this article delve into the details of Yale’s general collections.

Figure 2 displays the three major collection expenditure types over the past fifteen years. The data show a compound annual growth rate (CAGR) of 16 percent for e-resources. The CAGR for monographs is flat (0%) and for print serials it is negative (down 11%).

The data demonstrate the marked shift from print to electronic formats; the transition from print to electronic journals is nearly complete. In contrast to the dominance of e-journals over print journals, as of fiscal year 2015, the monographic spend was 85 percent on print, 15 percent on e-books. The data in figure 3 also show that the electronic resources category has grown beyond e-journals; it includes e-books, numeric and geospatial datasets, digitized archives, and multimedia collections. In 2013, expenditure codes were revised to provide more granular distinctions among e-resources, thus revealing new trends in e-resource acquisitions.

The new, more descriptive expenditure codes also enable analysis of collections spend by content type (monograph, serial) rather than delivery container (print journal, e-book, and so on). On figure 2, the expenditure trends chart, the shift to digital formats overwhelms the display and suggests—misleadingly—decreasing financial support of monograph collections toward expenditures on serials. The chart tells a different story when the data are visualized by content type in figure 3. As a percentage, spending on the monograph and serial categories had been relatively steady in proportion to one another. The significant changes are the emergence of new types of digital resources, such as datasets, digitized archival resources, and streaming audio and video services.
Approval Plans
In fiscal year 2015, the library acquired more than 50 percent of its monographs through approval plans covering the United States, United Kingdom, Canada, and Western European countries. The assessment model for these acquisitions includes number of titles acquired, expenditures, and circulation rates.

![FIGURE 3](image)

**FIGURE 3**
Yale University Library: Collection Expenditures by Content Type

![FIGURE 4](image)

**FIGURE 4**
Yale University Library: Approval Plan Acquisitions (United States/United Kingdom Vendors)

Note: The chart displays the circulation frequency of approval materials added by year of acquisition. If the item circulated between the date it was added and Dec. 31, 2015, it is represented in the ‘Patron Circulation’ percentage. The dotted line represents the total number of approval items received during that acquisition year.
The analysis showed a significant difference in the circulation frequency of U.S./U.K. acquisitions and the Western European acquisitions. The U.S./U.K. acquisitions are largely English-language materials and have a good chance of circulating at least once. Of all the materials received on approval in 2003, 70 percent of them have since been checked out by a user.

The collections from Western European vendors, mainly materials published in German, French, and Italian, have never demonstrated a high circulation rate. As the library has faced budget pressures, these assessment models have informed a nuanced approach to print monograph collecting rather than across-the-board reductions.

**Circulation Trends**

Figure 6 shows that over the past ten years, circulation has declined for the major constituencies: since 2004, undergraduate circulation is down 46 percent, graduate student circulation is down 20 percent, and faculty circulation is down 31 percent. Only the resource sharing category (the collections loaned to other institutions through BorrowDirect and interlibrary loan) has risen since 2004, by 46 percent. The decline in overall circulation rates was already widely known, but observing the decreases among groups that have traditionally been the heaviest users of print—faculty and graduate students—was eye-opening.  

**Apples and Oranges**

While tracking expenditures on print and electronic collections is fairly straightforward, as shown in figures 2 and 3, assessing user engagement with different formats poses challenges. The graph in figure 7 uses standard measures: COUNTER statistics for electronic resources (e-journals and e-books) and circulation activity for print collections. The differences in scale are stark. In 2014, there were more than 8 million article downloads (tracked by the COUNTER JR1 reports), 2.75 million book chapter/section downloads (COUNTER BR2 reports), and just over 380,000 circulations for print collections.
Looking at circulation and e-resource usage statistics together may seem awkward. Circulation of a print monograph is not the same as downloading a research article. It also may not be worth trying to develop a formula that translates one metric to the other. However, looking at the data together provides a perspective on the differences in scale; the usage of electronic resources dwarfs that of traditional print collections. Visualizing these different metrics of use, for different types of materials, on a single chart provides...
a way of looking at what Lorcan Dempsey calls “the facilitated collection, a coordinated mix of local, external and collaborative services assembled around user needs.”

