Purdue University
Purdue e-Pubs

Charleston Library Conference

Distinctive Collections: The Space Between "General" and "Special" Collections and Implications for Collection Development

Daniel Dollar
Yale University, daniel.dollar@yale.edu

Gregory Eow
gregory.eow@yale.edu

Julie Linden
Yale University, julie.linden@yale.edu

Melissa Grafe
Yale University, melissa.grafe@yale.edu

Follow this and additional works at: http://docs.lib.purdue.edu/charleston

Part of the Library and Information Science Commons

An indexed, print copy of the Proceedings is also available for purchase at: http://www.thepress.purdue.edu/series/charleston.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences.

Distinctive Collections: The Space Between "General" and "Special" Collections and Implications for Collection Development

Daniel Dollar, Director of Collection Development, Yale University Library
Gregory Eow, Associate Director of Collection Development and Kaplanoff Librarian for American, British, and Commonwealth History, Yale University Library
Julie Linden, Head of Collection Management, Center for Science and Social Science Information, Yale University Library
Melissa Grafe, John R. Bumstead Librarian for Medical History, Cushing/Whitney Medical Library, Yale University

Abstract

Many libraries separate collection development activities into two broad categories, that of "general" collections versus "special" collections. Although this makes for a clean distinction between two areas of library activity (roughly the work of librarians as distinct and separate from that of archivists), in between these two poles lie "distinctive collections"—items that are neither especially rare nor unique (special), but are also not run-of-the-mill monographs or journals. Government documents, numeric datasets, ephemera, area collections, audiovisual media, born-digital materials—these are all recognized subsets of library collections with their own frameworks (more or less developed) for acquisition, cataloging/metadata, preservation, inter-institutional collaboration. Falling as they do somewhere between general and the special collections, these distinctive collections are often overlooked in traditional collection development and public service activities. This session presents an overview of how distinctive collections and their management fit into the overall collection profile of a library.

A Conceptual Framework for Building Collections

The Yale University Library is working to create a conceptual framework to guide collection building in a time of change. In the past, the Yale University Library operated as a loose confederation of over 20 largely autonomous libraries, many with their own policies and procedures for collection development and public services. Two exogenous developments, however, have forced the Yale University Library system to increasingly operate, in the words of University Librarian Susan Gibbons, as "one library."

First, there was (and continues to be) the challenges posed by the proliferation of digital resources. In 10 years, the percentage of the collection budget spent on digital resources at the Yale Library has gone from just under 20% in fiscal year (FY)03 to just under 60% in FY12. The lifecycle of e-resources lend themselves to centralized management, from negotiation, to licensing, to access through a discovery layer. These justifications for central management of e-resources, however, are largely moot, since the market for digital resources is already highly centralized in the hands of a relatively small number of players. In FY12, less than 5% of the vendors that contract with the Yale University Library accounted for over 80% of the library's spend on electronic resources. The second exogenous development was the 2008 financial crisis, which presented the Yale Library with double-digit budget cuts for the first time in recent memory.

The proliferation of digital resources and simultaneous cuts in materials budgets produce what one might label the problem of austerity in the digital age. This challenge led the Yale Library to look for efficiencies and productivity gains through the centralization and automation of collection development. That the management of e-resources could be handled through a more centralized workflow was widely understood. But

print collections can also be automated. For instance, item-by-item selection of print monographs is a resource-intensive procurement strategy that is hard to justify in an era when many publications are available on demand and effectively never go out of print. Affecting as it does both electronic and print resources, it was clear that the problem of austerity in the digital age necessitated a top-to-bottom rethinking of the library's previously decentralized collection development structure.

From one vantage point, it looked as if there would be two collection development philosophies. On the one hand, there would be the collection development philosophy of "general" collections, which would rely on centrally negotiated, big packages of electronic resources and a mix of approval plans and patron requests to build print collections. On the other hand, there would be the collection development activities of the special collections units, where collection development archivists and curators painstakingly build heavily curated collections through the cultivation of donations as well as the purchase of material on the rare book and manuscript market (including the use of auctions as a procurement tool). However, we find that this

level of abstraction—with general collections on one end of the collection development spectrum and special collections on the other—is not capacious enough to account for the collection development activity of a major research institution. A vast amount of material that is not special in the sense of consisting of manuscript or archival material cannot be accounted for in an automated, streamlined, and centralized collection development apparatus. For instance, area studies materials, art books, datasets, and much of the material falling somewhere between the general and special ends of the spectrum cannot be accounted for in the automated collection development vision.

