ALA American Library Association

LIBRARY SPACES AND SMART BUILDINGS

TECHNOLOGY, METRICS, AND ITERATIVE DESIGN

Jason Griffey

Library Technology Reports
Expert Guides to Library Systems and Services

JAN 2018
Vol. 54 / No. 1
ISSN 0024-2586

Library Technology
R E P O R T S

Expert Guides to Library Systems and Services

Library Spaces and Smart Buildings: Technology, Metrics, and Iterative Design

Edited by Jason Griffey

ALA TechSource
alatechsource.org

American Library Association

Library Technology REPORTS

ALA TechSource purchases fund advocacy, awareness, and accreditation programs for library professionals worldwide.

Volume 54, Number 1

Library Spaces and Smart Buildings: Technology, Metrics, and Iterative Design

ISBN: 978-0-8389-1610-0

American Library Association

50 East Huron St.
Chicago, IL 60611-2795 USA
alatechsource.org
800-545-2433, ext. 4299
312-944-6780
312-280-5275 (fax)

Advertising Representative

Samantha Imburgia
simburgia@ala.org
312-280-3244

Editor

Samantha Imburgia
simburgia@ala.org
312-280-3244

Copy Editor

Judith Lauber

Production

Tim Clifford

Editorial Assistant

Colton Ursiny

Cover Design

Alejandra Diaz

Library Technology Reports (ISSN 0024-2586) is published eight times a year (January, March, April, June, July, September, October, and December) by American Library Association, 50 E. Huron St., Chicago, IL 60611. It is managed by ALA TechSource, a unit of the publishing department of ALA. Periodical postage paid at Chicago, Illinois, and at additional mailing offices. POSTMASTER: Send address changes to Library Technology Reports, 50 E. Huron St., Chicago, IL 60611.

ALA TechSource
alatechsource.org

About the Editor

Jason Griffey is a librarian, technologist, consultant, writer, and speaker. He is the founder and principal at Evenly Distributed, a technology consulting and creation firm for libraries, museums, educational institutions, and other nonprofits. Griffey is an Affiliate Fellow at the Berkman Klein Center for Internet and Society at Harvard University and was formerly an associate professor and Head of Library Information Technology at the University of Tennessee at Chattanooga. Griffey was a winner of the Knight Foundation News Challenge for Libraries in 2015 for the Measure the Future project (http://measurethefuture.net), an open hardware project designed to provide actionable use metrics for library spaces. Griffey is also the creator and director of the LibraryBox Project (http://librarybox.us), an open-source portable digital file distribution system. Griffey has written and spoken internationally on topics such as the future of technology and libraries, personal electronics in the library, privacy, copyright, and intellectual property.

Abstract

We are on the edge of a huge set of technological changes that will alter how we can measure library 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. In *Library Technology Reports* (vol. 54, no. 1), "Library Spaces and Smart Buildings: Technology, Metrics, and Iterative Design," I'll explore these technologies and provide librarians and other interested parties with a look into what's possible in the current state of technology for smart library buildings. Looking at three different projects that involved space metrics and analysis in libraries, this report shows how Virginia Tech; Concordia University in Montreal, Quebec, Canada; and the Measure the Future project are using technological tools to analyze library spaces to improve their environment for their users. Virginia Tech is researching how furniture movement acts as a stand-in for patron activity. Concordia University experimented with a project that monitored noise levels. The Measure the Future project 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.

Subscriptions

alatechsource.org/subscribe

Contents

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).

How to Measure the Future

Jason Griffey

There have been a lot of words written about the Measure the Future project, and even more presentations given about it over the last two years. What I hope to accomplish here in this issue of *Library Technology Reports* is not so much to revisit what's been said about the project. I also don't want to write a pitch or marketing message for the project. What I really want to do is to tell the story of Measure the Future, from the initial ideas that started it on its way to the current state and where we hope to be in the future. Part of what I want Measure the Future to do is help libraries tell better stories about themselves, to have data to back up the stories that they tell, and hopefully to have stories to tell that they didn't even know about. The stories that libraries tell about themselves have changed over the years, and I think they will continue to change and evolve even further over the next several years. If I want Measure the Future to be a part of telling those stories, maybe I need to tell its story first.

Measure the Future
http://measurethefuture.net

Setting the Stage

It was 2011 or maybe 2012. I was gathering statistics as part of my role as head of information technology at the library at the University of Tennessee at Chattanooga, and I was getting more and more fed up with numbers by the minute. The sorts of things that we tracked, according to our requirements for accreditation and for ACRL, were useful only to compare us to other institutions in ways that didn't seem particularly meaningful in the modern age. These numbers didn't tell me much about how to make our services and spaces better for our patrons. And even numbers that might guide us in better acquisitions practices were incredibly difficult to pull from the morass of database vendors.

I started thinking, trying to consider carefully what might be important for libraries to know, what might give us insights into how patrons were using libraries. The more I thought about what was likely to be important and what was also a huge gap in our knowledge, the more I was convinced that we needed lots more data about our building use. The library building is, in most cases, the library's most valuable fixed asset. The library building is a huge aspect of the library's net worth, and yet we don't focus our attention on how it's used in the same way we look at our materials. So how could we start to better understand how our buildings were being used?

Moreover, as collections shift from physical to digital, communicating the importance of the physical building to those who oversee the funding for libraries is a key for future growth.

That question was the central one around which I began to brainstorm ideas. It was immediately obvious to me that since we didn't have really any statistics about use of the building (other than gate counts of number of people who walked in), some thought about what to gather and how it might be gathered was the first order of business. I realized that what I wanted was a system that would tell me how people were using the space. Were they coming into the library, collapsing into a chair, and not moving for hours? Were they just using the building as a pass-through to another location? Were patrons using the stacks to browse for items at all? Did patrons use our big tables in groups, or was it just one person camping all day? All of these questions seemed valuable to try and sort through.

The next step was to try and see if there was a way to capture that data. Luckily, I had some experience with small electronics platforms, like the Arduino and the Raspberry Pi, and so I knew there were at least a dozen ways one could approach the problem. Measuring occupancy per room was part of the challenge, but we also needed to be able to parse what the people were doing in the space. Not individually, perhaps, but as a collective, what sorts of actions are people doing while they are in the library space? This makes door sensors and other point-of-contact sensors difficult from a logistical point of view. There are just too many points of contact to wire all of them in any given space, much less across different libraries.

A number of commercial entities at the time were using cellphone signals as a stand-in for "individual" movement. By using Wi-Fi or Bluetooth signals or both from mobile phones moving around a space, you can track where people are and what they are doing with a high degree of accuracy. Big box retailers and the like have used this technology for years, and it was starting to trickle into the cost range where libraries were beginning to play with Bluetooth beacons and other types of tracking technology. After looking at the options, I abandoned this idea for my project for several reasons. The largest is that it simply cannot assure privacy in a way that I was comfortable with implementing inside of a library. There are mechanisms for "anonymizing" data from mobile connectivity, but (especially at the time) I didn't feel that they were enough for me to be confident in protecting the identity of patrons in the library.

The privacy issue compounded with the somewhat obvious problem—if we track people with cell phones, we have information only about people with cell phones. This ignores many of the patrons of public libraries, like children, the homeless, recent immigrants, and more. Putting together a system for making decisions about library services and then ignoring swaths of the community that would likely benefit the most from library services did not seem like the wisest course of action. Between privacy issues and this selection bias issue, I took Wi-Fi and Bluetooth tracking as a method for our new tool off the table.

My thoughts turned to imaging. What was the possibility of using some kind of image sensor to capture the whole space at once and analyze how patrons were moving? If we could do this without actually taking pictures, just by capturing the location of people in the space without any identification, then it would pass the security test (more on that later). Images would also gather everyone equally, without bias, across the types of technology the patrons used. One possibility was using an infrared sensor that looked for body heat, but after putting together a quick demo using a standard webcam as a data source, I realized that it was possible to use computer vision to solve this problem without the added expense of the infrared camera.

Decisions and Solutions

Once there was a demo in place, the project applied to the 2015 Knight Foundation's News Challenge for Libraries, a grant round for funding ideas that would benefit libraries around the US.[1] The newly named Measure the Future project was one of eight winners of the News Challenge, which gave us funding for initial development of the project and supported the development through the fall of 2016.

