Introduction

Jason Griffey

"When a measure becomes a target, it ceases to be a good measure."

—Goodhart's Law[1]

In the history of libraries, there have been a variety of overarching reasons to collect and analyze usage statistics: their use in determining future actions, their utility in summarizing efforts and activity for funders, and their function for comparative purposes from library to library, among others. Examples of these instances include the ability to track physical item circulation and base future acquisitions on popular materials and using longitudinal circulation statistics to show demand and argue for increases in funding. The specific things that libraries have measured toward these ends have changed over time, and never more so than in the last twenty years.

As the world of information access has pivoted from physical goods to digital screens, the ways that libraries measure themselves have changed. Information use was once measured fairly straightforwardly, by just counting the materials that were circulated to patrons or used in the library. Count the things that are used, add that count, and you get something like usage of the collection. From this relatively simple measure, much can be determined if those counts have additional facets applied to them, such as which books, from what topic areas, for how long, and the like.

Aside from circulation and material tracking, the other common measure for libraries over the last several decades has been the classic door count. How many people come into your building a day? When combined with material usage, this gives you the ability to look at things like circulations per person, another common metric in library reports. It's also a simple measure of how busy a space or building might be, which allows for lots of maintenance and staffing decisions to be made.

For years, these sorts of measurements have been used to evaluate library building usage. They have been, in some libraries (albeit too few), supplemented by observational data, often gathered through sampling during representative times of year. These sorts of sociological studies have been done both by library staff directly and by experts brought in to help the library understand its space usage. Sociologically driven observational studies are a fantastic tool for understanding behavior, but they are limited in many ways. They are always time-limited and rely on statistical validity to be able to generalize the data. Another drawback is that they are biased towards the things that the observer is looking for, and nonstandard or misunderstood behaviors may be miscoded in the study. And, as always, there is the ever-present threat of observer bias.

We are on the edge of a huge set of technological changes that will alter how we can measure our spaces. New advances in sensor technology, artificial intelligence and machine learning, computer vision, and more have brought the ability to monitor spaces in ways that were previously unthinkable. And the near future of this technology will be even more radical, enabling possibilities such as tracking every object in a space constantly at all times. And even people!

Why Measure Spaces

Libraries have always paid attention to their buildings, and funders of libraries (whether local communities or individual philanthropists) have always wanted the spaces to be special in some way. Grand spaces, impressive spaces, the sorts of spaces that inspire awe and reverence—truly the "cathedral of the book," as they have been called. Or modern and sleek, state-of-the-art technological marvels in their own right, places like Dokk1 in Aarhus, Denmark, or the Hunt Library at North Carolina State University that have captured the collective library fascination with space and building.[2] This is only fitting since the physical building is often the single most economically valuable thing that the library owns, even beyond the collection it houses within. A community has an enormous amount of value tied up in the physical space of the library. But the attention paid to the usage of the physical space has not always been balanced with that worth.

It's relatively easy to measure transactional elements. Tracking and measuring items that are interacted with is much easier than trying to tell what people are doing inside a space. When someone checks out a book, clicks a link on a database, or walks through the door, that's a fairly easy action to measure. Answering more complicated questions about library use is much harder. What do patrons pay attention to as they walk through the building? Where do patrons choose to sit, and, more importantly, why do they choose those specific places? Are your spaces more conducive to people sitting by themselves or to groups? How do groups affect other uses of the library?

I believe that having solid numbers behind the use of our spaces is the future of library statistics, especially as they relate to proving worth to funders, citizens, boards, academic provosts, and others who ultimately hold control of the funding streams to libraries. As information seeking increasingly runs to other, nonlibrary sources, the traditional material-based metrics no longer appropriately measure a library's worth to its community. Data collected by sensors and analyzed over time will give librarians far more of an ability to answer questions about use, even when they may not have considered the question yet. Big data gives rise to emergent patterns that are not always expected a priori, and the ability to ask questions of ambient data about a library space is enormously powerful both for understanding current use and for planning future use.