We propose a third category, that of "distinctive" collections, to account for the materials that sit uncomfortably in the middle ground between general and special collections. Using the means or tool of procurement as the identifying marker for whether material is classified as either general or distinctive, general collections are those materials that are acquired through highly automatic and centralized means, as outlined in Figure 1.

Figure 1. "General" Collections

"Distinctive" Collections

- Resources where the management workflows at present are not efficiently consolidated:
  - Subject expertise is required for individual item selection decisions.
  - Limited print runs, geographic, copyright or other restrains requiring continued "just-in-case" purchasing.
  - Restrictive funding sources requiring close compliance monitoring.
  - Evidence still needed, in the form of qualitative and quantitative data, to support collection activities.

Yale UNIVERSITY LIBRARY

Figure 2. "Distinctive" Collections

Distinctive collections, on the other hand, are those materials that for one reason or another need more attention in the selection or procurement process, as outlined in Figure 2.

Case Study: Numeric Data and the Center for Science and Social Science Information (CSSSI)

Some library collections, such as area studies, much foreign language material, and limited print run arts books, are readily identifiable as falling in the distinctive collections category. However, the collection development activities of Yale University Library's Center for Science and Social Science Information (CSSSI) reveal that the distinctive collection category has use even in an area where one might expect collection building to be far along the process of being automated—in the natural, physical, and social sciences.

Core collections are CSSSI's bread and butter—these are online journals, databases, and monographs from major academic and some trade publishers. By acquiring much of this content via approval plans and packages, CSSSI librarians do not lavish time on title-by-title selection. Just as important to CSSSI is a type of

distinctive material that can be time consuming to acquire and manage: numeric data, which are machine-readable data files intended for analysis in statistical software packages, such as Excel, SPSS, and R.

Data are "having a moment," and rightly so. Libraries, including Yale's, are increasingly involved in supporting scientists and social scientists to manage the data they collect, analyze, store, describe, and disseminate—and here we bring our expertise in metadata and digital infrastructure to bear. It is an exciting time to be a data librarian and to work closely with researchers on data collections that are, on a continuum of general to special collections, truly special—gathered and compiled by researchers at this institution (maybe in collaboration with others)—unique, and mattering right now to whole teams of researchers and their funders.

Libraries also deal with numeric datasets that are not institutionally created, datasets that we acquire and, typically, license—from governments, intergovernmental organizations, nongovernmental organizations, commercial outfits large and small—sometimes one-person operations. On the general end of the continuum

are data acquisitions that are well-established and held widely in research libraries; they often work on a subscription model, much like a full-text database, or a "membership" model. Institutional membership in ICPSR, the Inter-university Consortium for Political and Social Research, provides access to ICPSR's vast archive of numeric datasets. A senior thesis writer wanting to analyze public opinion in the European Union can get the raw data from the Eurobarometer surveys from ICPSR. A graduate student modeling the biogeochemical effects of cropping in the Great Plains will find the perfect dataset in ICPSR. CSSSI librarians consider ICPSR to be "core;" we renew our membership every year; we load records for its datasets into our catalog; we promote it; we understand how to manage it as a continuing resource.

We also acquire and manage data in ways or from sources that are not so run-of-the-mill, and these are the distinctive collections. We collect such data in response to requests from our faculty and students. A few examples of data we've collected over the past year will illustrate the range of subjects and providers we deal with:

- Datasets on combat air activities in the Vietnam War, acquired from the U.S. National Archives and Records Administration;

- Italian national legislative election results from the early 1920s, acquired from an Italian research institute;

- Census of Population and Housing data from the Philippines National Statistics Office.

Acquisitions like these are difficult to routinize or streamline. Because we generally acquire these datasets upon request, it is hard to budget for them. In the social sciences collections budget we do set some funds aside for data requests, but we may exhaust those funds early in the fiscal year, or we may find ourselves toward the end of the fiscal year with money to spend. Therefore, based on known, ongoing demand, we do target particular areas for proactive collection development.

An example of the proactive collection development that we can do involves another distinctive collection at Yale: our South Asia collection. Yale's India Initiative (www.yale.edu/macmillan/southasia/indiainitiative.htm) is playing out in faculty hires, post-doctoral appointments, curriculum, and programs. The library has responded by creating a new position and hiring a full-time, permanent librarian to develop and promote the South Asia Collection. The social science librarians collaborate with the South Asia librarian on a range of collections issues. For example, on the general end of the collections continuum, we reviewed and adjusted our US/UK book approval plan to make sure our coverage met the needs of the India Initiative. We also collaborate on distinctive collection development, focused on data, both in response to user requests and in anticipation of usage. For example, upon request, we've purchased state and district-level India census data, with GIS boundaries. We know that when we buy sub-district boundaries, or major cities, those datasets will get used; they'll be a good investment of our collections funds.