Our first set of decisions revolved around which hardware to settle on. We needed a microcomputer, something that was capable of running some amount of computer vision locally on the sensor itself. This was due to an early decision that was made to start by having any processing of the location data done locally, on board the device itself. This was partially because it was slightly easier to build and could be realized faster. It was also done because we realized very early that in order to make libraries comfortable with installing cameras in their spaces, there had to be a good security story to tell. The best security story is that the data is collected locally, processed locally, never leaves your building, and doesn't include any information about your patrons—and so that's the tool we set out to build even though in some ways it was more difficult than other possible solutions.

The obvious answer for which microcomputer platform to use was the Raspberry Pi, the most popular small computer in the world. The only problem was that at the time, the current Raspberry Pi model (Model 2) didn't include wireless networking by default. In order to get Wi-Fi, you had to buy a separate USB Wi-Fi adapter and then hope that it was stable and ran well on the operating system—neither of which was an assumption I was willing to make. USB Wi-Fi dongles are notorious for their flakiness, and for a device that I was hoping to install in libraries around the country, I needed something far more reliable. We looked for a board that would run the software needed, that had Wi-Fi on board, and that was low-power enough to not need any sort of special attention paid to it over time. We found that in the Intel Edison and began development of the alpha units in 2015.

The other aspect of the project that is worth calling attention to is that it is being built using open source code, and all of the code that we have developed is also being released via an open source license on our

Github repository. We are using standard tools in the development and are sticking with standard web technologies for the user interface. Raspian, OpenCV, Go, Python, React.js, and the other tools used to build this project are well understood and openly supported, with no proprietary or controlled code that can cause issues with some vendor software. The data is in a standard JSON format, and libraries that implement Measure the Future have direct access to the sensors and software. They also own the data they collect; it isn't collected by Measure the Future without permission and request. We are dedicated to making these tools as widely available as possible in order to enable libraries everywhere to be able to test and implement them. Using open and easily available hardware, 3-D printable cases that are made available for reproduction, and open source code that is licensed for sharing and reuse is the best way to do this.

Measure the Future Github Repository
https://github.com/MeasureTheFuture

Design and Development

The next goal of the project was to develop software that would use a standard webcam attached to the Edison to act as a sensor and gather data points as people move through a space. Capturing the location and duration of movement for each recording individual as they move through a space and recording those data points to a database was the first order of business. Clinton Freeman, a developer located in Cairns, Australia, was recommended to me as someone with the technical background to be able to pull this off. Clinton had worked with both health care and libraries in the past and had a great grounding in the sort of privacy issues that arise from using cameras in public and how libraries, librarians, and patrons might react to them. Clinton understood from the beginning the sort of issues we needed to avoid and quickly became the primary developer of the project.

Measure the Future gathered information from two initial partners in the design stage of the project, the State University of New York at Potsdam library, directed by Jenica Rogers, and the Meridian Public Library in Idaho, directed by Gretchen Caserotti. Both were involved in early discussions that set the path for the project development and initial goals. Among other librarians who helped in the initial design phases, particularly in some of the key early thinking, Andromeda Yelton was invaluable. She helped in thinking hard about the privacy model we should follow and in the development of the early UI and UX models for the project.

Security

Several security principles arose from these early discussions. The first was that the alpha units would concentrate on gathering the information and acting as a distribution point for the gathered statistics with no central server architecture. The sensors wouldn't yet talk to a central server due to complexity and implementation difficulties in local libraries. Instead they would act as individual "islands" of data gathering, and libraries could query the individual sensors to see a current heat map of the space or to download the data for analysis. It was clear even in these early stages that the end game for the project needed to be a central visualization and data analysis server that would gather multiple sensors in multiple branches together in one interface. That complexity, however, was well beyond the minimum viable product stage, and we wanted to prove worth before we embarked on that much more involved and difficult process.

The second principle was linked to the privacy issues inherent in gathering data about patrons in a library. We decided that a standing goal would be to never gather any information that could be used to personally identify individuals. This approach complicates many aspects of the project, not the least of which is that as a result of this decision, we are forced into a corner with the way we interpret and can present data about patrons in the space. If the system can't tell Person 1 from Person 2, it has no way of determining if Person 1 enters and exits the area being measured. It simply says "oh look, another person," and counts Person 1 as another unique patron. This means that "patron counts" using Measure the Future are necessarily fuzzy, but the other options for dealing with the issue all led to the potential for patron identification, especially if multiple types of data for a given time period existed. So we made the conscious choice to make our data slightly less precise in service of being extra cautious about patron privacy. I think that's the correct call to make, although it is an incredibly common request from libraries I have spoken with about the project.

The way I describe our approach to security is that we are attempting to measure the space, not individual library users. We're dealing with aggregate movement data and anonymous individuals with no visual information stored for later analysis. We're not even saving the "blob size" information because that could theoretically be used to de-anonymize someone in specific circumstances. Instead, we store only the center location of the identified blob, reducing the ability to identify individuals. We store data in fifteen-minute "buckets" of data as well, in order to prevent identification attacks that rely on precise timing of individuals in spaces. This doesn't reduce the value of the aggregate data, nor even of the movement data; it just

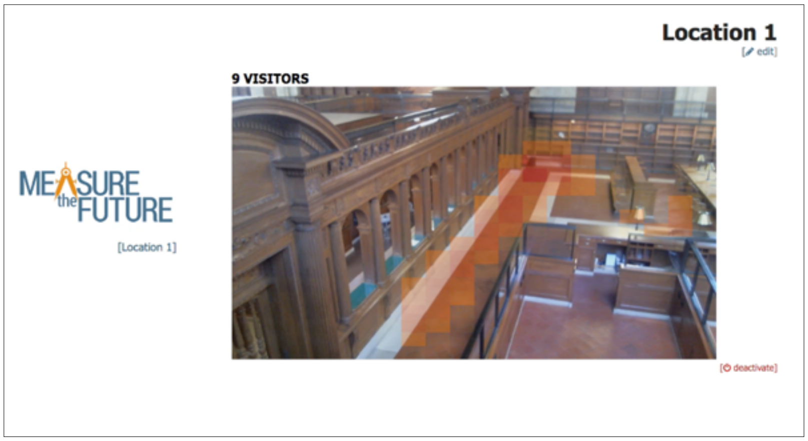

Figure 2.1
Early Measure the Future Interface

prevents precise identification of individual patrons.

Technically, we also ensure that the connections between the sensor and the device used by librarians to view and download data are secured via WEP2 and strong passwords, as well as strong passwords at the system level. It isn't exaggerating to say that we spent nearly as much time discussing and modeling our security plan as we did designing the rest of the system. Moreover, as we move into our more connected Beta development round, we will maintain this focus on security, even as we move to a more cloud-based data visualization and aggregation service.

How Measure the Future Works

Measure the Future works by using a webcam as a sensor for a computer vision system running on a microcomputer. The webcam is placed in a position such that it can have a vantage point to "watch" the space, which normally means as vertical and overhead as we can get. Most installations have been high and at an angle, not truly overhead, although more is better than less for the camera to be able to capture accurate data. The system is calibrated by taking a single reference image, preferably when the space is clear of people. Once calibrated, the sensor is switched into Measurement mode, where it is actively capturing data about movement through the space (see Figure 2.1).

Once per second, the system checks the image sensor in the camera and compares it to the calibration image. Areas that are different are analyzed for size, and if it fits within the settings boundaries, then the different area is identified as a computer vision "blob." Believe it or not, a blob is actually a technical term in computer vision work and designates a contiguous area of pixels that the system should keep track of, identify, or watch. The size of a blob is variable and can be adjusted in the settings panel in order to prevent either false positives (huge shadow moves across the room due to a window) or false negatives (missing people because the sensor is far away and they appear too small).

A blob is identified as soon as it enters the frame, and while it is in the frame, every second another data point is created that notes the location of the blob in X,Y coordinates that are mapped to the calibration image. Each data point is also timestamped with a duration of time. With the calibration image, coordinates, and timestamps, each blob can be tracked through the space in question. You can see how patrons move through the space, where they stop and linger, where they congregate, and where they never go. Over time, you can see what areas in your space are popular and what areas aren't used by patrons. You can query the data to tell you how many people stopped by the new book display and how long on average they spent there.

