

# Introduction

Jason Griffey

*Although expert systems may create new roles for librarians and free them for other professional tasks, the systems will in some ways encroach upon professional domains. They encourage librarians to familiarize themselves with expert systems, current research, and applications that may affect libraries.*

—S. E. B., “The Cutting Edge,” *American Libraries*<sup>1</sup>

Since long before the invention of the digital computer, humans have dreamed of nonhuman creatures and things that could reason and solve problems. In Greek mythology, there’s Talos, a bronze statue that protected Crete from invasion and pirates, watching for and destroying anyone that came into its path.<sup>2</sup> The stories of Golem and Frankenstein’s monster illustrate humans imagining what the creating of “life” would entail and portraying nonhuman things that they think are worthy of fear and repulsion. Jonathan Swift in *Gulliver’s Travels* imagines “the Engine,” a machine that is capable of writing books on its own.<sup>3</sup> One of the early physical automatons that drew crowds was the Turk, a mechanical man that was capable of playing chess against onlookers (later to be revealed to be a hoax, with a human hiding in the mechanism and playing the game).<sup>4</sup> All of these pre-digital, nonhuman thinking objects have a few things in common: they were all presented in a fantastical way, as extraordinary and special.

The creation of digitally programmable machines, starting as early as the early 1800s with Ada Lovelace and Charles Babbage, gave rise to another type of concern, related to the fear generated by Golem and

Frankenstein’s monster, but understood and even pursued by Babbage himself. It was, after all, his efforts in describing and categorizing labor that first led him to try to create his Difference Engine.<sup>5</sup> His goal? To separate what was necessary for humans to do in a working situation and to automate the remainder. The industrial revolution had already illustrated the future of mechanical engines to replace the physical output of people, and it seemed to Babbage that his Difference Engine might well replace at least some of the intellectual output of humans, and thus replace them. The Difference Engine was limited in its abilities, doing only mathematics, but of course Babbage had plans for an Analytical Engine that would be programmable in the ways that we now understand general-purpose computers to be. While these early computers pale in comparison to the most rudimentary understanding of digital computing today, they were the first machines used to externalize what was previously an internal analytical process of humans. They also pointed toward what would become a series of ever-changing goalposts in the world of computing and artificial intelligence (AI).

Shortly after the creation of the first electronic computer in the 1940s, people began to speculate what it would mean for a computer to be “intelligent” and laying out tests that would illustrate this. They began with competitive endeavors: games—first tic-tac-toe, and then checkers. For decades this continued as the standard concept of a test of intelligence for a computing device, although along the way other games were added to the “challenge” list, and each fell in time. First chess, with the IBM computer Deep

Blue in 1996, and then eventually the game of Go was overcome in 2016, with Google's DeepMind AI and AlphaGo system defeating the best human Go player. Famously, the father of AI, Alan Turing, proposed a competition between human and machine, wherein a conversation would take place.<sup>6</sup> If the human couldn't tell the difference between communications with another human and communications with a computer, then the computer should rightly be described as intelligent. In each of these cases, the question of determining human intelligence from nonhuman intelligence is at issue, as that is the key to knowing if it is possible for nonhuman intelligence to exist.

What changes in our world when these nonhuman intelligences are no longer unique, or special, or even particularly rare? Clay Shirky once said, "Communications tools don't get socially interesting until they get technologically boring."<sup>7</sup> I think we can generalize even further and say that technology in general doesn't get socially interesting until it becomes boring. AI and machine learning are becoming so much a part of modern technological experience that often people don't realize what they are experiencing is a machine learning system. Everyone who owns a smartphone, which in 2018 is 77 percent of the US population,<sup>8</sup> has an AI system in their pocket, because both Google and Apple use AI and machine learning extensively in their mobile devices. AI is used in everything from giving driving directions to identifying objects and scenery in photographs, not to mention the systems behind each company's artificial agent systems (Google Assistant and Siri, respectively). While we are admittedly still far from strong AI, the ubiquity of weak AI, machine learning, and other new human-like decision-making systems is both deeply concerning and wonderful.

