The Case for Algorithmic Scepticism in Law

Abstract: While we are aware that algorithms are at play in the background of much of what we do in our day-to-day life they are still widely misunderstood, says **Sarah A Sutherland**, the founder and principal consultant at Parallax Information Consulting. Here Sarah talks us through how we should approach our technology adoption by considering what algorithms are being used for and what the underlying technology and data that supports them actually is.

Keywords: informatics; technology; data; ethics

Algorithms have become ubiquitous in our society, yet they are widely misunderstood. Many of these misunderstandings arise from widespread lack of understanding of the technical basis for what algorithms are and how they function, but even experts often don't understand how they work, only that they do in many situations. This lack of understanding means that there are both rational and irrational calls for caution as they are adopted further. To balance the benefits from technology adoption and caution, we need to approach these issues carefully and consider what algorithms are being used for and what underlying technology and data supports them.

Algorithms' functionalities used to be more defined, but they are frequently not hard coded anymore. Instead, they are designed to find patterns in underlying data using free association and then follow them in new situations.
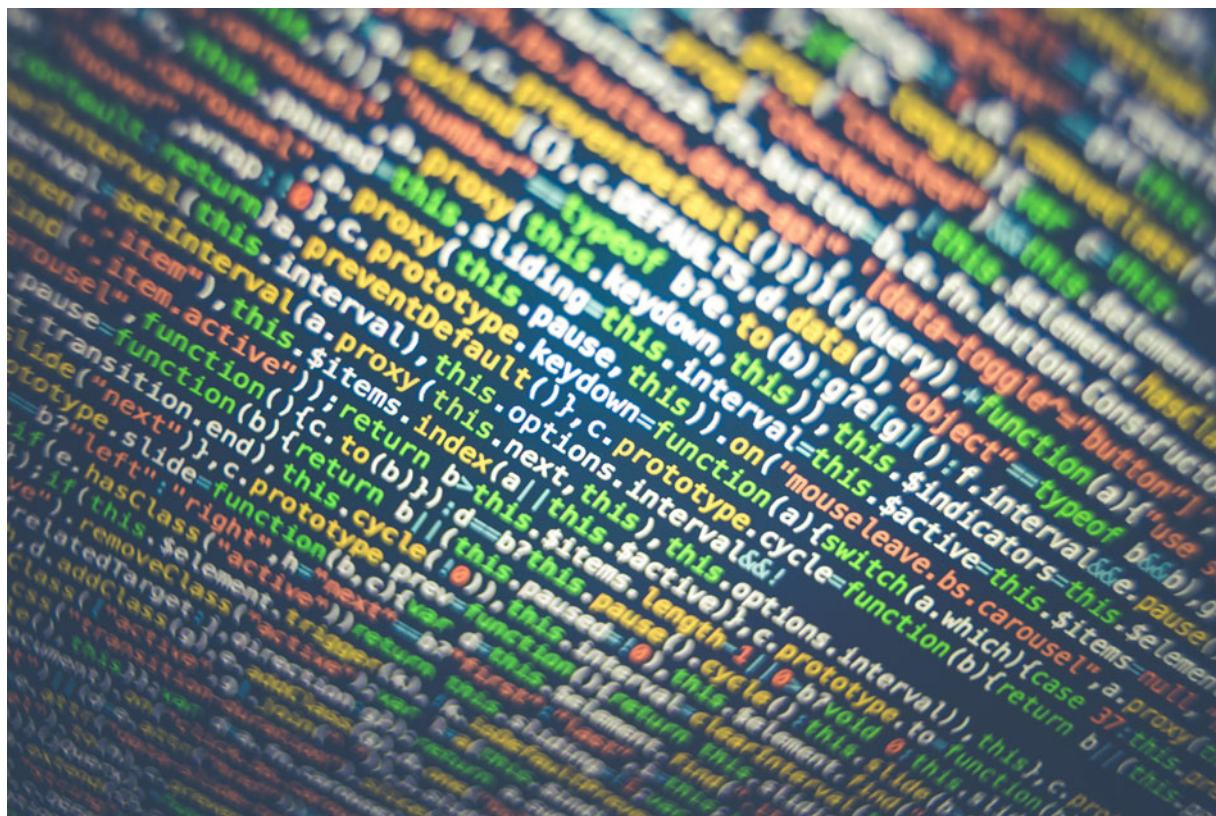
We are generally aware what algorithms are, but do we really understand what they do and how they do it?

They don't have clear definitions for how they operate, so the main way to assess whether their results are appropriate is to assess them against fresh data and see whether the patterns are useful outside the initial training set.