Discussion

Iterations of these and related data analyses and visualizations have been shared with internal stakeholders and external groups. Within the organization, clear connections are being made between data analysis and collection development decisions. Engagement with analyses has resulted in requests for more data and information that may be used to delve deeper. Access to reports and data dashboards is part of an initiative to encourage selectors and liaisons to explore the data with their own questions. Librarians with expertise in particular subjects, languages, and imprint areas bring important context to these developing data stories.

Declining circulation data has influenced a variety of decisions. Circulation limits have been loosened for some previously restricted collections, loan periods have been extended for particular patron groups, and more material types have become eligible for resource sharing (the only category that has seen a circulation increase in the past five years). New on-demand scanning services encourage engagement with print collections to an audience that enjoys the convenience of downloadable content. These initiatives are examples of how the concept of “collections as a service” can be realized.

As for collection development, the decrease in print circulation may serve as a cautionary indicator, as Mark Sandler bluntly puts it:

If students and faculty members are rejecting library offerings in the tens of thousands and millions, it might be time to reconsider.... Cramming more and more unwanted books into a finite space can hardly pass muster as a rational business strategy, or as an effective means for advancing scholarly communication.

In determining how to more appropriately focus library monograph expenditures to meet changing user behavior and expectations, the model being developed examines three important facets: language, subject, and publisher.

The analysis of Western European-language monographs has focused on print acquisitions, given the limited holdings of e-books in these languages. Titles acquired on approval from Western European countries circulate at much lower rates than titles acquired on U.S./U.K. approval plans. Over the past two years, the library reduced commitments to Western European approval plans by revising publisher lists and prioritizing subject areas.

The decrease in print circulation cuts across all disciplines, although the downward trend for arts and humanities subject classifications lags behind the decrease in STEM subjects. There was consideration of when arts and humanities books circulate—that is, while social science and STEM books may have an immediate audience, since currency is an important factor in the value of the monograph, arts and humanities titles may take some years to be discovered and used. Ladwig and Miller’s study of first-circulation patterns at Cornell and Notre Dame found that “despite the book’s subject, if it has not circulated at least once within a few years after acquisition, the book is unlikely to find its reader.”

The impossibilities of perfecting predictive collection development, and particularly of trying to anticipate future needs, should be acknowledged. Evaluating purchase request data is a part of the assessment program. Some of these requests would eventually be acquired through the standard approval process. However, there are also requests that would never have been considered if not suggested by a patron. The data strongly suggest pulling back on some approval plan acquisitions and directing resources to the improvement of other fulfillment channels.
Although, as noted above, it is important not to equate e-book usage statistics with circulation statistics, the patterns of print circulation and e-book usage are remarkably similar. For example, an analysis by LC class reveals that arts and humanities books are heavily used in both formats. Among the highest performing subjects in the ebrary Academic Complete e-book package are LC classes B (Philosophy, Psychology, and Religion), D, E, F (History), M (Music), and N (Fine Arts). Analysis by publisher showed remarkably consistent patterns for certain presses. For example, Oxford University Press (OUP) books enjoy strong usage in print, on the Oxford Scholarship Online platform, and within the ebrary Academic Complete package. In 2015, the library launched a pilot to acquire print and digital monographs from OUP. Rebecca Seger, Senior Director of Institutional Sales for OUP, noted both the current state of user behavior and the inevitable trend: “As research moves towards a digital future, this program allows Yale to sustainably manage that transition and satisfy user needs for the format of choice—print and/or digital.”