Measuring and reporting a library's worth to its local community is a tough thing to do well. There has been, over the last few decades, an acceptance of reporting the worth of a library as something like return on investment (ROI). ROI is a measure of profitability used in the economic study of business and is reasonably easy for everyone to understand. ROI is usually communicated in terms of money spent to support libraries and money returned to the community as a result of this expenditure. "Every $1 invested in the library returns $*X* to the local community" is the normal sort of phrasing, and while it's an eye-catching way of justifying library spending, in my opinion, it is also a dangerous one. ROI is by its nature an economic measure and one rooted heavily in the concept that money is to be invested for return rather than used to generate public and civic good. If libraries lean too heavily on the rhetoric of profit and return, we lose the messaging of generalized public good, of raising communities to be better than they were, of trying to approach greater goods like equity, justice, and an informed and educated populace.

It is also dangerous because, of course, if you accidentally create targets rather than measures, as Goodhart so pithily put it, you run into trouble.[3] By relying only on numbers like ROI, you are only ever allowed to go up, and dips in those numbers must be explained. Rather than being understood as a community good, you are a community investment, and if that investment is depreciating, then it must be fixed in some manner. If libraries have more robust descriptive stories to tell about their impacts, especially if we can tell those stories with quantitative data and do not fall prey to simple economic measures, I believe we will be far more able to thrive, even in economically turbulent times.

So what can we measure instead? What sorts of numbers should we be reporting? I don't think there is a single answer to those questions just yet, but the goal of this issue of *Library Technology Reports* is to illustrate what some potential answers might look like. Much of what's now possible to measure is due to the rise of inexpensive connected sensors and other technologies. Let's take a look at those and what they might enable us to gather and report.

Technology

The statistics most commonly used to judge a library's success can be collected in an automated fashion. This only makes sense, as automated counts are collected without staff attention being necessary and can be collected over long periods of time, allowing for comparative analysis that is more difficult with other collection methods. The progress of technology enabling more and more data to be collected automatically is a big part of the new potential for "smart" spaces and buildings in libraries. The incredible rise of the mobile phone as the primary computing platform for the world has helped to drive down the cost of a number of technologies and enable the measurement of things in the world in ways that would have been science fiction only a few years ago.

Modern smartphones are a wonderland of different sensors. Accelerometers that measure movement, light and infrared sensors that see light levels and distances, microphones that measure sound levels and cancel extraneous noise, cameras that can take incredibly detailed images for later analysis, and much more are in a modern smartphone. The explosive adoption of mobile phones also means that the cost of these individual components has fallen through the floor and that they are available for other projects at a very reasonable cost. Combine one or more of these sensors with an inexpensive microcomputer or microcontroller platform like the Raspberry Pi, the Beaglebone, the Arduino, or one of a dozen more, and you've got a data collection device. All of these are now so cheap that it's almost trivial to work with them, and hardware is almost never the limiting expense for computing at this point in history.

Raspberry Pi
https://www.raspberrypi.org

Beaglebone
http://beagleboard.org/bone

Arduino
https://www.arduino.cc

Internet of Things

The current shorthand for the sorts of devices we're talking about is the Internet of Things, where the ultimate state of being is that computing and communications device costs go to nearly zero, which enables every object in the world to be connected to the internet. This would have the effect of making every object in the world a sensor, enabling everything from your water bottle to your pencil to report to a server somewhere its current status, location, and the like. If this sounds like a dystopia to you, I'm not sure you're wrong, but that's definitely where Moore's Law is pushing us.[4]

Every microphone is also a speaker, and every camera is also an interface. It is often the case that something we consider a sensor can also act upon the world and that the things we put out are not only passive collectors of data but can also relay actions to other systems that make those systems more efficient. This is at the core of the idea of a *smart building*, where the structure itself has a robust set of sensors and controllers that are all interconnected and inform the holistic management of the building. If you're reading this and are over the age of about twenty, you probably remember your first interaction with automatic or motion-activated lights that come on when you enter a room and go off after not sensing movement for a preset amount of time. These were the early, early predecessors of the smart building, where the environment automagically adapts itself to the presence of a person or people in it. This technology expanded quickly into heating and air conditioning units, where the presence of people determined whether a space was heated or cooled. It's not a surprise that these were the first few bits of a building that were automated, in that lights, heat, and cooling are all at the top of expenses for upkeep of a building. Managing them more efficiently is a huge cost savings to building managers.