## Definitions

You may have noticed that there is quite a gap between "plays a game well" and "can have a conversation" when it comes to AI. This illustrates one of the fundamental divisions in AI research—the difference between what is sometimes called strong versus weak, or general versus applied, AI. In this section, we're going to walk through a series of rough definitions of AI.

Initially, I suppose we should define AI itself. The term *artificial intelligence* was coined in 1955 by John McCarthy.<sup>9</sup> It's used to denote any sort of intelligence that doesn't arise through natural processes, or where intelligence can be understood to be created. Human intelligence is usually used as the counterpoint to AI, although animal intelligence also comes up as a comparative in the literature. Colloquially, it refers to computer programs making decisions and judgments

that appear to be something that humans would be required for, such as recognizing objects, animals, or even individuals in photographs. Understanding and summarizing a long text passage would be another example where an AI system might perform a feat of "reasoning" that would count.

This is distinct in some ways from *machine learning*, where a specific type of AI system is capable of being trained, taught, or programmed without direct human action. A machine learning system is one where the AI is given data to consume, and that data determines the way in which the system responds. This can be one-time programming, as when a machine learning system is trained to identify certain patterns through exposure to that pattern in a large data set. It can also be iterative, where the system is designed to take its own output as a data source, checking itself and reprogramming itself as it goes. Systems can even be designed as pairs or groups, where a series of machine learning systems each learn from the other, in either cooperative or competitive ways.

The last phrase that one is likely to find in current literature about AI is *neural network*, or just *neural net*. This is a type of computer that is designed to mimic the physical structure of neurons in the human brain in its circuitry or logic. Rather than reporting decisions in simple binary on-or-off states, neural nets collectively pass along "weights" of decisions from one to the other, making best guesses as they process data, in a way that is modeled after biological processes. This makes neural nets a specific type of machine learning system, which in turn is one type of AI system.

A related concept from the history of information and library science is that of *fuzzy logic*. If you search LIS literature for early AI work, you'll find a lot of articles referencing fuzzy logic as a concept and using it to prototype research tools. Mostly these tools revolved around the same sort of tools that are currently being prototyped using newer AI techniques, services like similarity matching between subject headings and automated cataloging based on simple semantic analysis. Fuzzy logic refers to logical operations that don't have simple Boolean values of *true* or *false*, but instead have a reliability rating expressed as a value between 0 and 1. These values allow for different sorts of logical decision-making to take place, in a manner very similar to what neural nets do today.

For modern library and information science, I would recommend using *artificial intelligence* as the very broad category and sticking with *machine learning* for referring to specific systems. This is the convention that I will attempt to stick to for the remainder of this issue of *Library Technology Reports*, using AI only where I mean the concept or practice very broadly applied. In most cases what I will be referring to are machine learning systems that perform specific tasks.

## Current State of AI Technology

In the modern world, AI is everywhere. It's used in modern video games to control the actions of non-player characters, in analyzing texts to provide summaries for readers, and in determining whether or not a photo has a cat in it. Much of modern technology has, somewhere in the background, some form of AI or machine learning at work, making decisions and turning inputs into outputs. Ubiquity has made AI somewhat boring in the way Shirky posited, and cloud computing and connected devices have hidden AI systems, not obvious to users, on the edges of our computing efforts.

