Collaborative Librarianship

Volume 10 | Issue 2 Article 3

9-12-2018

The Baked-in Bias of Algorithms

Lori Bowen Ayre
Galecia Group, lori.ayre@galecia.com

Jim Craner
Galecia Group, jim@galecia.com

The Baked-in Bias of Algorithms

Abstract

Algorithms are created by and used by humans in software programs and in everyday tasks. They are composed of input data, a series of steps, and output. When it comes to computer algorithms, we often see the results of algorithms but we don't see the processing steps or the input data that has determined the output. It is important to be aware that all these components are subject to mistakes and biases - the input data as well as the processing steps.For this reason, we should seek transparency in the algorithms that are put to use and which affect our lives every day.

Keywords

algorithm, bias, transparency, inadvertent injustice

Technology Matters

The Baked-In Bias of Algorithms

Lori Bowen Ayre (lori.ayre@galecia.com)
The Galecia Group

Jim Craner (jim@galecia.com)
Collection Development & Operations Manager and Drupal Developer, The Galecia Group

Abstract

Algorithms are created by and used by humans in software programs and in everyday tasks. They are composed of input data, a series of steps, and output. When it comes to computer algorithms, we often see the results of algorithms but we don't see the processing steps or the input data that has determined the output. It is important to be aware that all these components are subject to mistakes and biases - the input data as well as the processing steps. For this reason, we should seek transparency in the algorithms that are put to use and which affect our lives every day.

Keywords: algorithm, bias, transparency, inadvertent injustice

One of the latest buzzwords is "algorithm," usually used in a technical context. Simply defined, an algorithm is just a sequence of steps. The Simple English Wikipedia cites a recipe as a good example of an algorithm: starting with the ingredients (input), certain steps are performed in a certain order (algorithm), resulting in a complete dish (output). Computers use algorithms in the form of software programs that define those steps and process input data, resulting in output data. But algorithms are in use in every aspect of our lives, not just in our computers and recipes.

Film recommendations offered by Netflix and book recommendations from your library's ILS are based on algorithms using the viewing patterns of millions of people (Netflix) or your patrons (your ILS). Bibliocommons, a popular discovery layer product, uses the circulation activities of all Bibliocommons users in its "recommendations" algorithm. And as a librarian, you use algorithms every time you alphabetize a book or shelve Harry Potter DVDs in the proper sequel order.

Another algorithm of which you may be aware is *Library Journal*'s (*LJ*) Index of Public Library Service. This algorithm involves analyzing a number of service metrics, population figures, and budget expenditures to yield a numerical "score" so that libraries can compare themselves to other libraries in their budget peer group. The various inputs – library visits, circulation and e-circulation, public computer usage, program attendance, and service population – are fed into a series of calculations yielding a single numeric score.[1]

But the *LJ* editorial staff goes to great pains in their FAQ to note that they do **not** "measure the quality, excellence, effectiveness, value, or appropriateness of library services." To do so would involve weighting each output, assigning a value to certain qualities, ultimately favoring a certain type of library over others. In their

words, they are trying **not** to endorse a certain strategic objective, such as "*'library as place' versus remote library use versus community outreach and engagement.*"[2] In fact, they are intentionally striving to ensure a particular neutrality and they are very clear and transparent about how they gather and synthesize their data, and the possible negative aspects of their decisions.

Internet content filters provide another great example of algorithms with positive and negative effects. Filters can be effective at blocking out undesirable content but they are not very accurate. The worst example is that of an overzealous automated censor that blocks web pages based on keywords. These filters can end up blocking legitimate educational information, such as breast cancer treatment resources, due to the appearance of the word "breast."[3]

The problem is that algorithms are not usually as transparent as the *LJ*'s Index or content filters that block based on keywords. The inputs and processing steps are often proprietary. It is often the case that all we know about the algorithm is the output.

WIRED reports that U.S. states are using algorithmic computer systems – developed, controlled, and kept secret by corporate developers – to determine criminal sentences and parole lengths but nobody knows how these computer systems determine someone to be "high risk."[4]

Websites now offer automated mortgage loan decisions. We assume those decisions are based on logic and fair data but what if some of that data is rooted in the old real estate concept of "redlining," or discriminating against certain races of people?