It's a short leap from "turn on the lights" to "adjust the temperature" based on whether someone is in the room. It's a longer and harder problem to do more finely detailed actions, from customized temperature controls based on the number of people in a space, to truly individualized services that respond to who someone is, not just their raw presence. You can think back to various science fiction examples to imagine a situation where a room might "know" who someone is and adjust lighting and temperature, play music (or not), lower the blinds, and the like, based on that person's specific preferences. These sorts of things are possible at this point for private residences using commercial smart home technology. For example, at my own home, when I approach my house in my car, the lights on my porch come on and the thermostat inside sets the temperature from "away" to whatever I have set as comfortable for the time of year. The porch lights turn themselves off a few hours after sunset, and when I lock my front door, the temperature automatically sets itself to an "away" mode.

All of these interactions are easily done in a home, but public spaces are enormously more complicated. While a private home has a known set of users (in general), a public space can be used by literally anyone in the community. This makes individual personalization very difficult, although for some library types some level of this identification could be done. For example, I have seen academic libraries where students must use their ID cards to enter, and the card swipe or tap triggers a sort of "welcome" display on a panel in front of the doors. The example I saw at the University of Technology in Sydney, Australia, welcomed the student by name ("Hello, Jason!") and reminded the student where in the building the materials specific to his or her major were ("The books on biology are on the 3rd floor, to the left, call numbers *X* through *Y*"). This was several years ago, and I can easily imagine an extension of this sort of smart building where the library could prompt the user for other sorts of resources and even maybe adapt to the user's presence.

New Computing Abilities

The other huge advance in smart systems is the growth in computing power over the last decade. Computing power gets better on a mostly predictable schedule, but we've crossed a line in what very inexpensive computers can do that seems like a sea change in capabilities. Two of these areas that will be transformative over the next several years are computer vision and the area of artificial intelligence, machine learning, and expert systems. These are related, and all of them blur in interesting ways, and this is sometimes difficult to explain. All of them, however, are focused on getting computers to increasingly do things that previously were the domain of human judgment.

Let's start with the easiest to explain, computer vision. Computer vision systems work with still or video images and attempt to recognize or classify things in the images, creating metadata about what's in them. This is related to a specific type of image recognition with which librarians might be familiar—optical character recognition (OCR). OCR systems take photos of text, recognize the letterforms, and transform them into text that can be manipulated by computer systems. OCR counts as a form of computer vision, but these days the phrase is used to refer to systems designed to be much more general in their object identification rather than being limited to just text. For instance, the facial identification that most popular photo systems use (Apple Photos, Google Photos, Flickr, Instagram, SnapChat, and more) is a form of computer vision. The ability to take multiple photos and have the computer tell you that these two have the same person in them is one type of computer vision recognition.

The more interesting things happening these days, though, are when computer vision systems are expanded into machine learning systems and aren't programmed directly but instead are trained on existing photo sets. Let's say you wanted to have a system that would answer the question "Is there a cat in any of these photos?" The modern way to tackle this would be to feed photos of cats to a computer vision and machine learning system and tell the system that all the photos have cats in them. The system itself then builds an identification system for things called "cats," and when you give it further photos, it should be able to label the contents appropriately either "cat" or "no cat." This is far more powerful than having to describe painstakingly to the computer what "catness" is.

This gives rise to being able to use cameras for a variety of statistical data gathering because you can now throw the images into a computer vision system that will extract from them the data you wish to capture. Later in this report we will discuss at length the Measure the Future project, which uses computer vision to show how library spaces are being used by patrons. Systems such as this will be more and more prevalent over time, and the use of similar systems will likely be a part of smart buildings before much longer.

The long game for the Internet of Things is far stranger than adjusting temperatures and turning on lights. It's also going to be used for far more than customizing services to patrons. I'll go into some of the potential for this technology later in chapter 5 ("Future Directions").