Let us examine two different models of using AI and machine learning to see what I mean. Both of the most popular smartphone operating systems in the world extensively use machine learning, but they do so using very different methods and architectures. Android, the operating system used by the majority of smartphones, is written by Google. Leveraging the strengths of its maker, Android's use of AI involves using the device as a sort of appendage, a sensor package that records, measures, and collects information, which is then sent upstream to servers that use billions of data points collected from tens of millions of users as input for their machine learning systems. These collected data sets are then used to produce weights for the machine learning system that analyzes photos and attempts to understand what the photos represent. Your photos are both included in Android's larger data set and analyzed against your other photos. When you ask an Android phone to show you pictures from the beach, what is actually happening in the background is an extensive set of complex data exchanges between your local phone and Google's servers, comparing your photos to the billions in its "photos" data set via its machine learning system and resulting in your phone showing you the pictures that the AI decided were most likely to be related to the concept "beach."

This methodology has several advantages and disadvantages. Since Google has billions of photos to weigh, and millions of people helping it train its AI, the decisions that the AI makes are generally very good. You can do complicated queries, such as "Show me photos from Florida on the beach with ice cream," and the AI will likely succeed in doing just that. Because the system is always iterating on itself, learning new weights as new photos are entered and described by people, new objects and events are added to the recognition engine as well. On the other hand, because it is using "public" training sets, and building its decisions based on the actions of everyone using their systems, bias and prejudice will be introduced to the system to the same degree it is present in public. There have been several examples of this

surfacing, but none more horrifying than when the Google Assistant began to label photos of Black people as "Gorillas."<sup>10</sup>

In contrast, Apple has chosen to model its AI and machine learning efforts differently. It does its analysis and weighting of your photos (as well as other data, but photos is the easiest category to explain) locally, on the devices themselves. If you have an iPad or iPhone, you can do similar sorts of searches as on an Android phone, for example, "Show me pictures of the beach." But instead of the weighting and training of the machine learning system happening on Apple's servers somewhere, it all happens locally on the devices. Apple installs models and weights from training sets that it has worked on remotely to your phone, but your data and pictures aren't a part of that data set. Your local devices use the same machine learning algorithms to include your photos in Apple's preset weights, but those aren't then pushed to Apple's servers to influence others' analysis.

This also has some advantages and disadvantages, although different ones than Google's approach. Because each data set is analyzed locally, there is no shared decision-making as there is with Google. This means that each device has to do the computing heavy lifting itself, rather than relying on remote servers for the bulk of the work. If you've ever reinstalled iOS and wondered why for the first day or so your battery life is terrible and Settings reports that Photos is using more battery than everything else combined, this would be why. When the system doesn't have a pre-existing set of search indexes for your photos, it burns battery life via the AI to create one. It also means that rather than having identical libraries across devices, each device might have slightly different indexing since it's happening entirely local to the individual machine.

The advantages of localized machine learning is seen in enormous gains in privacy and security of information. If you don't need to send photos and data back and forth from server to client, and if providers don't need to store and host data, the attack surface for the data and risk of privacy issues are hugely reduced. Continuing the example of photo libraries, Apple doesn't have access to the photos directly because of the methodology it uses to store and transmit data from your phone to its iCloud servers. According to the iOS 12 security paper, for instance, "Each file is broken into chunks and encrypted by iCloud using AES-128 and a key derived from each chunk's contents that utilizes SHA-256. The keys and the file's metadata are stored by Apple in the user's iCloud account. The encrypted chunks of the file are stored, without any user-identifying information or the keys, using both Apple and third-party storage services."<sup>11</sup>

This ensures that Apple doesn't have information that can compromise a user's privacy, even though it

might be less ideal for certain machine learning tasks. It is, I hope, obvious why this methodology difference might be of interest to libraries. As libraries and library vendors move into developing AI and machine learning systems, we should be very sensitive to the privacy implications of collecting and storing data needed to train and update those systems. As with existing systems where we outsource data collection and retention to vendors, libraries need to be very aware of the mechanisms by which that data is protected and how it may be shared with others through training sets. Where libraries can provide local analysis in the style of Apple and iOS, they should.

