Against the Grain

Volume 24 | Issue 4 Article 9

September 2012

Data-Informed Collection Management at the NCSU Libraries

Hilary Davis
NCSU Libraries, hilary_davis@ncsu.edu

Annette Day
NCSU Libraries, annette_day@ncsu.edu

Recommended Citation
Davis, Hilary and Day, Annette (2012) "Data-Informed Collection Management at the NCSU Libraries," *Against the Grain*: Vol. 24: Iss. 4, Article 9.
DOI: https://doi.org/10.7771/2380-176X.6177

Data-informed Collection Management at the NCSU Libraries

by **Hilary Davis** (Associate Head, Collection Management, NCSU Libraries) <hilary_davis@ncsu.edu>

and **Annette Day** (Head, Collection Management, NCSU Libraries) <annette_day@ncsu.edu>

Motivating Factors

North Carolina State University (NCSU) is a public, land-grant university with a focus on science and technology. The **NCSU** Libraries' collection consists of 4.4 million volumes, with an annual collection budget that ranges from $8.5 million to $10.5 million depending on the fiscal climate. This collection is essential to the research and teaching of our institution, but declining real budgets, growing demand, and shifting campus research emphases has made it increasingly important for us to understand and articulate our collections use, relevance, and value to campus. Making the best collections decisions with the resources available to us is at the center of what we do, and data is integral to those decisions. This article will describe our data-informed approach to collections at **NCSU**, highlighting the types of data we gather, the tools we use, and outcomes we've achieved with this approach.

What We are Measuring: The Why and How

Establishing a program of data gathering and analysis is a continuing cross-departmental effort at the **NCSU** Libraries. Crafting best practices for harvesting and assessing library collections data, creating and maintaining solid documentation around our best practices, and identifying who will assume responsibility for specific data elements going forward (e.g., a Collection Metrics Working Group, an ERM Committee, many subgroups of our ILS Management Committee) is an ongoing challenge.

One outcome of these collaborations is the identification of a core set of metrics that we think are valuable enough to warrant the investment to collect, analyze, and archive from year-to-year. They cover the spectrum, from full-text downloads and turn-away data, to grant dollars awarded to researchers at **NCSU**. It was difficult to restrict ourselves to these core set of metrics. Our initial inclination was to gather as much data as possible to see what it told us. But our goal is to build a manageable and sustainable data program, and focusing on these metrics is taking us towards meeting that goal.

Below is a list of the most common data elements that we collect and assess from our library and vendor sources and the list of campus data that we collect. We generally harvest all of this data on an annual basis, with the exception of expenditure data, which is managed throughout the year.

Library & Vendor Data

Metric	Details	Source(s)
Collection expenditures	Paid costs for books, e-books, journals, databases	Sirsi (ILS), E-Matrix (ERM)
Collection formats	Material type including print, electronic, microform, etc.	Sirsi (ILS), E-Matrix (ERM)
Usage statistics - Journals	COUNTER and Non-COUNTER, full-text downloads (HTML, PDF)	Vendor-provided usage reports
Usage statistics - Databases	COUNTER and Non-COUNTER, searches, sessions, turnaways	Vendor-provided usage reports
Usage statistics - E-books	Section requests, title requests, chapter requests, searches, and sessions, accesses	Vendor-provided usage reports
Circulations and Renewals	Total circulations and renewals (including Reserves and device lending)	Sirsi (ILS)
Multiple Holds, Missing, Lost items	Monthly reports on items that have more than two holds in a month and items that are reported as Missing or Lost	Sirsi (ILS)
ILL transactions	Books and journals borrowed by patron type (e.g., faculty, undergraduate) and college affiliation	Illiad
Publications and Citations	NCSU-centric view of publishing activity and citing activity (citing behavior of our researchers as well as how often our researchers' papers are cited); 1981 to present	Local Journal Utilization Reports (LJUR), a fee-based dataset from Thomson Reuters
Impact Factor	Journal impact factor is the number of cites in a particular year (e.g., 2007) to articles published in the two preceding years (e.g., 2006 and 2005) divided by the number of published articles in that same time period (2006 and 2005)	Journal Citation Reports (Thomson Reuters), subscription-based data