Iterative Design

One of the goals of better understanding the physical spaces of the library is to work to improve them for patrons, and doing so using sensors and Internet-of-Things-style data gathering allows for continuous data gathering. This approach is in contrast to the sampled or staggered data that is used by some libraries now. The huge advantage of continuous data is that you can iterate much faster and test the physical space in the same way that you can test digital spaces now.

Amazon, Google, and all of the major websites do continuous A/B testing, presenting slightly different pages to users as often as every time the page is loaded. They track which are more effective, for whichever metric they are measuring (buy the item, click the button, find the thing faster), and change their pages for everyone based on this continuous improvement. Doing this continuous improvement work to physical spaces is difficult without appropriate data, but of course smart spaces solve this problem. A/B testing spaces, even if they are just measuring how people react to a new display and then changing it in response to the data, could be immensely powerful for improving how patrons see and use spaces.

In This Report

The goal of this issue of *Library Technology Reports* will be to give librarians and other interested parties a look into what's possible in the current state of technology for smart buildings, as well as to point in useful directions for the near future of the Internet of Things and other sensor technology. Part of the challenge in doing these sorts of projects in libraries, specifically, is that libraries have a much higher expectation of sensitivity to privacy and personal data than other situations, such as a corporate environment. The lengths that libraries should go to protect patrons from potential privacy leaks are enormous. Libraries should take privacy as a primary position and security of data gathering and handling as a duty to the people they serve. This makes these sorts of robust data collection endeavors very complicated. Data is toxic over time, and risks increase as more and more data is gathered, as it could be combined in order to de-anonymize patrons or otherwise place risk onto those we serve. Librarians should think very, very carefully about the privacy and security implications for data-gathering devices and, when working with commercial providers or other vendors, should insist that said providers have a security plan and have thought through what their stance is on data collection and retention.

This report will look at three different projects that involved space metrics and analysis in libraries: Virginia Tech; Concordia University Libraries in Montreal, Quebec, Canada; and the Measure the Future project. Each is using technological tools to analyze library spaces in order to make the environment better for its patrons. In the case of Virginia Tech, the library used furniture movement as a stand-in for patron activity. Concordia University Libraries was interested in helping patrons sort out where they wanted to be inside the library, and so it looked into monitoring and then displaying the sound levels for public areas in the library. Measure the Future is using computer vision to see how patrons move around in library spaces and derive "attention" measures from those movements while doing so with a strong protection on any sort of identification of patrons. Finally, we will look at what the next five to ten years of technological progress will bring and how that might change the possibilities for a smart library.

Library Metrics

Below is a list of papers, websites, presentations, news stories, and other resources that have touched on the idea of sensors and space measurement over the years. While these references aren't necessarily cited in this work, they point towards the concepts and ideas that brought this issue of *Library Technology Reports* together. They can provide you with a more thorough look at what's possible, what's been done, and where we should be headed in this area of understanding.

Ayre, Lori Bowen. "Wireless Tracking in the Library: Benefits, Threats, and Responsibilities." In *RFID Applications, Security and Privacy*, 229-243. Reading, MA: Addison Wesley, 2005.

Bradley, Jonathan, Patrick Tomlin, and Brian Mathews. "The Smart Commons: An Experiment in Sensor-Based Space Assessment of Learning Environments." In *At the Helm: Leading Transformation: The Proceedings of the ACRL 2017 Conference, March 22–25, 2017, Baltimore, Maryland*, Chicago: Association of College and Research Libraries, 2017: 637-647. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2017/TheSmartCommons.pdf

Buchanan, George R. "The Fused Library: Integrating Digital and Physical Libraries with Location-aware Sensors." In *Proceedings of the 10th Annual Joint Conference on Digital Libraries*, ACM, 2010.

Clasper, Emily. "Library Metrics and Measurement: Counting What Counts & Making it Matter." Presentation from Wild Wisconsin Winter Web Conference 2014. Posted on SlideShare, January 8, 2014. https://www.slideshare.net/eclasper1/library-metricsandmeasurement-wwwc14.