Of course, it is not possible to determine with numeric data what the qualitative experience of an e-book use is, any more than it can be presumed that a circulated book is read cover-to-cover or is returned with the spine uncracked. Rather than trying to equate these two metrics of “use,” it is preferable to shed some light on the different and separable functions of e-books, characterized by Carey C. Newman as delivery, discovery, and use. When a student requests a print version of the already licensed online book, does that mean that the online book served a discovery function, but that the print version is optimal for use? When a professor requests an online version of a print book held by the library, is that because the delivery of that title is more timely or device-convenient? These scenarios exemplify a dilemma for libraries: users want both print and online versions of the same books. Buying both formats inevitably means that libraries are paying more for the same content in different containers. The use-value may justify the expense, but then, of course, less money is available for other collection expenditures. Duplication of content across formats may prove a temporary measure during a transitory time.

The case of Oxford UP’s popularity at Yale illustrates the dominance of a few scholarly publishers (some university press, some trade) with strong branding and reputations for quality. The adjustment of the monograph purchasing plans to meet demonstrated and expressed user needs is just one choice that affects the scholarly communication marketplace. It is important that collection development practices reflect and respond to scholarly communication trends. A reduction in spending on print monographs does not automatically translate to the same level of expenditure on e-books. For example, the library supports the arts and humanities by purchasing digitized archival collections; in recent years, there is also investment in digitized audio, video, and image collections; and very recently the acquisition of “raw data” for digital humanities projects.

In the near term, the plan is to further develop the collection assessment evidence base in the following ways:

Purchase requests: Patron book purchase requests are received in various ways, typically via a central online form or via e-mail to any one of several dozen subject librarians. To better understand who (faculty, graduate students, undergraduates) requests what (subjects, languages, publishers, publication dates), acquisitions staff now tag resulting order records as patron requests. Analyzing purchase request data along with BorrowDirect and interlibrary loan request data may reveal patterns that lead to adjustments in approval plan profiles or other methods for improving predictive collecting models. Alternatively, a lack of patterns might suggest that predictive collecting is not only impossible to finesse, but also less important than robust, rapid-
response fulfillment channels, including resource-sharing networks and some form of demand-driven acquisitions.

Digitization services: Yale Library’s Scan and Deliver service, launched in 2013, now accounts for a segment of print monograph usage, and of course there are scanning stations located in public spaces throughout the library system. There is also an increased demand for digitized texts for use in digital humanities projects. The ease and use of such services are expected to increase, as a greater portion of collections material is scanned for researchers and students. Gathering and analyzing digitization data will help organizations better understand this type of user engagement with print collections.

Measures of print collection use: Because circulation data is not the sole measure of print usage, efforts are underway to examine other potentially helpful metrics. Browsing behavior, measured by reshelving statistics, is a classic metric of in-library print use. With half of the library’s tangible collections housed offsite, a significant portion of the collection is not available for traditional patron browsing. Is there a way to evaluate how many volumes are requested but never checked out?

When usage is uniformly recorded across the library system, such assessment techniques are relatively easy to scale. Understanding collection use in terms of scholarly outcomes—for example, citations within dissertations—is intriguing, but at present perhaps difficult to execute at a large scale.

Research and Assessment Agenda

Beyond what the ACRL “top trends” described as “the incorporation and integration of more continuous, ongoing, flexible, and sustainable review of collections,” that document’s call for “more holistic and agile approaches (both qualitative and quantitative)” to collection assessment prompts a range of questions that would benefit from multiple studies and broad-based discussions within the research library community:

1. How do libraries determine what scholarly monographs to collect in print, both at individual institutions and collaboratively?
2. When is it justified to duplicate content across formats at the expense of more varied and diverse acquisitions?
3. Are there efficient, scalable methods for coordinating monograph acquisitions across resource-sharing networks that minimize duplication while still reflecting local institutional needs?
4. How do institutions support arts and humanities as those disciplines transition at different rates from tangible to digital content and scholarly practices?
5. In light of low and declining circulation rates, how should institutions approach collecting Western European–language monographs? Are these books becoming a type of special collection?
6. Do institutions have a clear and thorough understanding of the various forces at play in the scholarly communication marketplace? How does—or should—such knowledge inform collection development decisions at individual libraries and in collaborative endeavors?
7. How do libraries respond in an era of information abundance and improving technology where the library has a lessening monopoly on access to information and where it has shifted a majority of its funding to removing publisher paywalls that can be circumvented by skilled and motivated readers?
8. The transition of the book from print to digital has both altered and unbundled the elements of delivery and use that were once all contained in a print book. How do libraries individually and collectively respond in this unbundled environment? Where should they be proactive, where should they accept market solutions, and how best do they advocate for the needs of their user communities?
Conclusion