Campus Data

Metric	Details	Source(s)
Grant dollars awarded	Grants dollars awarded, per department or unit	RADAR grants database (NCSU database)
# PhD students enrolled and degrees conferred	Counts of doctoral students enrolled and degrees conferred per department	University Planning & Analysis reports
# Masters students enrolled and degrees conferred	Counts of masters students enrolled and degrees conferred per department	University Planning & Analysis reports
# Undergrad students enrolled and degrees conferred	Counts of undergraduate students enrolled and degrees conferred per department	University Planning & Analysis reports
# Faculty/Post-docs	Counts of tenured, non-tenured, and other (e.g., librarians) faculty per department/unit	University Planning & Analysis reports

Tools to Blend Data

Processing and analyzing the data so that it can answer our questions requires a variety of tools. At the **NCSU** Libraries, we use a mixture of home-grown tools and off-the-shelf products. A sample of those tools and how we use them are described below.

E-Matrix (ERM)

E-Matrix is our homegrown ERM system for the entire portfolio of our serials, regardless of format (Figure 1). We use E-Matrix to analyze composition (print or electronic, package or single title subscription) cost, use of the serials collection, and use of the licenses for our resources. We can add evaluative data for each title, including information on faculty editors, faculty requests, publication and citation data by campus authors, cancellation proposals and cancellation appeals, and accreditation needs. Reports in E-Matrix enable us to analyze usage statistics and review holdings, orders, licenses, and bibliographic data. E-Matrix also drives

continued on page 20

Data-informed Collection ...
from page 18

our public journal and database lists. We are experimenting with the **EBSCO's** usage module, *EBSCONET Usage Consolidation*, to see if it can help us automate the harvesting and analysis of usage statistics (e.g., using Sushi) and if it can help efficiently calculate cost-per-use. We are in early stages of our testing. See Figure 1.

Collection Views

Collection Views is a novel system to demonstrate the value of the collection for specific user communities at **NCSU** (e.g., academic departments and colleges) and to conduct internal allocation assessments using library collections data (e.g., journals and monographs expenditures) and community data (e.g., number of faculty in a given department, grant dollars awarded). It has also been employed in successful advocacy efforts on campus for collections funding. The aim of Collection Views is to help us understand how our expenditures on resources relate to different departments and colleges at **NCSU**. To do this, we mapped departments and colleges to subject codes associated with collection purchases. By providing interactive visualizations within Collection Views, this tool helps us bring together previously disparate data to better understand and assess our collecting priorities for each campus group.