There are two ways that algorithms can go off the rails. One is that the data being used can be biased or inaccurate. Another is that the program itself can be biased. Some of these programmatic biases can be built into software purposefully, while at other times biases find their way into the algorithms accidentally – a form of "inadvertent injustice," so to speak.[5]

Consider the "source data" that goes into an algorithmic system to calculate prison and parole sentences as mentioned in the criminal justice example above. This data consists of a number of factors, some of which are based on the individual defendant's traits, and some of which are based on historical data. As Nick Thieme writes in an article about "computational injustice": "*AI's unique talent for finding patterns has only perpetuated our legal system's history of discrimination, for example. Since people of color are more likely to be stopped by police, more likely to be convicted by juries, and more likely to receive long sentences from human judges, the shared features identified are often race or proxies for race. Here, computational injustice codifies social injustice.*"[6] In other words, social bias and algorithmic bias can reinforce each other in a feedback loop – a vicious circle of injustice.

Speaking of vicious feedback loops, consider algorithm-driven portals like Google and Facebook. Both companies use an algorithm that presents new content based on previous choices. In other words, if a liberal person clicks on a news link from a presumed liberal source, shared by a liberal friend, then Facebook will be more likely to present more of that "liberal" content in the future. And vice versa. This leads to what is called "the filter bubble" effect, where people are put into a "silo" with little exposure to contrary opinions (or facts). As in the examples above, this can create an indefinite feedback loop. Given that over half of millennials and almost half of baby boomers get their news from Facebook, that is one vicious and far-reaching feedback loop.[7]

No matter how far removed computer algorithms seem from everyday life, they are not just a trusty series of 1s and 0s. The algorithms, the choice of data to use, how it is processed, the rules that are applied – these are all created by

people, with their respective history and biases and values. To rephrase Charlton Heston's memorable line from the 1973 cult classic *Soylent Green*, "Algorithms are people!"

As humans, we all have implicit biases. As we build these new systems – facial recognition, artificial intelligence, analytical algorithms – we're creating them in our own image, with these biases baked in. It is critical that we examine the logic *and* the humans behind them rather than trusting that "the computer must be right."

---

[1] *Library Journal (LJ)*, "The LJ Index: Score Calculation Algorithm," accessed 06 June 2018. https://lj.libraryjournal.com/americas-star-libraries-score-calculation-algorithm, archive available: http://web.archive.org/web/20171018025809/https://lj.libraryjournal.com/americas-star-libraries-score-calculation-algorithm.

[2] *Library Journal (LJ)*, "The LJ Index: Frequently Asked Questions," accessed 06 June 2018. https://lj.libraryjournal.com/stars-faq/, archive available: http://web.archive.org/web/20180102185241/http://lj.libraryjournal.com:80/stars-faq/.

[3] Organization Staff, "Internet Filters," National Coalition Against Censorship, accessed 20 July 2018, http://ncac.org/resource/internet-filters-2.

[4] Jason Tashea, "Courts Are Using AI To Sentence Criminals. That Must Stop Now," *WIRED*, published 17 April 2017, accessed 20 July 2018, https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/.

[5] Li Zhou, "Is Your Software Racist?," *Politico*, published 07 February 2018, accessed 20 July 2018, https://www.politico.com/agenda/story/2018/02/07/algorithmic-bias-software-recommendations-000631.

[6] Nick Thieme, "We Are Hard-Coding Injustices for Generations to Come," *Undark*, published 20 February 2018, accessed 20 July 2018, https://undark.org/article/ai-watchdog-computational-justice/.

[7] Amy Mitchell, Jeffrey Gottfried, and Katerina Eva Matsa, "Facebook Top Source for Political News Among Millennials," *Pew Research Center: Journalism & Media*, published 01 June 2015, accessed 20 July 2018, http://www.journalism.org/2015/06/01/facebook-top-source-for-political-news-among-millennials.