A Peek at Proprietary Algorithms

Software increasingly determines what people see online or even how long they might spend in jail, but few can access how such programs work.

Let's say you start using a music streaming service online. You pick a few songs you like, and the site creates a playlist of similar music for you based on that input, using its huge database of information classifying songs into different genres, along with data about other users' likes and dislikes. Each time the service plays a song, you rate it, and the service refines its statistical model of your preferences. This is *machine learning* at work, and it's one of computing's fundamental tools for helping people make choices in the face of information overload, no matter whether the task involves choosing songs in a catalog, classifying the overflow of posts and photographs on social media, filtering wanted email from spam, or even helping businesses sort the flood of résumés received in response to an advertisement.

To complete such tasks, software follows *algorithms,* which have been part of computing since the dawn of computer science: Ada Lovelace wrote the first one back in 1840. More recently, however, the term *algorithm* has evolved to mean "a self-contained step-by-step set of operations that computers and other 'smart' devices carry out to perform calculation, data processing, and automated reasoning tasks," according to the public policy council for the Association for Computing Machinery (ACM), the world's largest educational and scientific computing society.

Principles for Algorithmic Transparency and Accountability

1. **Awareness:** Owners, designers, builders, users, and other stakeholders of analytic systems should be aware of the possible biases involved in their design, implementation, and use and the potential harm that biases can cause to individuals and society.
2. **Access and Redress:** Regulators should encourage the adoption of mechanisms that enable questioning and redress for individuals and groups that are adversely affected by algorithmically informed decisions.
3. **Accountability:** Institutions should be held responsible for decisions made by the algorithms that they use, even if it is not feasible to explain in detail how the algorithms produce their results.
4. **Explanation:** Systems and institutions that use algorithmic decision-making are encouraged to produce explanations regarding both the procedures followed by the algorithm and the specific decisions that are made. This is particularly important in public policy contexts.
5. **Data Provenance:** A description of the way in which the training data was collected should be maintained by the builders of the algorithms, accompanied by an exploration of the potential biases induced by the human or algorithmic data-gathering process. Public scrutiny of the data provides maximum opportunity for corrections. However, concerns over privacy, protecting trade secrets, or revelation of analytics that might allow malicious actors to game the system can justify restricting access to qualified and authorized individuals.
6. **Auditability:** Models, algorithms, data, and decisions should be recorded so that they can be audited in cases where harm is suspected.
7. **Validation and Testing:** Institutions should use rigorous methods to validate their models and document those methods and results. In particular, they should routinely perform tests to assess and determine whether the model generates discriminatory harm. Institutions are encouraged to make the results of such tests public.

Source: https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf

Computer programs running advanced machine-learning algorithms are now everywhere, and are making decisions about people for which we often have little or no recourse. Go shopping online, and algorithms decide which products you see and what special deals you might be offered—and which are withheld. Apply to a university, and an algorithm may determine if a human even sees your admissions essays. In New York City and Santa Cruz, California, among other places, predictive policing algorithms may determine how often the police patrol a neighborhood. And if you commit a crime in Wisconsin, an algorithm might determine how much time you spend in prison.

These uses of algorithms all share several important characteristics. Like the music service scenario, they all appear to be based on advanced statistical machine-learning techniques, developed over the past two decades, which use large amounts of training data to create models that can then make predictions or perform classifications. In addition, these algorithms are proprietary: So far, the public (or regulatory officials acting on the public's behalf) has no right to inspect algorithmic implementations or the training data, even when the algorithms are used for public purposes. And finally, in many cases, the algorithm is being used by an organization with institutional power, to make decisions about people who frequently have no right to appeal if the algorithm makes a mistake.

On September 14, the ACM's U.S. Public Policy Council (USACM) held a panel discussion at the National Press Club in Washington, D.C., to discuss the effect of algorithmic decision-making on society and the technical underpinnings of algorithmic models. I moderated the panel, and I cochair the USACM's working group on Algorithmic Transparency and Accountability (ATA).

Jeanna Matthews, an associate professor of computer science at Clarkson University in New York and also a cochair of the USACM group, opened the panel's discussion by introducing the "Principles for Algorithmic Transparency and Accountability" (*see box*) that USACM issued earlier this year. Modeled on the various principles of fair information practice developed by privacy regulators worldwide over the past 40 years, these principles are designed to provide a framework for thinking about the challenges posed by the increasing use of algorithms in our society. The seven principles address many of the concerns that have been voiced about the growing use of algorithms, without placing limits on the possible beneficial uses of algorithms currently being explored.

Nicholas Diakopoulos, an assistant professor at Northwestern University, spoke about his website, Algorithm Tips (http://algorithmtips.org/), which tracks the growing use of algorithms by the U.S. government. The website aims to promote accountability and transparency, and specifically to assist investigative journalists in writing about these topics. The database contained more than 150 algorithms in September.

Dan Rubins, one of the two industry representatives on the panel, is cofounder of Legal Robot, an artificial intelligence startup that is using algorithms to analyze case law and contracts. Rubins said that his company is using the ATA principles as the framework for the transparency report that Legal Robot is publishing on its website. It is possible for companies to be open about how their technology works without giving up commercial advantage, Rubins said, but doing so requires focusing on where the company's added value happens to be. For example, he said, it can be difficult to move algorithms from one domain to another, because the training data may not be representative. One of Legal Robot's advantages, he says, is the way it has collected and labeled its training data, instead of the actual algorithms. For example, many programs that were developed to process English fail when presented with legal documents, because the vocabulary and style of language usage is so different.

Algorithmic Bias

Geoff Cohen, a vice president at Stroz Friedberg, an Aon company, has performed many computer forensic investigations of algorithms used by companies. He said that in recent years, U.S. patent law has weakened while legislation regarding trade secrets has become more powerful, and many companies have reacted by trying to keep their algorithms more secret. But inspection is possible, Cohen said, adding that inspection typically happens when an algorithmic-based company is being considered for acquisition or is the subject of a lawsuit. In the future, government regulators might instead perform such inspections.

Finally, Ansgar Koene, chair of a Standard for Algorithm Bias Considerations working group under the auspices of the Institute for Electrical and Electronics Engineers (IEEE), discussed the efforts under way to create an international standard that organizations could use to understand and eliminate unintentional algorithmic bias in their offerings. For instance, there has been increased attention to cases in which algorithms have developed "racist" classifying tendencies, some of which could be tied to underlying data sets that contained unknown biases. In other cases, algorithms have been shown to have unintended consequences, such as causing users to see on social media only news that conforms to their own political leanings, or repeatedly sending police to the same neighborhood because there were no reports of crime in others.

There is certainly a growing interest among computer scientists and academics in developing approaches for algorithmic transparency and accountability. Since 2014, for example, a workshop called Fairness, Accuracy, and Transparency in Machine Learning has explored this topic with significant mathematical rigor. Increasingly, academics are reaching out to corporations and policy makers, with the hope of establishing norms so society can benefit from algorithms while preventing individuals from being inappropriately harmed by them.

With the rise of the information economy, it would be unworkable to go back to a world in which nothing is filtered. Without algorithmic assistance, social media, Internet search engines, and even email would all be unusable. Companies are now exploring ways of using algorithms to filter not just spam, but also fake news and even hate speech. Just how far such filtering goes, and whether it is possible to turn it off, needs to be the subject of public debate. —*Simson L. Garfinkel*

Simson L. Garfinkel is an adjunct faculty member at George Mason University, where he teach digital forensics. He is the author of Database Nation: The Death of Privacy in the 21st Century *(O'Reilly, 2000). Internet: https://simson.net/*