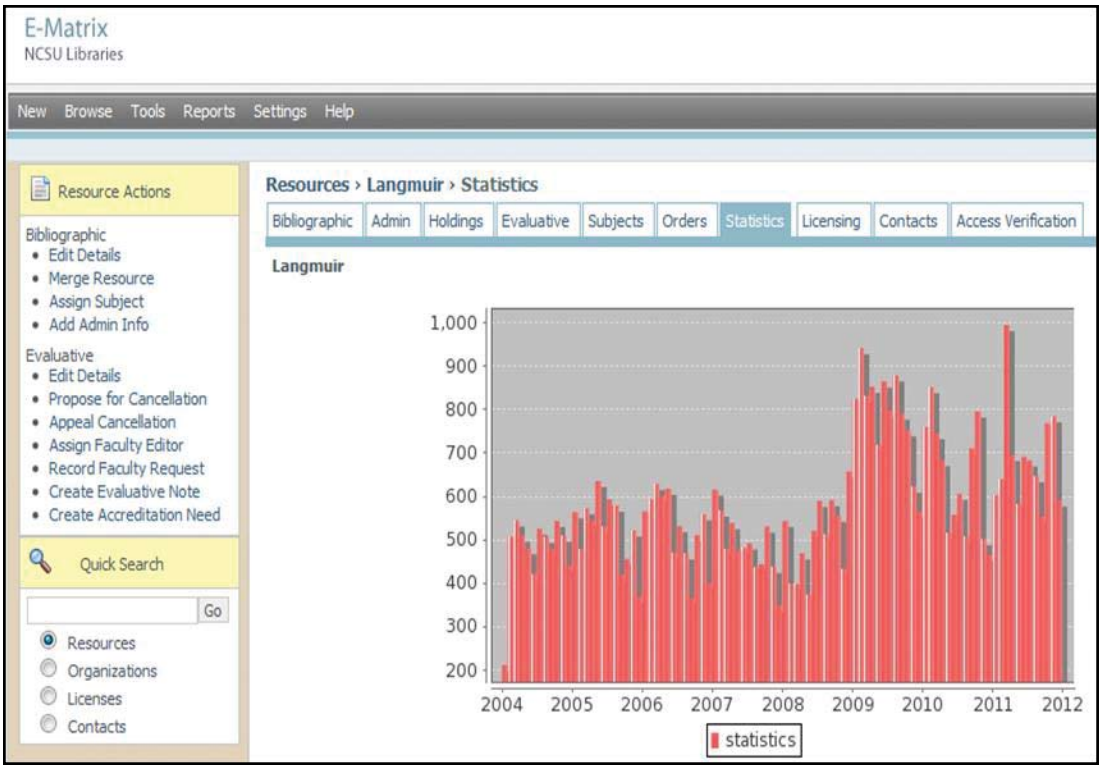

Figure 1: Screenshot of statistics display in E-matrix

SAS (Statistical Analysis Software)

In 2010, the collection management department began utilizing SAS programming and analysis to support collection assessment projects. SAS is a valuable tool because it allows us to slice and dice large amounts of data quickly. For example, one project used SAS Project Management to analyze a twelve-year series of print items and examined the correlation between an item's years in the collection and its circulation status. (**John Vickery's** Print Item Usage Analysis: *http://www.lib.ncsu.edu/collectionmanagement/projects/print-item-usage-analysis/*).

Use Cases

Checking Our Assumptions and Fitting the Collection to Campus Needs

As previously described, Collection Views was built specifically to blend library data with campus data (Figure 2). By combining these data sources, we are better able to assess how our collection funds support different departments and disciplines on campus.

The visualizations bring the data to our fingertips so that we can see, for example, that our allocations for one department may be out of line with the number of graduate students and grant-funded projects of that department. Conversely, we may find that we are over-allocating library funds to a research program that may no longer be a priority for the University (e.g., evident by a decrease in enrolled graduate students or reduced grant funds). Looking at the data in this way prompts us to continually assess how our collection funds are allocated, and it enables us to maintain an appropriate balance in line with campus demographics and strengths (Figure 3). We have used this confluence of data in seeking budget increases from the University for resources that are vital to current research.

continued on page 22

against the grain people profile

Hilary Davis

Associate Head, Collection Management
North Carolina State University Libraries
Box 7111, Raleigh, NC 27695
Phone: (919) 513-0654 • Fax: (919) 513-1108
<hilary_davis@ncsu.edu>
http://www.lib.ncsu.edu/staff/hmdavis4/index.php

PROFESSIONAL CAREER AND ACTIVITIES: Currently, I am Associate Head of Collection Management at the **North Carolina State University Libraries**. My primary role is to work with faculty, students, and researchers in the physical sciences, and I lead the department's data analysis efforts for collection assessment. In 2008, I was named one of **Library Journal's** "Movers and Shakers." I hold an MLS from **University of Missouri-Columbia** and an MS in Biology from **University of Missouri-St. Louis**.

One of the main limitations of Collection Views is that it only gives us insight into how our past decisions (based on allocations made in each prior year) fit with campus demographics and grant income. There is little predictive power in the tool, but with more years of data, we can start to understand trends in how library allocations are distributed to support campus stakeholders.