In the current release, the default display for librarians using the system is a cumulative heat map of the space with controls for calibration and for downloading the sensor data locally. The data is stored on the sensor in a relational database, but the download link on the interface provides easy-to-use JSON formatted files and the calibration image in a zip file. This gives the library all it would need to do whatever sort of data analysis it would like, from advanced heatmaps (see Figure 2.2) to patrons counts to specific location queries.

Sensor units can be installed in fixed locations, for gathering data over time about a specific space, or they can be moved in a more tactical process of measuring specific locations or programs for limited times. Measuring the usage of the library reading room is a great use case, but so is gathering data on a new book display to see how patrons are interacting with it. As the system develops, I hope to see libraries using it in ways that we never expected. That is, for me, the measure of an interesting technology project. As William Gibson famously wrote, "the street finds its own uses" for technology.[2]

Alpha Testing

For our alpha testing of the system, the project had the opportunity to be a part of the reopening of the Rose Reading Room in the New York Public Library in the fall of 2016. We really could have found no bigger stage, nor larger room, in which to try the first installation of the sensors. Six of our alpha sensors based on the Edison were installed in the fall of 2016 and were left to run over the course of the fall and winter.

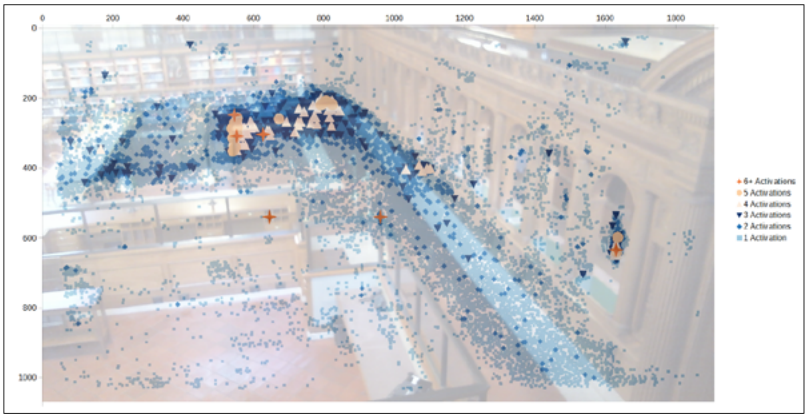

Figure 2.2
Detailed Heatmap of Measure the Future Data

is made, each sensor would check the time the browser had and set the time on the board accordingly.

This seemed like a good solution to the issue, and in testing it seemed to work beautifully. We could set up a new sensor, start collecting data, download the data, and the timestamps were all correct. When we did the initial setup of the sensors in NYPL, we calibrated and tested the units and started collecting data, checked the data, and everything looked great. NYPL staff collected data over the next few days, and again in checking the data for dates (downloading and checking the beginning of the file and then scrolling to the end to compare timestamps), everything looked great—until, of course, we started doing visualizations. When we put the data into a visualization, the timestamps didn't make any sense at all, and so we dug in to see what was going on.

What we discovered was one of the strangest bugs that I've dealt with in my time building hardware like this. The sensors had, it turns out, been turned off at night with the lights in the room—they were on the same circuit, and when the master for the room was turned off, so were the sensors. They then came back on when the lights were turned on in the morning and began recording data again. But because they had been power-cycled, they no longer knew the correct time and so timestamped beginning with Linux start time (January 1, 1970)—until, of course, someone connected to the sensor, at which point the system took the browser time and began applying it, so that if you looked at the last several hundred data points, they would be timestamped correctly. This was a data bug that existed only when you weren't looking.

It became apparent that part of our troubleshooting of all of our alpha issues would have to be a careful analysis of the platform we had chosen. The Edison had fallen down on the processing side of things, and even with refined computer vision techniques, it was likely that we would run into other hardware issues. Meanwhile the Raspberry Pi foundation had announced the Model 3 version of its hardware in early 2016, and by fall they were finally becoming available for purchase. The Raspberry Pi Model 3 dealt with a lot of the issues that had caused us to decide against it early in our development, primarily by putting Wi-Fi onboard rather than relying on external adapters. With more processing power, more storage, and onboard Wi-Fi, the Raspberry Pi Model 3 seemed like the answer to our issues—except that we'd have to start almost from scratch in porting code from one platform to the other.

The installation was a bit overkill for the rooms, in that we could have covered the same space with fewer sensors, but we were being careful and ensuring we'd have some fallback if we found issues with hardware or software—you can never be too careful with alpha systems. Two sensors were placed in the Bill Blass catalog room, and two each in Rose Reading Room North and Rose Reading Room South (see Figure 2.3).

It was apparent quickly that there were issues with the Edison platform. Initial testing had been done in very limited traffic areas, and when the Edison attempted to keep up with the traffic in one of the busiest library rooms in the country, and during the single busiest period, the sheer volume of computation needed swamped the microcomputer and caused every type of computing issue possible. Over the course of the first few months, we saw I/O throughput errors, disk errors, and in one case the processor on one of the Edisons overheated. We worked our way through many of the issues and began digging into the data to try and do some more focused data analysis. That's when we found the most interesting bug of our alpha testing.

Perhaps obviously, the data that we were gathering depended on having accurate timestamps. The Intel Edison, however, doesn't have an onboard clock for keeping time separately from being on a network. This isn't unusual among microcomputers these days; the Raspberry Pi has the same limitation. But this meant that we needed a way to set the time on the sensors that didn't rely on them having access to the internet. Remember, these were never going to connect to the wider internet once installed; they were going to connect directly to a laptop or tablet that the librarians were using to monitor and download information. Our solution, which is the same used in another open source project I run called LibraryBox, is to scrape the time from the browser during the calibration step. When the initial connection to a laptop

Figure 2.3
Measure the Future sensors at NYPL

After evaluating options, it became obvious that moving to the Raspberry Pi–based hardware configuration was indeed the best option, and so began the development of the Measure the Future Beta program.

Beta

Through the spring and summer of 2017, we focused on moving everything to the new hardware while ensuring that we solved the problems that were identified in the alpha testing. We solved the lack of a clock for accurate timestamping by adding one physically to the Raspberry Pi. One of the advantages of the platform is that it is so popular that it has a huge variety of additional components that can be added to the base model. Adding a battery-powered real-time clock gives us confirmed timestamps for all data collected, with no concerns about power cycling or other service interruptions. By late summer, we had tested our new sensor units and confirmed that they were ready for testing in the real world.

Enter our new beta partners, the libraries at the University of Rochester in Rochester, New York; the Carnegie Library in Pittsburgh, Pennsylvania; and the Boston University Law Library. They will join NYPL, Meridian, and SUNY Potsdam as testbeds for our beta hardware, which is rolling out over the course of the fall of 2017. In addition to the updated hardware, the beta development will continue on the software side, pushing toward the launch of the cloud-based visualization and analysis tool. This new visualization tool is needed for multiple reasons, most of which boil down to user experience and system capabilities.

For libraries with multiple sensors, having all of the data in a single place and interface is clearly a better experience. In addition, we want to be able to cross-reference sensor-to-sensor data and generally have a more holistic look at building usage, rather than individual room usage, as quickly as we can. There are also visualizations and analysis of the data that we simply can't do on the sensor unit itself. The Raspberry Pi is a big step up from the Edison, but it doesn't compare in processing power to a cloud-based server where we can throw almost unlimited amounts of processing power at a particular set of data. The data we're collecting grows pretty quickly, as you can imagine. Every second we're capturing the position and timestamp for everyone in the room, all day long. Over months and months, the only reasonable way to handle that much data and deal with it all at once is to put it onto a proper server and have much more powerful processors deal with it.

With more power to throw at the data, especially longitudinal data over months and eventually years, we hope to be able to surface patterns of use that would be invisible via other data collection methods. Our beta partners will be the first to see the power of that data, and over the next six months, we will be developing the next stage for Measure the Future.

Conclusion

At the time of writing, Measure the Future has one beta site live and is running on the latest iteration of our sensor hardware, with two more location scheduled to go live in the next two weeks and another two in the following month. By the end of 2017, we should have our latest hardware in all six of our partner libraries, all of them collecting data locally. Early in 2018, we will begin moving those that wish from local data collection and visualization over to our cloud service. It's possible that not all beta sites will want to share their data remotely in any way, which is totally understandable. If they wish to implement a local instance of the Measure the Future cloud, they will be able to do that because of our open source nature. I believe, however, that our security model will be such that most libraries will choose to share their data with the project through our cloud portal.