The above discussion describes two different methodologies for doing work using AI systems and focuses on object and image recognition in photos as the function of the machine learning system. This is only one of dozens and dozens of uses to which AI and machine learning systems are being applied in modern technology. Very broadly, one could maybe categorize uses of AI as “analysis and synthesis of media” in current tech, as so many systems are being designed to do recognition and semantic analysis work. The examples above of iOS and Android analyzing photos for objects is a common use case, and it’s easy to see that type of system being useful for libraries and archives in creating basic metadata from digitization projects. AI systems can be trained to recognize locations from a single photograph, not only in the terms of the subject of the photo, but also of where the photographer was likely standing (based on angle, geography, and more). These systems could be enormously useful in making the processing of archives and collections more quickly findable.

Similar types of systems are being developed for video, where the series of photographs that make up video are analyzed and dissected for a variety of different pieces of information, depending on the need. These can be helpful, in the case of something like HomeCourt, an iOS app that watches video of players on a basketball court and tracks position, form, shooting percentages, and more in order to help players learn from their workouts. Or they can be potentially harmful, in cases where they enable nearly real-time tracking of individuals through a store, mall, or even down city streets.

*HomeCourt*  
<https://www.homecourt.ai>

## Problems and Biases

While AI and machine learning systems will provide untold benefits to libraries, the risks and concerns that

have arisen over the last several years in regard to AI systems should give us significant pause. AI is only as good as its training data and the weighting that is given to the system as it learns to make decisions. If that data is biased, contains bad examples of decision-making, or is simply collected in such a way that it isn’t representative of the entirety of the problem set that will be asked of the system in the end, that system is going to produce broken, biased, and bad outputs. These may reflect social issues, where data could cause the AI system to be racist in its decision-making, or classist, or sexist . . . any sort of nonbalanced inputs can cause the outputs to reinforce the negative. We’ve seen this from the largest technology companies in the world, and unless we are very careful about how we implement AI in library work, we risk doing serious damage to serving our patrons.

Part of the difficulty in predicting and policing bias in AI systems is that they are often “black box” systems, where a great deal of what is being computed is inaccessible to human understanding. Neural nets, for example, are incredibly complex, with millions of interrelated weights being calculated for a given query, and with each query possibly being given different weights. They are not predictable in a precise way, so while they can be trained to operate within a given range of likely outcomes, they are simply not directly predictable in the way that classical algorithmic computing is typically understood. For a given neural net, and a given training set, and a given query, one could build a statistical model of the likelihood of outcomes, but not predict with certainty what that outcome might be.

This means that when biases are present in training data, the effects they might have on queries and outcomes may not be directly predictable. In many cases, bias can be seen only after the fact, which is far too late when dealing in data and outcomes that can affect patrons. These systems must be tested, the training data must be collected with care and understanding, and the systems themselves tuned and trained iteratively and evaluated and assessed carefully. More than ever, knowing how and what an outside vendor could be doing in the training stages is critical to understanding the system as a whole. My lack of trust that this will happen as AI systems are developed for libraries is one reason I believe libraries themselves should be working on these systems.

## Goals of Report

This report will attempt to outline some of the background of AI and machine learning systems and argue that the near future of library work will be enormously impacted and perhaps forever changed as a result of these systems becoming commonplace. It will do so

through both essays on theory and predictions of the future of these systems in libraries, and also through essays on current events and systems being developed in and by libraries right now in 2018. In these current event chapters, a variety of librarians will discuss their own projects, how they implemented AI and to what ends, and what they see as useful for the future of libraries in considering AI systems and services.

First up, chapter 2 is an essay relating the development and design of, to my knowledge, the first machine learning system developed by a library and deployed to production in a library anywhere in the US. The system is HAMLET (How about Machine Learning Enhanced Theses) by Andromeda Yelton, currently a developer at the Berkman Klein Center for Internet and Society at Harvard. At MIT, when she created and developed HAMLET, the system was a turning point in my own understanding of what machine learning might enable in libraries. HAMLET's story is a great one for illustrating what can be done with very little time and a lot of talent.

