Douglas Black

An accidental datahound

Transitioning skills to experience and application

I didn't set out to become a data-driven librarian. The concept was still new when I was in library school, and smitten as I was with the human element of information structures and information seeking, the idea seemed detached, focused on mere numbers divorced from the daily face-to-face reality of a reference librarian.

Well.

In my first professional position, my supervisor told me about a study performed at that library a few years before that was prompted by librarian complaints about too many lengthy phone calls with distance students while patrons at the desk awaited assistance. A four-month record of telephone interactions showed that the vast majority of phone reference calls lasted less than ten minutes, indicating that the librarians were mistaken. However, the data incidentally revealed something else. While the number of lengthy calls (more than 20 minutes) was the lowest of all call-length categories, the total time spent on those calls exceeded that for other call lengths. The librarians were in fact spending more total time on those long calls than on the more numerous shorter ones, explaining and supporting their viewpoint.

It's a timeworn adage in collection-development circles that a budget is just a snapshot of a moment in time. That study showed me that any set of numbers is a snapshot, which is always taken from a particular angle, with framing and composition. Snapshots can also be viewed from multiple angles, as can our services and our perceptions of how we provide them.

The second insight from that study was that we record reference questions *as answered,* not describing the reference interview or the process of working toward a solution with the patron. So I spent two weeks recording detailed descriptions of my every transaction. The narrative data provided a richer picture of what happened at the desk, recording the question as asked and as answered, with as much of the negotiation and problem-solving as possible from start to finish. What I learned was that useful information can be obtained not just from raw numbers and back-office analysis, but from qualitative methods as well, and that qualitative data can often help place quantitative data in useful context. What happens at the desk amounts to stories, which often have morals. (I also learned that qualitative data can be tremendously laborious to collect and analyze.)

Interlibrary loan (ILL) data is often used to help with selecting material, but it has other applications, as well. In reviewing transaction data to identify the heaviest user groups (in a small liberal arts college, it was senior thesis writers by a wide margin), I also found what ILL staff know: A lot of borrowing requests are cancelled because

Douglas Black is head of collections management at Middlebury College, email: dblack@middlebury.edu

the library already owns the material. On reflection, I realized that those requests represented instances in which patrons didn't successfully find the needed material on their own. While it's rarely passed to other librarians, that data could help identify opportunities for outreach and instruction, or perhaps for modifying a library's website organization or finding tools. This project showed me how data can be useful beyond its obvious applications, and that there are more possible connections among library operations than we might think.

Even usage reports can provide insight into user behavior, especially when viewed in comparison to each other. For instance, COUNTER database reports show searches conducted in a database's own platform, but not discovery-layer searches linking to that database's records.[1] Thus, low search numbers against high record views or full-text downloads could mean one of two things: The content is being located via discovery layer, or users are examining an unusual number of results.

In either case, the content is being used. Many more searches than results viewed can mean only that patrons are searching the database but not looking at much. The content isn't useful, they don't know how to use the database, or they're eventually finding exactly what they want with rare tenacity and without examining many interim results (more investigation would be needed, as they say). What I learned here was that relationships among usage data can help identify opportunities to improve our resources and services, akin to open communications between people in relationships.

Finally, one major review of databases and journal packages required compiling a mass of data on historical costs, usage, and other notes presented to users in the form of a Qualtrics survey—not just to gather input, but also to educate users on what we offer and what it costs. Data can flow in multiple directions beyond mere reporting and internal decision-making. In this case, I discovered that not only could it feed back into collections decisions, it could also serve as part of our user-education efforts.

Overall, data work has taught me the deep degree to which various library operations can function together as a whole, as fluid and faceted as human behavior in seeking and creating information. To paraphrase Ranganathan, the library is a wholly integral organism, and data are for using and sharing.

Numbers themselves aren't the real goal of developing and analyzing data. I'm uncomfortable with using firm cost-per-use thresholds for database decisions, sticking rigidly to subject allocations, or using circulation as the sole determinant of a given title's value (I believe that checkouts represent only one form of usage). Professional judgement still matters more than anything. Rather, data can be one starting point for understanding the conversations that academic libraries fundamentally promote: conversations between readers and authors, ideas, and other readers; between faculty and students; between the library and faculty; and among a host of other constituencies and entities, all aimed at creating knowledge and learning. Developing and analyzing data is itself a form of inquiry into how we do what we do and the ways in which we can make it matter.

Had you asked me all those years ago, I would never have predicted this path. For me, data work has become a lens through which we can view some of the workings of humans seeking interactions with ideas and with each other. It's one way to see the library as an organic whole and as an integral part of the campus community. That dry, impersonal concept has turned out to be an engaging, fascinating, even personal way to explore and communicate about how we perceive, use, and manage information for our patrons and our libraries—and for ourselves.

Note

1. See www.projectcounter.org, which provides standards for reporting usage.