Cintron, Johnathan, Devlyn Courtier, and John DeLooper. "Testing Three Types of Raspberry Pi People Counters." *Code4Lib* 38 (2017). http://journal.code4lib.org/articles/12947.

Dimitrov, Konstantin. "Arduino/Genuino 101 BLE Thermometer with TMP102 and Blynk." Arduino Project Hub. January 14, 2017. https://create.arduino.cc/projecthub/TheGadgetBoy/arduino-genuino-101-ble-thermometer-with-tmp102-and-blynk-ab5984?ref=platform&ref_id=424_recent___&offset=29.

Frayer, Lauren. "High-Temp Sensors Help Old Port City Leap into Smart Future." *Parallels*, NPR, June 4, 2013. http://www.npr.org/sections/parallels/2013/06/04/188370672/Sensors-Transform-Old-Spanish-Port-Into-New-Smart-City.

Gil-Garcia, J. Ramon, Natalie Helbig, and Adegboyega Ojo. "Being Smart: Emerging Technologies and Innovation in the Public Sector." *Government Information Quarterly* 31 (2014): I1-I8.

Given, Lisa M., and Gloria J. Leckie. "'Sweeping' the Library: Mapping the Social Activity Space of the Public Library." *Library & Information Science Research* 25, no. 4 (2003): 365-385.

Greenwalt, R. Toby. "In Search of Better Metrics." The Wired Library, Public Libraries Online, May 7, 2013. http://publiclibrariesonline.org/2013/05/in-search-of-better-metrics.

Hyman, Jacob Andrew. *Computer Vision Based People Tracking for Motivating Behavior in Public Spaces*. Thesis (M. Eng.), Massachusetts Institute of Technology, 2003.

"interlibnet." "Proving Our Worth: Library Measurement and Metrics." April 14, 2015, International Librarians Network. https://interlibnet.org/2015/04/14/proving-our-worth-library-measurement-and-metrics.

Kuznetsov, Stacey, and Eric Paulos. "Participatory Sensing in Public Spaces: Activating Urban Surfaces with Sensor Probes." In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, ACM, 2010.

Linn, Allison. "Decades of Computer Vision Research, One 'Swiss Army Knife.'" *The AI Blog*, Microsoft, March 30, 2016. https://blogs.microsoft.com/ai/2016/03/30/decades-of-computer-vision-research-one-swiss-army-knife.

McLaughlin, Jenna. "Tor Could Protect Your Smart Fridge from Spies and Hackers." The Intercept, July 20, 2016. https://theintercept.com/2016/07/20/tor-could-protect-your-smart-fridge-from-spies-and-hackers/.

Molnar, David, and David Wagner, "Privacy and Security in Library RFID: Issues, Practices, and Architectures." In *Proceedings of the 11th ACM Conference on Computer and Communications Security*, ACM, 2004.

Pundsack, Karen. "21st Century Library Measuring Sticks." *Public Libraries Online*, November 23, 2015.

http://publiclibrariesonline.org/2015/11/21st-century-library-measuring-sticks/.
"SensorTape." Vimeo video, 0:31. Posted by Responsive Environments. February 12, 2016. https://vimeo.com/155159411
Yan, Wei, and David A. Forsyth. "Learning the Behavior of Users in a Public Space through Video Tracking." In *Proceedings of IEEE Workshop on Applications of Computer Vision (WACV)*, 2005: 370-377.

Notes

1. Marilyn Strathern, "'Improving Ratings': Audit in the British University System," *European Review* 5, no. 3 (July 1997): 308, http://journals.cambridge.org/abstract_S1062798700002660.
2. See Kristin Hohenadel's article "The Library of the Future Is in Denmark," Slate, August 25, 2016, http://www.slate.com/blogs/the_eye/2016/08/25/dokk1_in_aarhus_denmark_is_the_best_new_public_library_of_2016.html and James B. Hunt, Jr. Library at North Carolina State University's video, "The Library of the Future" https://www.ncsu.edu/huntlibrary/watch/.
3. Strathern, "'Improving Ratings,'" 308.
4. Gordon E. Moore, "Cramming More Components onto Integrated Circuits," *Electronics* 38, no 8 (April 19, 1965).