It is not radical to believe that the print book will become a special collections format along with manuscripts and incunabula. The e-book user experience will continue to improve. Text and data-mining methodologies will evolve into common tools for information discovery. The scholarly monograph, which has been in crisis for the past twenty years, will eventually give way, perhaps within a transformed tenure system, perhaps to other types of credentialing systems for scholarly communication.

Fifteen years out, the collection assessment models that have been established at Yale Library should help track whether e-book expenditures have overtaken print monograph expenditures; whether the acquisitions of print scholarly monographs have decreased, but the titles that are acquired show higher circulation rates, because collecting is more optimally aligned to the university’s teaching and research needs; and whether e-book usage continues to climb or, like recent e-journal usage, levels out. The model enables the ability to drill down into details (perhaps the library continues to collect print in a particular subject in a particular language because of continued steady circulation) and to extend to analyze the increase in emerging formats and new types of use (perhaps the library accelerates the investment in raw data for digital humanities). The key is to embrace a robust assessment culture to guide thoughtful movement toward “the facilitated collection,” as described by Lorcan Dempsey. In this environment, collection development evolves beyond a “print logic” driven by information scarcity to “network logic,” in which the library facilitates access to data and information, regardless of source or format, to meet user needs. Collections are a service, and, with service to our respective institutions and the broader scholarly community as a guiding principle and a willingness to examine and assess the steps we take, there is much exciting work ahead.

The authors thank our readers for their feedback on drafts of this paper: Heather Gendron (Yale University Library), Allen Townsend (Yale University Library), and David Magier (Princeton University Library), as well as the anonymous reviewers of College & Research Libraries.

Notes


8. “About the Library,” Yale University Library, available online at http://guides.library.yale.edu/about [accessed 16 December 2016].


17. Ibid., 16.

18. As of this writing, the thirteen Ivy Plus institutions are: Cornell, Columbia, Harvard, Yale, University of Pennsylvania, Princeton, Brown, Dartmouth, Duke, Johns Hopkins, University of Chicago, MIT, and Stanford. For an overview and history of the BorrowDirect service, see www.borrowdirect.org.


20. These focused collaborative collections include contemporary composers’ scores and Brazilian monographs.


25. Kelly Barrick and Themba Flowers, “‘This Place Is Awesome’: Program, Service, and Staff Planning for the Center for Science and Social Science Information” (NERCOMP Annual Conference 2013, Providence, Rhode Island, March 12, 2013), available online at www.educause.edu/sites/default/files/library/presentations/NC13/SESS02/ThisPlaceisAwesome.pdf [accessed 26 November 2017].


27. Trends in declining print circulation at research libraries are summarized by Rick Anderson,

28. Eighteen of Yale’s licensed e-journal platforms are represented here; while the largest packages are included, a complete tally of all the JRI statistics would yield a total even higher than 8 million article downloads. The significant growth in the e-book downloads over the past six years corresponds with the growth in Yale Library’s e-book acquisitions. The e-book section downloads total presented here does not include popular platforms that do not provide COUNTER statistics, such as Books 2x7, Safari Books Online, and OverDrive.


33. Yale Library has canceled e-book collection subscriptions from continental Europe because of low use.

34. Ladwig and Miller, “Are First-Circulation Patterns for Monographs in the Humanities Different from the Sciences?” 77.


38. Of course, this method does not capture information about what patrons request that the library does not purchase, for whatever reasons.


46. Ibid.