Additionally, the analyses in Collection Views cannot tell us everything we need know about differences in research and teaching needs across departments, such as the different ways that campus departments make use of library resources. Some collecting areas may be important because they are historical strengths of our collection, even if they do not provide immediate support to particular departments. Because the mapping between departments and fund codes were created by librarians, these can significantly affect the results.

Making Difficult Decisions

In fiscal year 2009/2010, the **NCSU** Libraries faced substantial cuts to its collections budget. As part of those cuts, journal subscriptions had to be canceled. To make the best decisions on which journals to cancel, the Libraries needed to gather as much campus feedback as possible on its list of 1,112 journals proposed for cancellation. The Libraries designed and built the Collections Review tool, a Web form where users could easily record and submit their responses to the proposed cancellation list (Figure 4). The form presented key data points to enable the campus to make decisions on keeping or canceling a title and adding features to help users filter and manage the data.

We used two methods for processing all of the feedback. For the first method, we weighted the rankings by the community of users who provided feedback based on how closely their research and teaching subject areas matched the journal subject areas. This approach was to help minimize the tendency of users to want to cancel journals that were not relevant to their research and teaching (e.g., a biology researcher may have issued a suggestion to cancel all history journals). We supplemented this method by factoring in other data, such as journal impact factors and citation and publication patterns (from the Local Journal Utilization Report LJUR)). At the end of the day, we ended up cancelling 499 subscriptions and supplementing journal access with aggregator databases such as **EBSCO's** *Academic Search Premier* and *Business Source Premier*. More details about the 2009 Collections Review and methodology, including examples, can be found here: *http://www.lib.ncsu.edu/collectionmanagement/projects/collectionsreview/*.

Looking for Trends

The collection budgets cuts in 2009/2010 were not limited to the serials review described earlier. We also needed to make cuts to monograph acquisitions. Our strategy to lessen the impact of those cuts was to make sure our monograph purchases were highly targeted. We wanted to identify areas where we could scale back our selection, especially regarding approval plan coverage of particular call number ranges. We took ten years of item-level usage statistics of all circulating monographs added during the fiscal years 1997/1998 – 2006/2007 (as recorded in our ILS) and mapped that data to the Libraries' approval plan.