At that point, the goal will be to look for patterns of similarity and difference between libraries. Identifying patterns across libraries is something that I believe could be incredibly useful, especially for space planning for renovations and new library buildings. Ultimately, our hope is that the data leads to libraries being able to understand how their patrons want to use their spaces, allows for iterative testing of spaces to make them ever better for their local communities, and gives libraries the information they need to tell the stories needed to ensure their continued funding.

Notes

1. John Bracken, "Knights News Challenge: Libraries Closes Sept. 30," September 10, 2014, Knight Foundation, https://www.knightfoundation.org/articles/bracken-knight-news-challenge-libraries-offers-25-million-innovative-ideas.
2. William Gibson, *Burning Chrome* (New York: Ace Books, 1982).

Raspberry Pi and Arduino Prototype

Measuring and Displaying Noise Levels to Enhance User Experience in an Academic Library

Janice Yu Chen Kung*

Problems associated with noise in academic libraries are an ongoing concern for patrons and library administration. Noise disruptions come from numerous sources including people, cell phones, sounds from eating, and audio players. Noise studies from Nigerian universities found that some of the major sources of noise were environmental factors such as automobiles, airplanes, and equipment (photocopiers, scanners, outdoor lawn mowers, air conditioners, and ceiling fans).[1] Based on the 2013 LibQUAL+ survey, noise and the ensuing lack of quiet study space continue to be challenges faced by Concordia University Libraries.[2] When I worked as a business reference librarian in Concordia University's Webster Library in collaboration with the web services librarian, we felt that the noise issue was an interesting problem to solve.

Noise Studies in Academic Libraries

Noise is a prevalent problem in academic libraries, and it is one of the major complaints from students. There have been a number of studies that collected subjective and objective data to measure noise levels.[3] It was important to capture both forms of data because objective noise data, collected through sound-monitoring devices, may not necessarily reflect how patrons perceive the noise level.[4] Loudness is subjective; what may be noisy to one person is acceptable to another.[5] Examples of collecting subjective data include the administration of questionnaires or surveys. Objectively, sound and noise are measured with a metric called the decibel. The decibel scale reads the sound pressure and translates the range of sound to a logarithmic scale.[6] The range varies from 0 to 140 decibels (dB), where 0 dB is the threshold of hearing, normal speech registers at 60 dB, and 120 dB is the noise level near a jet aircraft engine.[7] Luyben and colleagues found subjective data to be the better measurement of the two because information collected from patrons "reflected only noise that was perceived as annoying" and the electromechanical system was not discriminatory in the type of sounds generated, including dropped books,

* **Janice Yu Chen Kung** is a Public Services Librarian in the John W. Scott Health Sciences Library at the University of Alberta. At the time of the project, she was the Reference and Subject Librarian (Business) at Concordia University, Montreal.

jacket zippers, chairs bumped into tables, and other random sounds.[8]

Academic libraries explored numerous interventions to reduce noise. Noise-level zoning was one of the common strategies, including furniture rearrangement on different floors of the library. In one study, researchers removed tables and upholstered chairs from the central area and relocated them to other areas of the floor, but they found that the reallocation produced no measurable reductions in noise.[9] Crumpton discussed the benefits of reducing clusters of furniture and carefully selecting furniture such as carrels and cubicle-like walls to minimize group socialization.[10] The separation of printer and copier rooms and group study rooms from the quiet study areas focused on space allocation in containing noise levels to specific areas in the library.[11]

Other intervention strategies have also been explored, including staff monitoring by students, library staff, guards, or campus security.[12] Hronek conducted a study to determine if reducing light levels would minimize the amount of noise made by patrons when they entered the library but concluded that reducing light levels had no significant impact on noise levels.[13] Libraries enforced policies and procedures in creative ways. One method of communication involved the staff handing out cards (slips or bookmarks).[14] For instance, one message read, "Don't be Cellfish! Please set your cell phone to vibrate."[15] A number of libraries used signage, posted policies on their websites, created handouts, or used a combination of these interventions to inform patrons of their policies.[16] The key was to maintain consistent messaging for the policies to have credibility.

McGill University developed a creative noise intervention project with NoiseSign, an electronic monitoring device that measured the current noise level of particular areas in the library.[17] The researchers established an acceptable noise threshold, and when the noise threshold was reached, the LED sign would light up. They hypothesized that the sign would provide real-time feedback to inform students that they were being too loud, which would facilitate self-monitoring among the students. However, their findings showed that the intervention did not significantly change the amount of noise generated.

The common theme across all studies demonstrates that interventions tend to not produce measurable results in reducing noise. Students also responded negatively to the interventions. Some found it more difficult to complete their work, and some students were upset with the change and felt that library staff were encroaching upon their personal study areas.[18] Rather than implementing an intervention to minimize noise, we wanted to implement a solution that would inform users about the particular noise levels of different areas in the library, provide real-time and objective, quantitative feedback on noise, and allow patrons to choose which environment they prefer.

Method

Our goal is to have decibel measurement data visualized on screens to enable visitors to see the noise levels in each area of the library. This display would allow visitors to choose the area with the right amount of noise for their purposes (e.g., two students working quietly together would go to a semi-silent area; one student going to read a book would want to pick the quietest area in the library). In addition, decibel levels taken at regular intervals would be sent to a database, which could be queried in order to make informed and targeted interventions.

To implement this project, we first worked on a proof-of-concept prototype that would use sensors to measure decibel levels and quantify what is "silent" versus what is "quiet." The parts used to build the prototype included Arduino and Raspberry Pi components, a microphone sensor on the Arduino, and a computer monitor.

Implementing the Prototype

Step 1

Due to our limited knowledge of Arduinos and Raspberry Pis, we needed to have a better understanding of how they work. Reading and working through the exercises from the books *Getting Started with Arduino* and *Getting Started with Raspberry Pi* helped guide the project.[19]

Step 2

In *Getting Started with Arduino*, there is a sample exercise that teaches you how to add a light sensor to the Arduino, which is a microcontroller (small computer) dedicated to one specific purpose. We tested this out and after successful implementation of the light sensor exercise, we changed the sensor to a microphone to measure noise levels. In order to communicate the decibel measurement readings from the microphone, the sensor was connected to the Arduino on a circuit. The Arduino continually measured the decibel levels in an area by running one program on a continuous loop as long as it remained on. Using the Arduino was ideal since there was no need to build circuits from basic components, it is very affordable, and it is open source. Arduinos also come with an IDE (Integrated Development Environment) software, the suite of software that is needed with which to code the program. For example, the following code was used for the Arduino programming:

```
int DIG = 8;
   int ANA = A0;
   int sound = 0;
   int start = 0;
   int DELAY = 1000;

   voidsetup()\{
   Serial.begin(9600);
   start = millis();
   }

   void loop()\{
   // put your main code here, to run
      repeatedly:
   sound = analogRead(ANA);
   Serial.println(sound);
   delay(DELAY);
   }
```

As shown in the code above, variables must be declared first by defining some settings. There are always two parts in the programming. The *setup* defines the serial port and starts the clock. This is executed only once when the device is powered on. The *loop* retrieves the reading from the A0 connection and sends the information over the serial port, then waits one second and repeats, ad infinitum.

Step 3

The Arduino was then connected on a port to a Raspberry Pi computer, which listened to the Arduino, read the sensor data (i.e., volume from the microphone), added a timestamp, and output a data file, all using the Python programming language.

Step 4

Python communicates information to the world by using a web framework. We used Flask, a type of Python web framework, to turn the Raspberry Pi into a basic web server that sent the sensor value and timestamp to a webpage. The data file was in JSON format (though it could be XML, too), which enables several output functions, including the generation of real-time displays on screens or kiosks in the library and the website (via HTML5, jQuery, and Google Charts), and writes to text files to produce reports. Figure 3.1 outlines the schematic of the prototype.

Challenges and Lessons Learned

During the initial stages of the project, there were challenges in learning how to use command line, understanding networking and IPs, learning Python, and implementing technology-based changes in an academic institution. Discussions were needed with the university's library administration to allow our prototype project to move forward.

Calibrating the microphone was a challenge. The readings from the sensor were not in decibels, and it required many modifications to find the right calibration so that the sensor reading corresponded accurately with the decibel measurement.