There are many examples of artificial intelligence systems that use data elements which we think are irrelevant. This raises the question about what kinds of applications in a legal environment are acceptable when this kind of uncertainty about why a decision was made may ultimately be unavoidable. Even where systems are designed to give reasons for their outputs, they often express the reasons in forms such as per cent weightings assigned to different elements based on what the system interprets to be most closely correlated for the appropriate outcome.

Whether a particular policy intervention or attribute of a party in a court matter is correlated with certain outcomes is easy to identify. However, we are generally not able to identify causation, which requires knowing what would happen in one set of circumstances as opposed to another. In the legal sector, generally, we only know what happened instead of all possible outcomes. We don't have evidence for counterfactuals which would allow us to identify causation. Resolving this would likely require the adoption of experimental methodologies in the legal space, and it is unlikely that this will be resolved in the near term.

However, it is possible that future developments in data infrastructure such as law as code will be used in different types of algorithmic systems which will be more transparent.

Identifying causation may seem abstract, but as algorithmic systems become more integrated with the legal system and potentially more tied to outcomes in the world, it will be important to consider how much we are willing to accept results that make recommendations based on criteria we think inappropriate or irrelevant.

Generally, these concerns are less pressing in applications like search or writing assists as they are presenting results to users in a way that's designed to be convenient while still allowing users to manually sort and revise to ensure that the work is appropriately carried out. This requires a certain sophistication on the part of users, but generally people are comfortable disregarding irrelevant prompts. In contrast, systems that make recommendations about how particular decisions should be made are more high stakes, and their user experience design may encourage more reliance on them than is appropriate.

Algorithms are often the focus of development work, but the underlying data is equally or more important for quality of outputs. Part of the reason for this is that developers have more control over the algorithms they work with than what data is available.

In contrast, many of the algorithms commonly used in legal applications based on machine learning are open source and widely available. With sufficient expertise, technical infrastructure, and data infrastructure, they can be widely adopted – it's the implementation and set-up that's special and unique in each system. And the recent developments driven by large language models are built on systems than can be accessed for a fee, though the expense to do so may be substantial.

One of the major challenges for data use in law is the structure of the law itself. Generally, legal documents are not structured in a way that is conducive to use as data for several reasons:

- They are written in complex language

- Individual documents contain information pertaining to different areas of law

- They change and whole bodies of law become irrelevant, and

- Caselaw doesn't follow mathematical requirements for statistical analysis

Part of the reason for these issues is that the written law was mostly designed as a record of good process in governance rather than a way to communicate content. Leaving these essential issues aside, another issue continues to be pure availability of the law for analysis as:

- It may not be in usable formats, and

- Issuing bodies may not make it available

Many developers see case law as something that should be available as open data, but generally government rules for open data exclude anything with personal or business information. The law is so atypical as a source of data that it's been difficult to know how to handle it in many jurisdictions.

Part of the promise of machine learning is that systems will constantly be updated, unlike older data driven systems that required extensive retooling. This ease of integrating live feedback can lead to issues like Microsoft's Twitter Bot that became racist due to the training data it received in 2016.[1] This means that systems need to be retested and audited each time they are updated rather than updated with live data.

Very small changes in underlying data have large effects on outcomes, and there really is no way to easily manage them. This means that systems will likely always have significant differences between them. These changes also mean that results will not be static, instead they will be regularly updated. It can be hoped that this is done in a considered way with appropriate review, but the systems are so complex that there's not always a good way to assess accuracy aside from comparing predicted outcomes against targets. Often the targets are what would be expected to happen in a human based system, using techniques like audits and statistical review of results to assess whether they are appropriate, which allows for considerable problems with existing biases being integrated into new computational systems.

When deciding how to use algorithms and the systems they are based on, it is important to remain sceptical and think critically about the consequences of potential outcomes. Generally, I am less concerned about uses that provide results that are fully reviewed by sophisticated users than those that provide a gloss of quantitative methodology and assurance that is not warranted by the underlying system.

Endnote

[1] James Vincent, "Twitter Taught Microsoft's AI Chatbot to Be a Racist Asshole in Less than a Day," The Verge, March 24, 2016, <www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>

Biography

Sarah A Sutherland is the founder and principal consultant at Parallax Information Consulting where she focuses on legal data strategy. Her book *Legal Data and Information in Practice: How Data and the Law Interact* was published by Routledge in 2022. She was named one of the Fastcase 50 innovators, visionaries and leaders in legal technology in 2022 and received the Denis Marshall Award for Excellence in Law Librarianship from the Canadian Association of Law Libraries in 2023.