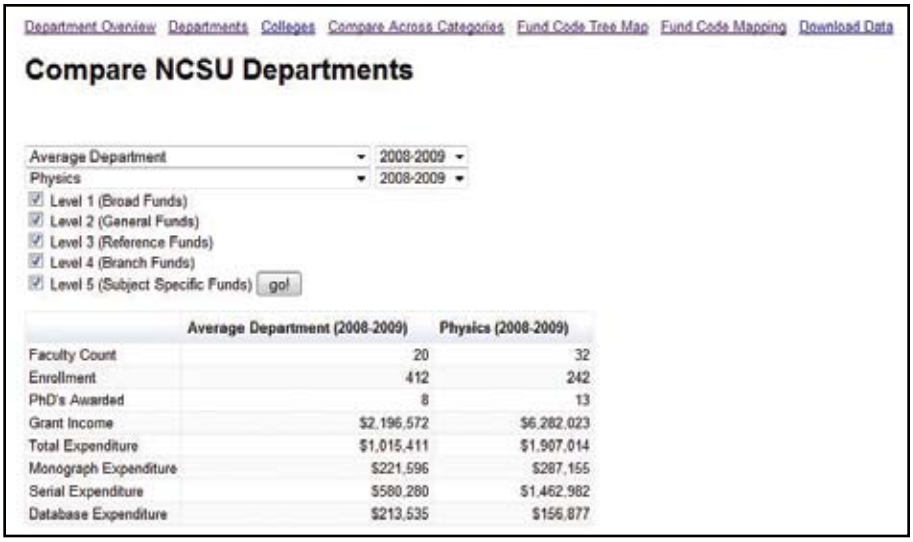

Figure 2: Collections Views Data Portal

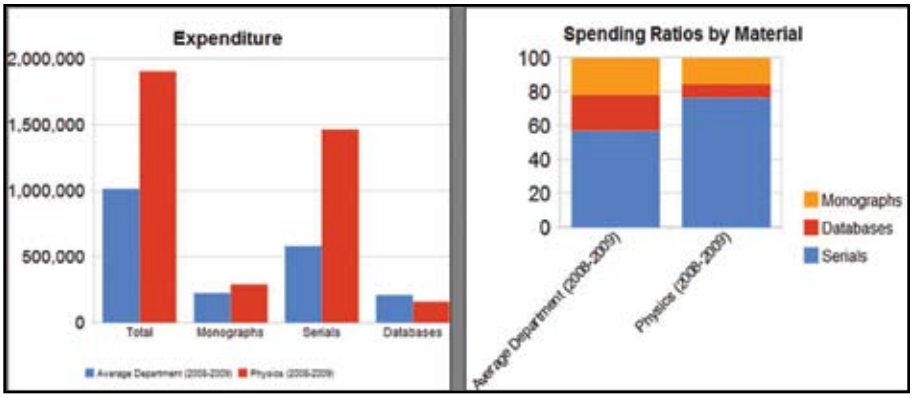

Figure 3: Visualizations comparing the Libraries support for the Physics department compared to the Average department at NCSU

Collection managers reviewed the data and tried to identify patterns in the low-use areas to determine if there were specific call number ranges, publishers, or content levels that weren't circulating. The end result was that we were able to adjust our approval plan and make more targeted firm order selections based on these patterns. We were able to meet our targeted reduction of 20 percent in monographic acquisitions and minimize the immediate impact on our patrons.

The driver for this project was to deal with our budget cut, but we also knew that this study could provide us with information to help us shape our growing patron-driven acquisition (PDA) program. The analysis of the data provided us with several call number ranges in non-core disciplines where there was a clear mix of circulation rates ranging from high to none, and no discernible patterns (e.g., publisher, date range, content level) to explain this variation in use. Putting these call number ranges into our PDA program seemed the next logical step, as we would avoid making speculative purchases while still providing access to these materials.

This project provided us with a rich seam of data that we will to continue to mine. Further uses will include comparisons of use patterns across disciplines, assessment of how historic use may predict future use, and if print use is a predictor of electronic use.

Future Plans and Strategies

While the coordination of library data analysis and reporting activities is assigned to one person, we have been making this kind of work core to every collection management librarian. Our goal is to empower every collection management librarian to have deep competency with searching and running reports in tools such as our ILS, SirsiDynix Symphony, the LJURs, Collection Views, and our ERM system, E-Matrix. In this en-

continued on page 24

Data-informed Collection ...
from page 22

vironment of data-informed collection management, it has become vital for everyone in our unit to work with the data in common productivity tools, such as Microsoft Excel and Access, to be able to interpret meaning from the data and translate that to decisions that impact our stakeholders.

The **NCSU** Libraries have embraced the data-informed approach, and it is now the cornerstone of our collecting program. We strive to keep learning new tools and techniques and integrate those into our processes. Currently, we are working on broadening the use of SAS in the department and utilizing visualization tools more effectively to better articulate the data.

against the grain **people profile**

Annette Day

Head, Collection Management
North Carolina State University Libraries
Box 7111, Raleigh, NC 27695
Phone: (919) 513-2129 • Fax: (919) 513-1108
<annette_day@ncsu.edu>
http://www.lib.ncsu.edu/staff/apday/index.php

PROFESSIONAL CAREER AND ACTIVITIES: I am Head of Collection Management at North Carolina State University. Originally from the UK, I have worked at **University College London**, **University of Warwick**, and the **University of Coventry** before transferring across the Atlantic to Philadelphia and the **University of Pennsylvania**. I then headed down south to my current position at **North Carolina State University**. I earned my MLS from **Leeds Metropolitan University** and hold a BS in Computer Science and Mathematics from **Leeds University**.