Locating the proper sensor was key to the project. We wanted to obtain a reading that we could reliably translate into a decibel reading. The microphone needed to measure sound in the pitch range we were interested in and give feedback on the amplitude of the noise in the room. We started this project with a small sensor meant for Arduinos, but it wasn't sensitive enough to obtain reliable readings.

We have created the prototype and are still in the testing phase of this project. In order to continue the testing phase, several challenges need to be addressed. Finding the right microphone sensor is critical, and this could be accomplished in one of two ways: using a more sensitive microphone sensor that is compatible with the Arduino or incorporating a programmable gain amplifier into the prototype. A programmable gain amplifier allows the device to measure small voltages with increased resolution, which could increase the strength of the signal and make the microphone more sensitive in picking up noise levels. The internet connectivity will be attached to the Raspberry Pi component by using a USB wireless stick, but it is uncertain how the device will be able to connect to the university's network. Networking issues in academic environments are generally caused by the security measures in place (i.e., being locked down), and as a result, working closely with colleagues from the IT department may alleviate some of the internet connectivity challenges.

Since the prototype will be situated in a public environment, tamper-proofing the device is necessary. One potential solution is to place the pieces in the ceiling boards so that they are hidden and out of the way. The placement of the prototype will also require some preliminary testing to determine whether or not the proximity of the microphone sensor and users who are generating noise is acceptable. If the distance is too far for the microphone to detect noise, the accuracy of the decibel reading would be at risk. Collecting data points and displaying the information on the web will also require thoughtful planning and execution.

Next Steps and Future Opportunities

Since the prototype has not been tested in a library setting yet (the prototype has only been tested in the private residence of a home), there are a number of locations in the library environment that still need to be tested such as group study rooms, large study

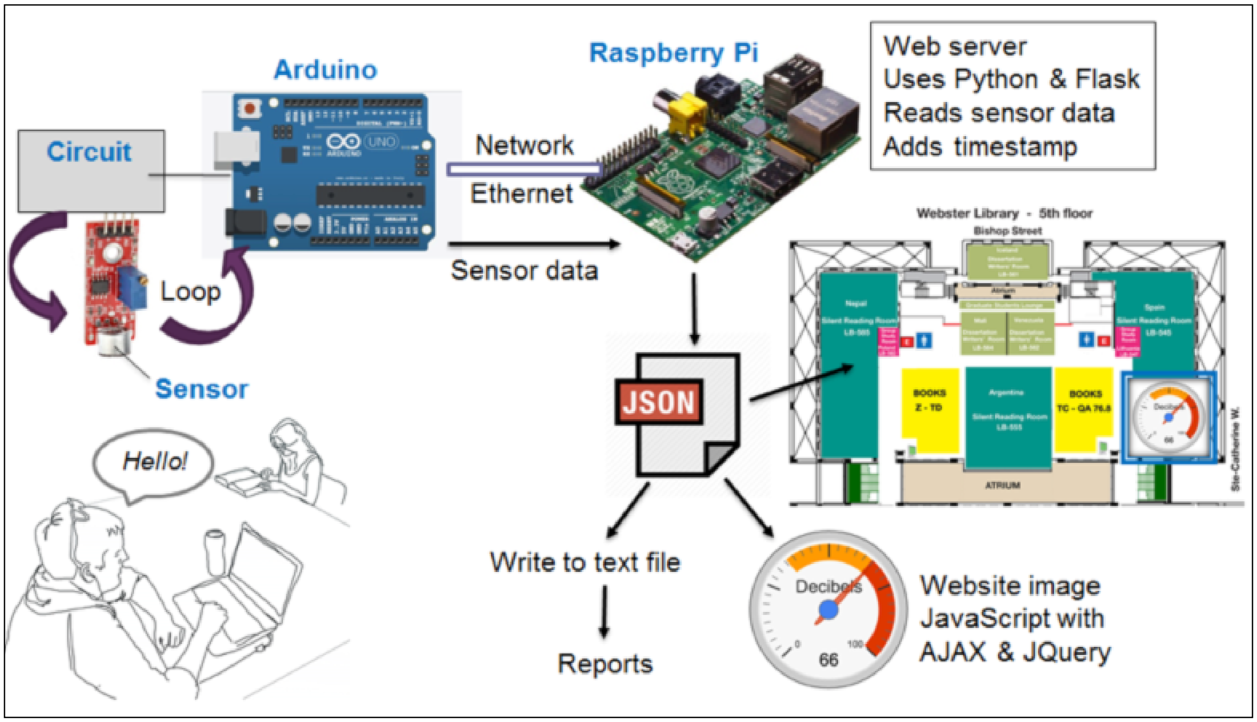

Figure 3.1
Schematic of the prototype

halls, computer labs, or collaborative spaces (e.g. makerspaces). The unpredictability of noise levels in such library spaces make them ideal sites to measure noise due to the variability in decibel readings likely to be captured at different times of the day. As in many noise intervention studies that have been done in the past, qualitative data may be helpful as a basis for comparison with quantitative data. Therefore, it would also be helpful to gather feedback from library patrons through surveys and questionnaires to determine what they perceive as the current noise levels. Having noise levels projected onto library display monitors and the library website will also require some assessment. For example, have patrons noticed the information being displayed, and what do they feel about the real-time information about noise levels? Could the information help inform their decisions on where to go in the library?

While the sensor created in this project was a prototype, it offers many possibilities for noise management in the future. This project provides a way of experimenting with "makerspace" tools such as the Raspberry Pi and Arduino to solve real-world problems for libraries. Other opportunities exist with this type of technology such as the adaptation of temperature sensors, integrating user interactivity where they may provide ratings to the real-time readings, and incorporating Raspberry Pi and Arduino with noise-cancelling technologies. Results and conclusions drawn from the pilot project will help inform library policies on space planning, library services, and enhancing the user experience.

Acknowledgement

The author would like to acknowledge Pamela Carson, Web Services Librarian from Concordia University, Montreal. This publication is a continuation of our collaboration on a paper presented at the ACCESS 2016 Conference in Fredericton, New Brunswick.

Notes

1. A. S. Aremu, J. O. Omoniyi, and T. Saka, "Indoor Noise in Academic Libraries: A Case Study of University of Ilorin Main Library, Nigeria," *African Journal of Library, Archives and Information Science* 25, no. 1 (2015): 5–14.
2. Concordia University Library, "2013 LibQUAL+ Survey Responses," About the Library, last updated November 13, 2015, http://library.concordia.ca/about/libqual/2013/responses.php?guid=space.
3. Charles P. Bird and Dawn D. Puglisi, "Noise Reduction in an Undergraduate Library," *Journal of Academic Librarianship* 10, no. 5 (1984): 272–277; Jessica Lange, Andrea Miller-Nesbitt, and Sarah Severson, "Reducing Noise in the Academic Library: The Effectiveness of Installing Noise Meters," *Library Hi Tech* 34, no. 1 (2016): 45–63; Paul D. Luyben, Leonard Cohen, Rebecca Conger, and Selby U. Gration, "Reducing Noise

