

Against the Grain

Manuscript 8336

How to Use Data and Assessment to Automate Collections Workflows and Decisions

Daniel M. Dollar

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>



Part of the [Library and Information Science Commons](#)

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries.
Please contact epubs@purdue.edu for additional information.

How to Use Data and Assessment to Automate Collections Workflows and Decisions

by **Daniel M. Dollar** (Associate University Librarian for Collections, Preservation and Digital Scholarship, Yale University)
 <daniel.dollar@yale.edu>

When considering how to use data and assessment to automate collections workflows and decisions, a good starting point is to think about the typical lifecycle of collections. **Galadriel Chilton**, the **Ivy Plus Libraries Confederation's** Director of Collection Initiatives, has developed the following graphic showing various phases of this lifecycle within the context of library operational groups.

more immediate steps we can take using today's technologies, an example of the emerging potential for machine learning and predictive analytics to alter how libraries respond to user needs is discussed in an article by **Ryan Litsey** and **Weston Mauldin** in the January 2018 issue of *Journal of Academic Librarianship* (<https://doi.org/10.1016/j.aclib.2017.09.004>).

Assessment is essential to fully leverage the power of data. Libraries must develop a culture

invest more time in outreach and instruction. We tracked the progress of this effort, in part, by observing book acquisition trends by firm orders (i.e., individual title purchases) versus approvals. Download statistics, coupled with circulation data, informed our move toward increased eBook acquisitions through e-preferred approval plan profiles and publisher agreements. This effort included an arrangement with a major university press for print and online access to their books using a funding model that allows for a managed shift to e-only by subject area. Data analysis has also demonstrated the advantages of acquiring shelf-ready print books as a means of improving request-to-delivery workflows and getting books more quickly into the hands of users. These findings led us to redirect collection funding to cover shelf-ready costs for English language acquisitions, and thereby use the collection development budget to more fully fund the total cost of acquisition.

Data-informed assessment is only going to play a larger role as collection development and management decisions happen in a networked environment. The **Ivy Plus Libraries Confederation (IPLC)** is actively exploring how to develop collections at network scale. We have initiated several collaborative collection development arrangements focused on specific subject or regional areas. To fully assess these initiatives and engage in more ambitious projects requires the development of a collection assessment program by the partnership. As a first step, a dataset feasibility study is underway to pull together a defined five-year set of bibliographic and holdings records for single-part monographs for analysis. The study is attempting to answer a prescribed set of questions using data from a subset of the **IPLC** institutions, before scaling up to the full partnership to inform large-scale assessment efforts aimed at shaping prospective acquisitions and retrospective retention decisions. Potential outcomes include the development of shared approval plans and a shared print program.

It is important to note here that the collections of individual libraries were never comprehensive, and all libraries hold non-rare, distinctive published materials that are not widely held. To maintain a collection lifecycle focused on users, a robust assessment effort at the library, and increasingly the network level, is essential. It is better for users served by the **IPLC** institutions to have access to the partnership's collection of 90-plus million volumes versus the collections of any one individual library, and through resource-sharing networks, this benefit extends to the broader scholarly community. While shared print initiatives allow networks to maintain bibliographic diversity as individual libraries manage down

continued on page 17



Shared with permission of the author

While this collections lifecycle could apply to all types of libraries, this article focuses on academic libraries. The graphic helps capture the idea of collections as a service directed toward the user communities in the center. The graphic also notes the actions associated with each phase of the lifecycle — actions that are informed by quantitative and qualitative data. How do we marshal this data to guide and improve the service that libraries provide through their collections?

Libraries have been using data to evaluate services and operations for a long time. Numerous articles in the library science literature discuss how to assess workflows and procedures, and find ways to automate and/or introduce other efficiencies. Technological changes are expanding the amount of data available, as well as our ability to analyze it for evaluative and predictive purposes, and artificial intelligence is poised to revolutionize library operations in the coming years. While this article focuses on

of assessment where it is habitual to ask for the relevant data and consider data requirements for assessment before launching a new project or service. Such a culture requires professionals who can determine key research questions, understand the relevant data sources (and their limitations), obtain the data, import it into environments where analysis can occur, and present it to stakeholders and administrators. Often this is an iterative process, with each analysis leading to additional questions. “Analysis paralysis” — analyzing data without using it to inform decisions — is a concern, but can be mitigated by having clear goals and objectives to be informed by a given assessment effort.

At the **Yale University Library**, we have worked to build such a culture. For example, we used circulation data to shape approval plan profiles used to acquire monographs (i.e., books). One objective was to decrease the amount of time subject librarians spent doing title-by-title book selection, so they could

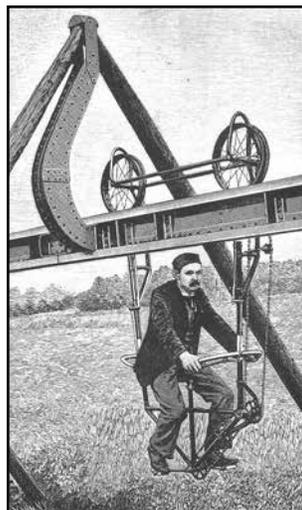
How to Use Data and Assessment ... from page 16

their print collections over time. These efforts will not be fully successful without robust assessment efforts informing and influencing collection development decisions.