Distinctive collections such as datasets can pose licensing challenges, especially when the data vendor is not accustomed to dealing with academic libraries. The vendor may be used to dealing with an individual or a research team, and it may take multiple conversations to help the vendor understand that the library pays for campus-wide access, which we may provide by circulating physical media (a dataset on a DVD) or on the campus network (available to current Yale affiliates by login). It's important for the librarians who review and negotiate license agreements to understand these aspects—so if the data librarian isn't the one handling the license negotiation, they stay in the loop during the license review process and "translate" between the library and the data vendor if needed.

Other aspects of data collection development and management may fall to either end of the general-to-special collections continuum. For example, datasets can be cataloged in MARC

and/or can be described with domain-specific metadata, given the infrastructure and justification for doing so. I would encourage you, in your libraries, to streamline and routinize what you can. The data librarian should not have to be more involved in license negotiation, cataloging, or access troubleshooting than, say, the business librarian would with business databases. Let the e-resources, cataloging, and acquisitions staff handle as much as they can. The data librarian should talk with the digital preservation specialists, so that datasets as digital assets worth preserving are not overlooked, and so that the particularities of datasets—such as file formats and accompanying technical documentation—are understood. But there's no reason that datasets should be excluded from the library's overall digital preservation strategy, or that the preservation work should fall completely to the data librarian. The data librarian may well have her hands full with traditional subject specialist responsibilities—communicating with faculty and students, instruction, research consultations, promoting resources, and so on—as well as grappling with emerging data management needs on your campus. And she'll need time to evaluate, often in collaboration with colleagues, the unusual, esoteric, necessary data acquisitions that will make your collections distinctive to your faculty and students.

Case Study: The Medical Historical Library, Harvey Cushing/John Hay Whitney Medical Library

The Medical Historical Library, Harvey Cushing/John Hay Whitney Medical Library at Yale, presents a case study of the three aspects of collecting: general, special, and distinctive. The library contains a large diversity of material, with approximately 140,000 volumes, circulating and non-circulating; 7,000 prints, posters, and drawings; 600 medical/scientific instruments; photographs; archives; manuscripts; paintings; pamphlets; video; and film. The Historical Library is digitizing many of its collections, both on its own and as part of a consortium of medical libraries that contributes to the Medical Heritage Library online. The Historical Library is also trying to wrestle with "born-digital" materials in the

course of collecting papers. In short, as a manager for this collection and a subject specialist, I deal with all kinds of materials that come my way that can fall into any of the three categories, with a lot of overlap.

For general collections, I purchase 100–150 new book titles annually in support of the History of Medicine department, which shares space in the Historical Library. Subscriptions to journals are folded into package deals with the main part of the Yale Library system. Beyond selecting, with help of our approval plan vendor, reviews, and input from faculty, I do not have to do any other work with the general collections. As emphasized earlier in this presentation, the general collections are managed with a focus on efficiency to minimize time and talent investment.

The Historical Library mainly works with special collections. Special collections are considered rare, unique materials, some endangered, in a variety of languages, and increasingly digital. The workflow for collection development and management is similar to most special collections. As librarian and manager for the collection, I negotiate with dealers, review catalogs and auctions, and receive gifts that require processing and paperwork. The Historical Library uses endowments to acquire these materials, many with specific designations. As part of a larger effort to make collections less hidden, the Historical Library is increasingly promoting the materials online and through digitization.

In between the general and special is the distinctive collection, which is not unique, nor widely held, and may have an institutional context. For Yale Library, this might include locally-grown knowledge that is made digital, such as the "Historian's Eye"(historianseye.commons. yale.edu/about/).

As digitization increases, there will be more remixing of digital materials—a move from digitization to new knowledge, new forms, and born-digital. Scholars and students are increasingly using our collections in complex ways that we want to capture. These may include datasets, websites, blog posts, and other scholarly

output. Questions that challenge the library's collection policy include:

- How do we capture these new forms in collections? Who does this? How do we integrate in normal workflows (ILL, etc.)?

- How do we designate money for this? Where can time be found to capture these materials?

- Are distinctive collections defined by what people find as gems? Who defines distinctive, and where does it lie? How should we go about building collections going forward?

Distinctive materials come from a variety of sources, and may be incorporated into special or general collections, or off-site storage. One impulse is to lump materials that do not fit easily into the general schema as "rare" or "semi-rareish," but that may not work because the distinctive collections may fall outside normal special collection streams. The point of this panel is to think more broadly about general, special, and distinctive collections, and realize that categories are slippery and may change from one to another. A beginning point may be to start surveying what materials lay outside of general or special collections development planning and find the distinctive.