in a College Library," *College and Research Libraries* 42, no. 5 (September 1981): 470–481.
4. Luyben, Conger, and Gation, "Reducing Noise in a College Library," 470–481.
5. Lange, Miller-Nesbitt, and Severson, "Reducing Noise in the Academic Library: The Effectiveness of Installing Noise Meters," 45–63.
6. Charles M. Salter, "Acoustics for Libraries," Libris Design Project, US Institute of Museum and Library Services, 2002, accessed September 2, 2016, http://www.fau.usp.br/arquivos/disciplinas/au/aut0213/2015/Acousticslibraries.pdf.
7. Salter, "Acoustics for Libraries."
8. Luyben, Conger, and Gation, "Reducing Noise in a College Library," 470–481.
9. Ibid.
10. Michael A. Crumpton, "Sounding Off about Noise," *Community and Junior College Libraries* 13, no. 4 (2005): 93–103.
11. Janet E. Franks and Darla C. Asher, "Noise Management in Twenty-First Century Libraries: Case Studies of Four U.S. Academic Institutions," *New Review of Academic Librarianship* 20, no. 3 (2014): 320–31; D. Young and H. Finlay, "Redesigning Study Spaces: Noise-Level Zoning," *Library and Information Update* 5, no. 5 (2006): 40–41.
12. Bird and Puglisi, "Noise Reduction in an Undergraduate Library," 272–277; Crumpton, "Sounding Off about Noise," 93-103; Wanda V. Dole, "The Effectiveness of Guards in Reducing Library Noise," *Library & Archival Security* 9, no. 3-4 (1990): 23-36; Al Ntui, "Noise Sources and Levels at the University of Calabar Library, Calabar, Nigeria," *African Journal of Library, Archives and Information Science* 19, no. 1 (2009): 53-54; Young and Finlay, "Redesigning Study Spaces: Noise-level Zoning," 40-41.
13. Beth Hronek, "Using Lighting Levels to Control Sound Levels in a College Library," *College and Undergraduate Libraries* 4, no. 2 (1997): 25–28.
14. Shelley Heaton and Nancy Master, "No Phone Zone: Controlling Cell Phone Use in Academic Libraries," *Public Services Quarterly* 2, no. 4 (2006): 69–80; Katie M. Lever and James E. Katz, "Cell Phones in Campus Libraries: An Analysis of Policy Responses to an Invasive Mobile Technology," *Information Processing and Management* 43, no. 4 (July 2007): 1133–1139, https://doi.org/10.1016/j.ipm.2006.07.002.
15. Heaton and Master, "No Phone Zone: Controlling Cell Phone Use in Academic Libraries," 69–80.
16. Ibid.
17. Lange, Miller-Nesbitt, and Severson, "Reducing Noise in the Academic Library: The Effectiveness of Installing Noise Meters," 45–63.
18. Luyben, Conger, and Gation, "Reducing Noise in a College Library," 470–481.
19. Massimo Banzi, *Getting Started with Arduino*, 2nd ed. (San Francisco: Maker Media, 2011).
20. Matt Richardson and Shawn Wallace, *Getting Started with Raspberry Pi* (Maker Media, Inc: USA, 2013).

Building Intelligent Infrastructures

Steps toward Designing IoT-Enabled Library Facilities

Jonathan Bradley, Patrick Tomlin, and Brian Mathews*

"Only connect."

—E. M. Forester, 1910

Silently conversing objects surround us. From smartphones to Fitbits, invisible streams of data are coursing through and between the devices we hold in our hands or wear on our bodies. Our refrigerators, sensing their contents, churn out shopping lists or place orders on the web to replenish their stock; coffeemakers and lightbulbs now connect to Bluetooth and Wi-Fi networks. On a broader scale, the connectivity enabled by the Internet of Things (IoT) has been used to build "smart cities" with improved urban infrastructures and energy-efficient buildings. Sensors, beacons, accelerometers, and actuators: these and other components are the building blocks by which we increasingly digitize, organize, and personalize the physical world. As Jacob Morgan wrote for Forbes in 2014, "The new rule for the future is going to be, 'Anything that can be connected, will be connected.'"[1] The future is here.

Consider what IoT technology means for libraries. The traditional view of libraries as islands of automation, specialized expertise, and control over access to content holds less weight in a hyperconnected world. Yet libraries remain spaces immersed in data and data collection, a fact that IoT technology has the capability to harness in new ways. Imagine a library dashboard that not only tracks gate counts and usage of physical and digital collections, but also monitors the "health" and "fitness" of the building, from the cleanliness of bathrooms to the movement of furniture in areas of the library most heavily used for study or collaboration. Or imagine walking into a library commons and receiving recommendations on your phone about locations to sit based on the similarity of the research others are conducting nearby. Imagine a whiteboard that is able to push scholarly article recommendations based on the words, phrases, or diagrams written on

* **Jonathan Bradley** is the Innovative Technologies Coordinator for Learning Environments at University Libraries, Virginia Tech, where he evaluates and implements new technologies into library services. Projects that he has been a part of include building spaces for 3-D printing, virtual/augmented reality, and data visualization in the library. He is currently producing the hardware and software for the Smart Commons project, which gathers data about how patrons use library spaces by utilizing IoT technologies. **Patrick Tomlin** is the Director of Learning Environments at University Libraries, Virginia Tech. **Brian Mathews** is Associate Dean for Learning at University Libraries, Virginia Tech. He has previously serviced as an assistant university librarian at UC Santa Barbara. Brian frequently presents and writes on topics related to innovative technologies and organizational culture.

its surface. Enabling these connections—connections between people, and between people and devices—through IoT technology can empower librarians to make strategic decisions about library spaces and services and provide library users with a unique, personalized experience.

The migration to an IoT-enhanced library is a journey of multiple steps, of course. Some of these steps are infrastructural in nature, while others will require focusing on service design and the creation and delivery of a fluid user experience. Still others will entail the development of a system of software and algorithms to collect and aggregate library data in order to analyze it across space and time. This chapter provides a brief overview of our initial steps taken in such a direction. Over the past year, we applied IoT technology to the Newman Library at Virginia Tech in an attempt to better understand our users' interactions with its spaces. By tracking the movement of furniture in the library commons, we hoped to illuminate patterns of student work, examine the density of particular work areas, and ultimately create more effectively designed learning spaces and user experiences. Using accelerometers, motion detectors, force sensors, and Bluetooth beacons, we created a system for monitoring where and when furniture and equipment were moved, what study rooms were occupied, and how students interacted with them.

The project outlined here represents the first, preliminary steps in a much larger endeavor. Nevertheless, we believe it poses important questions for the study of library spaces and services at the outset. What metrics should frame the implementation of IoT devices? How do we get not only more data, but *better* data from the library itself? Can we effectively monitor the health of a building in terms of its physical condition? Is it possible to measure and articulate the fitness of our spaces in relation to the activities transpiring therein—that is, can IoT technology provide us with a more robust picture of the difference between the intentions for our spaces and how they are (or are not) actually used by library patrons? In short, can IoT technologies help us to better understand the nature of the interactions occurring in libraries and ultimately empower us to enhance the user experience in previously unknown ways?

We entered this experiment with an exploratory mind-set. Our purpose was both practical (What are the range of sensors available and how could we deploy them effectively?) and perspective-building (What types of data could we collect and what could it reveal about patterns in our learning environment?). Through this project we uncovered three overarching design challenges: battery life, programming language, and security. This section outlines the problems and offers some lessons learned.

Design Challenge 1: Battery Life

The most prominent design challenge while building the prototype for the Smart Commons module was battery life. According to the goals of the project, we wished to deploy numerous modules to chairs around our learning commons, meaning that maintenance would inevitably be a time-consuming job, and the battery life of modules could exponentially increase that maintenance time.

The original prototype had a battery life of just under one week, meaning ten modules would have to be located on the floor, removed from the chairs, opened, disassembled, charged, reassembled, and redeployed once again every week. Since the goal of the project was to eventually scale up and add more modules not only to our commons but to other branch libraries on campus as well, this model would not result in success. We determined that for it to be sustainable, battery life for a module would need to be closer to three months.

Lessons Learned

Anyone undertaking an IoT project in an academic setting needs to devote a great deal of thought to the hardware platform they will use. Boards like the Raspberry Pi, CHIP, and others in the family utilizing ARM processors seem like a good choice. They have huge communities of support, are easy to develop on, are cheap with lots of desirable features (built-in Wi-Fi, Bluetooth, etc.), can use multiple different programming languages, and are generally easy to obtain and get code running on. But they are poor choices for IoT projects because of the battery life of the device. As full computers, these boards draw power of a magnitude far greater than most embedded intelligent devices and simply aren't sustainable for a project that will not have wired power or need to be online for more than a few hours at a time.

At the same time, even when deciding to use a common IoT chip like an ESP8266, the type of board used merits examination. Development boards, like those in Adafruit's Feather series, are great for getting your project functional, but they may not be the best for the actual deployment. Many of these boards include features like onboard LEDs and USB serial bridges that help with development but that can hurt battery performance. Many onboard LEDs can be difficult to completely disable, and LEDs are a huge battery drain, even if running only while the chip boots from sleep. Other features can draw latent power even when not actively used.

For the second version of the Smart Commons module, we have switched from using a CHIP board to an ESP32 chip. Development is happening on an Adafruit Feather board, with the goal of having a custom PCB cut for the project that contains only the minimum features needed to stretch battery life out as far as possible.

Design Challenge 2: Programming Language

Developing for IoT offers a plethora of programming language choices, and the decision about what to use for an academic project is not always clear-cut. For the Smart Commons project, we went initially with JavaScript since it is far easier to find a programmer in the library for JavaScript than for a language like C or even Python. The hope was that using JavaScript would make the project more accessible to interested developers at other institutions. JavaScript also offers in Node.js packages an abundance of tools for IoT development, and the Smart Commons project relied heavily on the Johnny-Five robotics and Bleacon packages for interfacing with sensors and hardware.