Next, in chapter 3, we have an essay by Bohyun Kim, CTO and associate professor at the University of Rhode Island Libraries, where she discusses the launch of their Artificial Intelligence Lab, which is housed in the library on campus. The idea is similar to that of a makerspace in the library, where the strength comes from the neutrality of the space. The URI Libraries are bullish on the concept of AI and student-led development. It's a fantastic model that I hope other academic libraries adopt, and that perhaps public libraries could use as a model for community AI labs.

Finally, chapter 4 is an essay from Craig Boman, Discovery Services Librarian and assistant librarian at Miami University Libraries, which looks at his attempts to use a type of machine learning to build a system to assign formal subject headings to unclassified, full-text works. His experiments highlight both positive and negative outcomes from the experiment and suggest ways forward for others who would like to test this use for AI systems.

This report will conclude in chapter 5 with a discussion of possibilities and potentials for using AI in libraries and library science. AI is so ubiquitous at this point that there is no hope of being comprehensive in either recommendations or possibilities, but I hope the chapter is illustrative enough to point at the next five to ten years of development in the field and try and see where we are likely to most be benefited and

harmed by the explosion of this technology. I hope that this issue of *Library Technology Reports* precedes a significant expansion of efforts in this space by libraries in the same way that previous reports that I have written (on mobile technology, 3-D printing, and makerspaces) did. AI and machine learning systems have the potential to change basic functions within libraries, from cataloging to search to interfaces with patrons. And, like all emerging technologies, if we don't understand it, don't experiment with it, and don't build some of our own tools, we will be beholden to the commercial entities that trade our failures for our money.

## Notes

1. S. E. B., "The Cutting Edge," *American Libraries* 14, no. 11 (December 1983): 730, JSTOR, <https://www.jstor.org/stable/25626544>.
2. "Talos—Crete," Ancient Origins, February 16, 2013, <https://www.ancient-origins.net/myths-legends/talos-crete-00157>.
3. Jonathan Swift, *Gulliver's Travels* (New York: Signet Classic, 1983).
4. Ella Morton, "Object of Intrigue: The Turk, a Mechanical Chess Player That Unsettled the World," *Atlas Obscura*, August 18, 2015, <https://www.atlasobscura.com/articles/object-of-intrigue-the-turk>.
5. For more details, see "A Brief History," The Babbage Engine, Computer History Museum, accessed October 10, 2018, [www.computerhistory.org/babbage/history](http://www.computerhistory.org/babbage/history).
6. Noel Sharkey, "Alan Turing: The Experiment That Shaped Artificial Intelligence," *Technology News*, BBC, June 21, 2012, <https://www.bbc.com/news/technology-18475646>.
7. Clay Shirky, *Here Comes Everybody: The Power of Organizing without Organizations* (New York: Penguin Press, 2008), 105.
8. "Mobile Fact Sheet," Pew Research Center, February 5, 2018, [www.pewinternet.org/fact-sheet/mobile](http://www.pewinternet.org/fact-sheet/mobile).
9. Elaine Woo, "John McCarthy Dies at 84; The Father of Artificial Intelligence," *Los Angeles Times*, October 27, 2011, [www.latimes.com/local/obituaries/la-me-john-mccarthy-20111027-story.html](http://www.latimes.com/local/obituaries/la-me-john-mccarthy-20111027-story.html).
10. Jackie Snow, "Google Photos Still Has a Problem with Gorillas," *The Download, MIT Technology Review*, January 11, 2018, <https://www.technologyreview.com/the-download/609959/google-photos-still-has-a-problem-with-gorillas>.
11. Apple, *iOS Security: iOS 12* (Cupertino, CA: Apple, September 2018), 63, [https://www.apple.com/business/site/docs/iOS\\_Security\\_Guide.pdf](https://www.apple.com/business/site/docs/iOS_Security_Guide.pdf).