Applications and tools to analyze and visualize data are key to successful assessments. These tools must scale to large datasets, be regularly refreshed with new and corrected data, have security and access controls (where necessary), employ transparent or understandable algorithms, and be queryable to address evolving and novel questions. For example, at **Yale** we migrated from static monthly collection fund reports to weekly refreshed reports viewable through Tableau. Subject librarians who manage allocated collection funds have praised the more intuitive interface and up-to-date financial data in helping them more effectively monitor their allocations and be timelier with acquisition decisions.

At the network level, the **IPLC** is exploring applications and tools needed for collection management and development. A working group is engaged in this research with the goal of developing a suite of collection lifecycle tools to inform collaborative collection efforts. A hoped-for outcome would be a vendor-neutral selection tool, coupled with robust assessment data, to facilitate separate, coordinated, or joint collection building.

Libraries must embrace a world where assessment and applied technologies will play an increasing role in shaping collection workflows and processes. Vendors have a role to play in providing tools and the necessary data to inform local and networked operations. Data privacy (institutional and personal) and algorithm transparency are critical issues that libraries need to address with the vendor community. There must also be an understanding that libraries will increasingly acquire and manage collection materials in a network, say more like the way you think of branch library systems today. Ideally, libraries and vendors can work together to create products and pricing models viable at network scale, and available open access where possible. Libraries can realize workflow and economic efficiencies in how information resources are acquired, described, discovered, and preserved, while also working with vendors in a healthy scholarly communications marketplace where innovation continues, and the issues of data privacy, intellectual property, and algorithm transparency are addressed.



We have moved from the labor-intensive analog days to a digital environment where information resources in all formats (print and digital) can be provided to users at point of need, as well as made available for computational analysis. Libraries will continue to evolve in how they manage collections, working in collaborative networks and in mutually beneficial arrangements with publishers and vendors. Libraries must embrace a culture of assessment, locally and in close partnerships, to guide a wide range of decisions affecting all aspects of the collections lifecycle. The ultimate goal is to maintain

and improve service for our user communities, including the global scholarly community. Libraries are robust, versatile organizations, and we will continue to be so into a future increasingly enabled by data and technology where the services provided through library collections are developed, described, managed, analyzed, preserved, and open. 🐼

*Author's Note: I want to thank **Galadriel Chilton** for generously sharing the collections lifecycle graphic for use in this article. — **DD***

Progress and Pitfalls in Consortial eBook Acquisitions: The Orbis Cascade Alliance Experience

by **Abigail Bibee** (Technical Services Librarian, University of Arizona) <bibee@email.arizona.edu>

and **Andrea Langhurst Eickholt** (Collection Management Librarian, Eastern Washington University) <aeickholt@ewu.edu>

and **Jesse Holden** (Program Manager, Shared Content & Technical Services, Orbis Cascade Alliance) <jholden@orbiscascade.org>

The **Orbis Cascade Alliance** (“the **Alliance**”) is a consortium of 38 academic libraries in the Pacific Northwest, comprised of a diversity of institutions¹ serving over 275,000 students. In 2011, the **Alliance** began a program to develop a shared eBook collection. Initially built around a demand-driven acquisitions (DDA) plan run through **YBP** and **EBL**,² the program has evolved in recent years to capitalize on new opportunities and respond to challenges that have emerged in managing such a large program.

The \$1M eBook program is centrally funded via consortial dues and is intended to benefit all members. Like funding, management of the program is centralized through the **Alliance** “team” structure. The Ebook Standing Group (ESG) operates under the auspices of the Shared Content & Technical Services (SCTS) Team. Drawing on varied expertise from across the consortium, the ESG

is comprised of around ten members from **Alliance** institutions and the SCTS Program Manager.

Access and selection of eBooks is managed through the shared Alma ILS, which includes a consortial “network zone” (NZ), and Primo discovery service. Bibliographic records for discovery are centrally loaded into the Alma NZ, and therefore facilitate discovery and access for all **Alliance** member institutions via Primo. For each title accessible through the eBook program, a public note displays in Primo, distinguishing between “discovery” (not yet purchased) and purchased titles.

Managing a consortial process of selection, acquisition, and access to a shared collection requires the integration of several moving parts. New complications for the ESG developed as the eBook program expanded and plans diversified, especially in the 2017-18 year. Along the way, the group learned many les-

sons about implementing and simultaneously maintaining various eBook plans at a consortial level, which we outline below.

Acquisitions Models

Several models of eBook acquisition are available to consortia, but not all models are suited to every situation. Examples include DDA, evidence-based acquisition (EBA), package subscription, front and back-list purchases, and collection purchases. Selecting a plan or approach to meet participant needs is essential. As a result, a periodic review to ensure plans continue to meet needs over time is an important aspect of developing and maintaining a consortial collection.

Pitfall: It can be easy to initially underestimate how complex a seemingly simple plan may become when implemented in a consortial space.

continued on page 18