However, Node.js requires a full computer to run, which was fine for the ARM boards like Raspberry Pi and CHIP but does not translate to the small embedded chips like the ESP32. Programs using these lower-power chips are often coded in C or C++, which is a more difficult language to learn and offers more barriers to entry for an academic IoT project, in that the library has to have a C/C++ programmer to work on the project, which might not be a resource it has access to.

There are alternatives to using C/C++ on these IoT chips. By installing a different firmware, developers can enable a different language for development. Both the MongooseOS and Espruino firmwares support coding on IoT chips in JavaScript, and Micropython allows for Python coding on IoT platforms. However, these firmwares are just wrappers around the lower-level languages like C/C++, meaning that they are by their nature reactive to the underlying SDKs they are abstracting. This means that cutting-edge features often take a long time to gain support in these firmwares because they have to wait for the underlying SDK to code and stabilize a new feature, then the firmware's dev team (who are usually volunteers) have to code and test all of the wrappers before implementing.

Lessons Learned

The choice of programming language is a give and take in most situations, and the future of our Smart Commons project has been guided by this situation. The project has shifted from using JavaScript in the prototype module to using C++ as part of the Arduino IDE for the ESP32 that is at the core of version 2.0 of the project. This decision is the result of our need for cutting-edge features, namely BLE support for our chip. As of the writing of this paper, the only environment to have stable support for the BLE features of the ESP32 is the Espressif Systems (manufacturer of the ESP series of chips) SDK. However, since one of the goals is to make the code and the project as a whole as accessible as possible to other academic institutes, we will be monitoring other firmwares. We will likely be migrating the codebase to JavaScript or Python (or both) as the features we require become stable in those environments since JavaScript programmers are common and Python is currently the fastest growing programming language and easier to learn than C++.

For other groups pursuing similar IoT projects, we suggest using the most accessible language for your project that your technology needs allow. If you don't require cutting-edge features, Espruino or Micropython would likely easily meet your needs. However, if you do not plan to use open source and share your code with other groups and the project is intended for internal purposes only, then the choice boils down to the preferences of your internal development team.

Design Challenge 3: Security

IoT devices have received a great deal of attention recently because of security issues, which underscores the importance of taking extra steps to secure projects before they end up as part of a botnet that damages the health of the internet as a whole. For the Smart Commons project, we were warned by our central IT service that our campus has a large number of hackers attempting to gain access to the university's secure systems. Our IoT devices can't provide them that kind of access, but that doesn't mean that they can't be used as part of an attack or compromised to serve some other nefarious purpose. This meant we would need to make sure we took the security seriously from the outset.

In addition to software and networking security, physical hardware security was also a concern, meriting forethought to accomplish the project in a way that doesn't leave an IoT project completely open to attack. Our modules are in public places within the reach of patrons, meaning issues like theft and direct tampering also had to be considered when designing the modules. These are challenges that can be overcome, but it is our responsibility, as the stewards of data collected from our students and patrons, that we not be reactionary to attacks but instead proactive to mitigate as many risks ahead of time as possible.

Lessons Learned

In addition to the standard security practices, like changing all of the default login passwords to long, random strings, locking down unused ports, and securing API endpoints, there are a number of other practices that we put into place with the Smart Commons to head off attackers. First, we ensured that our data reporting used encryption for security. SSL certs are now free and easy to obtain, so there is no real excuse for transmitting data over insecure connections. It takes more work in the code, but it should be the default for an IoT project, even if you don't think the data you'll be transmitting is sensitive.

Second, we made the decision to have our module not act as a webserver. We knew that in addition to reporting data, the module would need to receive some information as well, mostly about alarm states and internal matters. Originally, the plan was to have each device act as a webserver and listen to information only for specific sources. However, webservers are targets for hackers and sources can be spoofed, so if at all possible, avoid having your device act as a server. We realized that with the data we needed to receive, we could just have the device check a state against our known data source during other operations instead of always listening and responding to incoming HTTP requests. While some IoT projects will inevitably require the device to act as a webserver, as you begin the project, you should consider whether or not your device could receive the needed information in some other way, like polling a trusted source periodically or grabbing the data during other operations.

In addition to software concerns, it is also important to think about physical access to the modules. It became clear to us during our design process that it would be easy for a patron to simply steal one of our modules from underneath a chair without some sort of physical security or to attempt to hack the device via physical connection to one of the ports on the chip. To mitigate these problems, we worked an alarm system into the design with button triggers that would set it off. We also designed a custom 3-D printed case for the module that would trigger the alarm if opened or if removed from the chair and would restrict access to things like ports and pins. With a bit of clever thinking, it is also possible to hide the screws that remove the case from the chair behind the case itself, making it so that one would have to open the case, thus setting off the alarm already, just to get access to the means for removing it from the chair completely. The alarm sounds only a rather quiet buzzer, just enough to let patrons know that they have done something wrong without disrupting an entire floor of students studying. More importantly, the alarm system also sends an email to the team informing them of the tampering and providing the device's last known location (gathered thanks to the BLE Beacon location monitoring). Additionally, it is advisable to purchase a board with encryptable flash space so that if someone does manage to run off with a device, that person will be unable to get access to the code and the API access or other sensitive information contained therein.

Metrics and Sensors

Due to the inexpensive nature of sensors and the wide variety already available for purchase, the metrics that can be gathered with IoT devices are nearly endless. The first prototype of the Smart Commons module tracked location-based data via Bluetooth, movement data through an accelerometer, and force data through a force-sensitive resistor. For the second iteration, the accelerometer was dropped from the design because the data gathered was deemed less useful than the location-based data being returned via the Bluetooth interface, and removing it increased battery life while reducing both cost and size. The second iteration of the Smart Commons was focused on refining the architecture, the battery life, and the size of the module, so no new sensors were added. As this module matures, the third iteration will add new functionality that we have identified as being desirable for assessing the health and fitness of the building.

Other sensors and metrics we intend to implement in the future include water leak detectors and sensors that can register humidity, temperature, air particulates, and barometric pressure. Each of these will assist in delivering information about the health of the building. From inoperable air conditioners to burst pipes, time-sensitive facilities information can be relayed to a dashboard immediately; adding additional sensors to the existing deployment of Smart Commons modules around the building is cheap and easy with tangible benefits for user experience and service design.

Additionally, we have been planning companion modules for the main Smart Commons module that can provide additional data to augment what is already gathered. These companion modules will utilize door open/close sensors, PIR (passive infrared) sensors, Velostat pressure-sensitive sheets, and thermal cameras to better track the fitness of the building. With these sensors we can better understand through anonymous data where students are at in the spaces, whether they are working together or separately, and the frequency with which they migrate to other places in the building for different task-based learning. The eventual goal would also be that these sensors could provide information on how students are using our services: Are they coming to the building to specifically use a service like one of the library's technology-oriented studios and staying to study, or do they use these services because they are already in the building doing other things?

These companion modules will likely require a different board architecture and power scheme than the main Smart Commons modules, but they will report back to the same data-gathering and dashboarding system, allowing them to augment the snapshots we get of the building's health and fitness. And with the low cost of these sensors, it is easy and cost-effective to add new ones as a new need for data gathering becomes clear. Most of these basic sensors are available for less than five dollars, and even most breakout boards carrying more complex sensors and interfacing can be had for less than fifteen dollars. By far the most expensive portion of the current Smart Commons project is the thermal camera purchased as a test prototype for tracking patrons in a space, at just over $200. A normal camera (costing around twenty-five dollars) could have been used for this, but a thermal camera was purchased to meet Virginia Tech's IT privacy guidelines for cameras in public spaces, which dictate that we should avoid having student faces captured on network-connected cameras.

IoT is still a rather fledgling technology despite the fact that many of these sensors have existed in some form or another for decades. As demand for them increases, the cost of sensors will continue to drop, and new sensors will be developed to meet emerging needs. This means the potential for data collection is huge; the question for librarians becomes less about what information could be gathered and more about the creation of purposeful, well-defined metrics and assessment strategies.

Conclusion

From checkout statistics to website analytics, libraries have long invested in data collection as a means of creating, measuring, and improving services. As more libraries have focused on assessing user experience and gauging the impact of their spaces, greater prominence has been given to user studies employing ethnographic strategies such as observations and interviews. What these approaches lack, however, are both the real-time results offered by IoT technology and the broader picture of the library it provides.