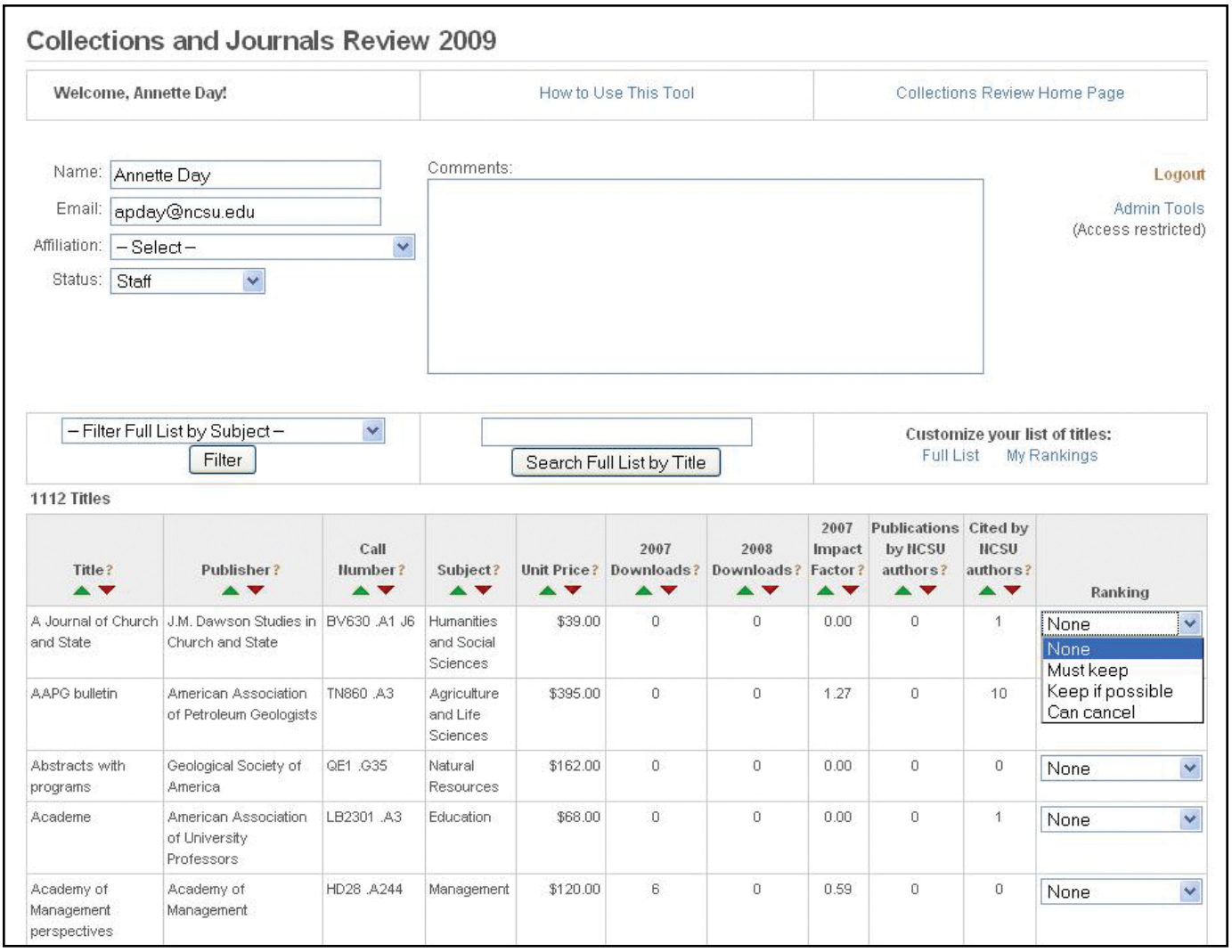

Figure 4: Screenshot of collections review web form used for the NCSU Libraries 2009 serials review