Library buildings are evolving. Now they can do much more than provide passive spaces for people to learn and work. It is when sensor-based applications and objects are aggregated to form a choreographable system that they have the potential to transform the library. True smart buildings are more than the sum of their IoT technologies—they utilize an intelligent infrastructure driven by an integrated network of systems and analytics. Similarly, building an intelligent infrastructure for libraries requires seeing them holistically, less like a container and more like a living organism in a state of constant flux and flow.

The University Libraries at Virginia Tech have started on this path. Each iteration brings us closer to realizing the potential of these sensor technologies. Since the IoT is still in an early stage, we are using each step to determine feasibility and the range of possibilities. Our goal is not only to better understand the health and fitness of our facilities and to ultimately improve services for our community, but we also aim to inspire other libraries to explore IoT and connect their buildings with ours.

Note

1. Jacob Morgan, "A Simple Explanation of 'The Internet of Things,'" Leadership/#NewTech, Forbes website, May 13, 2014, https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#30029f7d1d09.

Future Directions

Jason Griffey

There are many unknowns in the near future of technology, but two observational laws that continue to have predictive power are Moore's Law and Koomey's Law.[1] Moore's Law was coined by Gordon Moore, founder of Intel, after he observed that roughly every eighteen months the number of transistors on a silicon chip doubled, while at the same time the price for said chip was cut in half. This has the effect of doubling the computing ability and halving the price for computing power every year and a half. This means that computing power is one of the very few commercial resources that continually gets both better and less expensive over time. The companion law, Koomey's Law, operates on the same time frame, but instead of computing ability, it describes the amount of electricity needed to drive the chip in question. According to Koomey, every eighteen months the amount of energy needed to do a specific amount of computing is halved.

Humans are bad at understanding the difference in effect between linear and exponential change. To give just one fairly simple example, suppose we imagine as our baseline for computing a modern cellphone, say the iPhone 8. To buy an iPhone 8 costs $699. If we then apply Moore's Law to the phone as a whole (ignoring manufacturing costs—this is a very simple thought exercise, not a full accounting of the costs of production), we can extrapolate what the same amount of computing ability would cost in five, ten, or twenty years. To buy the same amount of computing power, complete with camera, wireless connectivity, and the like in five years will cost roughly ninety-two dollars; in ten years, twelve dollars; and in twenty years, only twenty-one cents. Yes, that's not a typo, that's twenty-one cents. And, of course, five years from that we're talking about fractional cents.

Why do we care about this change? Because the end game of the Internet of Things is that computing power and connectivity are so cheap that they are literally in every object manufactured. Literally everything will have the ability to be "smart"—every chair, every table, every book, every pencil, every piece of clothing, every disposable coffee cup. Eventually the expectation will be that objects in the world know where they are and are trackable or addressable in some way. The way we interact with objects will likely change as a result, and our understanding of things in our spaces will become far more nuanced and detailed than now.

For example, once the marginal cost of sensors drops below the average cost for human-powered shelf reading, it becomes an easy decision to sprinkle magic connectivity sensors over our books, making each of them a sensor and an agent of data collecting. Imagine, at any time, being able to query your entire collection for misshelved objects. Each book will be able to communicate with each book around it, with the Wi-Fi base stations in the building, with the shelves, and be able to know when it is out of place. Even more radical, maybe the entire concept of place falls away, because the book (or other object) will be able to tell the patron where it is, no matter where it happens to be shelved in the building. Ask for a book, and it will be able to not only tell you where it is, but it can also mesh with all the other books to lead you to it. No more "lost books" for patrons, since they will be able to look on a map and see where the book is in their house and have it reveal itself via an augmented reality overlay for their phone.

The world of data that will be available to us in ten to twenty years will be as large as we wish it to be. In fact, it may be too large for us to directly make sense of it all. My guess is that we will need to use machine learning systems to sort through the enormous mounds of data and help us understand the patterns and links between different points of data. The advantage is that if we can sort and analyze it appropriately, the data will be able to answer many, many questions about our spaces that we've not even dreamed of yet, hopefully allowing us to design better, more effective, and more useful spaces for our patrons.

At the same time, we need to be wary of falling into measurements becoming targets. I opened this report with a concept credited to economist Charles Goodhart, phrased by Mary Strathern, "When a measure becomes a target, it ceases to be a good measure."[2] We can see this over and over, not just in libraries, but in any organization. An organization will optimize around the measures that it is rewarded by, often causing negative effects in other areas. This is captured in the idea of perverse incentives, where an organization rewards the achievement of an assessment, only to realize that the achievement undermines the original goal. The classic example of this is known colloquially as the "Cobra effect," named after the probably apocryphal story of the British colonizers in India rewarding citizens for bringing in dead cobras in an attempt to control their deadly numbers in cities. Of course, the clever people of India were then incentivized to breed cobras in secret in order to maximize their profits.[3]

Libraries should be wary of the data they gather, especially as we move into the next decade or two of technological development. The combination of data being toxic to the privacy of our patrons and the risks of perverse incentives affecting decisions despite being warned by Goodhart about measures becoming targets is enough for me to caution libraries that wish to implement a data-heavy decision-making or planning process. I believe strongly in the power of data analysis to build a better future for libraries and our patrons. But if used poorly or unthoughtfully, the data we choose to collect could be our own set of cobras.

Conclusion

There is enormous potential for smart buildings to improve how libraries are viewed by their communities. There is also a huge threat presented by the addition of sensors to library spaces, in the form of destroying any semblance of privacy of the reading experience. This threat becomes larger the more that libraries outsource the collection of this environmental and usage data to outside vendors, especially those that trade in data outside of the library ecosystem. Libraries that start moving into this world need to be extremely careful to understand who controls the data about their spaces and where said data is going.

The risks for data collection aren't always obvious. One example that illustrates the challenge in threat modeling for the Internet of Things is from the Measure the Future project. By itself, the data that is collected by Measure the Future is innocuous and can't be tied to any particular patron. But if you have data about the movement of people in a space, and that space has only one person in it, then correlating that with another data source could serve to reveal the identity of the person browsing. If law enforcement shows up with a subpoena for all of the data that your library has for a particular period of time, then it is far better to not have the data for your patrons' browsing habits than it is to risk revealing their browsing behaviors. In this particular threat model, Measure the Future solves the problem by not actually recording the data in question if fewer than three people are in the frame, instead buffering the data and collapsing it into the next data bucket.

Like many technologies, the risk versus reward for smart spaces may take some time to settle out. I believe that it will settle into positive outcomes for those who choose to carefully integrate data collection into their physical surroundings, but it's equally clear that this must be done with care and thought about the risks to our patrons. It's important to think about these risks now, because as J. B. S. Haldane quipped, "I have no doubt that in reality the future will be vastly more surprising than anything I can imagine. Now my own suspicion is that the Universe is not only queerer than we suppose, but queerer than we can suppose."[4] That is certainly going to be true for technology over the next two decades.

Notes

1. Gordon E. Moore, "Cramming More Components onto Integrated Circuits," *Electronics* 38, no 8 (April 19, 1965); Jonathan Koomey, Stephen Berard, Marla Sanchez, and Henry Wong, "Implications of Historical Trends in the Electrical Efficiency of Computing," *IEEE Annals of the History of Computing* 33, no. 3 (March 29, 2010): 46–54, https://doi.org/10.1109/MAHC.2010.28.
2. Marilyn Strathern, "'Improving Ratings': Audit in the British University System," *European Review* 5, no. 3 (July 1997): 308.
3. "Cobra Effect," Wikipedia, last update October 6, 2017, https://en.wikipedia.org/wiki/Cobra_effect.
4. J.B.S. Haldane, *Possible Worlds: And Other Essays* [1927] (Chatto and Windus: London, 1932, reprint), 286.

Notes

Notes

Library Technology REPORTS

Upcoming Issues	
February/ March 54:2	**How to Stay on Top of Emerging Tech Trends** by David Lee King
April 54:3	**Privacy and Security Online** by Nicole Hennig

Subscribe

alatechsource.org/subscribe

Purchase single copies in the ALA Store

alastore.ala.org

ALA TechSource

alatechsource.org

ALA TechSource, a unit of the publishing department of the American Library Association