Open Algorithms: Moving Away from ‘Magic 8 Ball” Governance

Word Count: 1,883

Government increasingly uses algorithms that affect citizens' life outcomes. However, these algorithms often do not offer transparency into their operations and conclusions and can be systematically incorrect in ways detrimental to populations. To mitigate these challenges and to bolster innovation, state actors should require open access to all data and processes related to predictive algorithms used by government agents.

As technology advances, many state actors have explored the greater capabilities of predictive modeling and risk assessments algorithms (hereafter predictive algorithms) to assist in providing government services. Predictive algorithms, loosely defined, use data sets to predict the likelihood of future events. Oftentimes predictive algorithms couple statistical analysis with machine learning techniques and advanced computing power.

Predictive algorithms can offer a range of benefits. Their use may lower costs for agencies putting them to use, allowing the agencies to perform tasks more quickly or efficiently than if they were assigned to human operators or analysts. But predictive algorithms are also used because they are believed to reach conclusions more accurately than humans. By shifting through complex patterns in large data sets, predictive algorithms may uncover correlations or trends that humans never could find at scale. Proponents also suggest that through dispassionate analysis, predictive algorithms can limit human biases. While humans allocating rewards between citizens may be affected by unconscious racial bias or lookism, a machine might be able to discount these in favor of the most fair or efficient allocation of resources.

However, there are also risks in allowing predictive algorithms to determine outcomes of human populations. If inaccurate, the algorithms can inappropriately penalize or damage whole sectors of the population. This is likely if the data the algorithms operates on is itself biased or inaccurate. In one analysis, ProPublica suggested that algorithms used to predict the risk of

future crime is biased against certain racial groups.[1] The scope for harms to citizens is great since predictive algorithms have been used by various state and local actors for a variety of services across America, from assessing the likelihood of recidivism for criminal offenders to adjusting frequency of police patrols. Looking internationally demonstrates that there is a much wider universe of issues predictive algorithms could eventually be applied to. In the Netherlands, predictive algorithms have been used to determine which individuals are most likely to falsely claim unemployment benefits.[2]

What compounds the danger of such inaccuracy is that while some of these algorithms are created by IT experts or innovation hubs within government, many are produced by non-government actor. Rather than construct algorithms, in-house, government actors allow companies to bid for contracts to develop or utilize predictive algorithms to manage processes. This product formation is damaging in several ways. First, companies have incentives to minimize any problems with their algorithms. A finding of discrimination against women in an algorithm, for example, could cause an agency to terminate, or not renew, a contract for use of an algorithm. Second, companies supplying algorithms frequently lack transparency in clarifying which datasets and processes make up their algorithms. Companies often claim this is "proprietary information," and that the details of their algorithms must be kept private to protect their intellectual property and for fair competition. The lack of transparency means it is more difficult for other parties to determine whether there are flaws or biases in the predictive algorithms.

---

[1] Angwin, Julia, Larson, Jeff, Mattu, Surya and Lauren Kirchner. "Machine Bias." *ProPublica.* May 23, 2016.
[2] Metz, Cade and Adam Satariano. "An Algorithm That Grants Freedom or Takes It Away." *The New York Times.* February 6, 2020.

These misaligned incentives are not unknown to the public. Some have suggested purely market-based solutions to issues related to proprietary algorithms.[3] And government actors have begun to reckon with this too: a 2019 Idaho bill required that "all pretrial risk assessment algorithms shall be transparent, and all documents, records, and information used to build or validate the risk assessment shall be open to public inspection, auditing, and testing. No builder or user of a pretrial risk assessment algorithm may assert trade secret or other protections in order to quash discovery in a criminal matter." [4]

The proposal outlined below builds on Idaho's legislation, seeking to extend the benefits legislators envision from their limited bill to predictive algorithms of all kinds used across the United States. Under the proposed reform, actors at all levels of government should require predictive algorithms to be maximally transparency in publishing the data and processes through which they operate. Concerns over proprietary information or intellectual property would not be grounds to avoid publication of all details of algorithm operation.

There may be narrow grounds under which a government agency could allow for a right to obscure data or processes by which a predictive algorithm is created. This would be limited to algorithms involving state security or military operations that, if published, might endanger lives. Another possible use of partial obscurity is if a dataset must involve individuals' identities, government could allow for obstruction to keep people from being individually traceable or identifiable. However, these exceptions should be few and far between and incentives should exist to limit their use (such as the involved government actor petitioning a court or appropriate government body for approval).

---

[3] Garfinkel, Simon. "A Peek at Proprietary Algorithms." American Scientist, vol. 105, no. 6. November 2017.
[4] Judiciary, Rules and Administration Committee. "House Bill No. 118, Legislature of the State of Idaho." 2019. https://legislature.idaho.gov/sessioninfo/billbookmark/?yr=2019&bn=H0118.

Logistically, legislation on both the federal and state level mandating agencies institute this requirement is the best way to enact this policy. While perhaps government actors involving procurement might play a role, limitations to their capabilities make it unlikely they could be a primary motivator of this policy. The OMB's Office of Federal Procurement Policy sets broad procurement policy for the federal government and the GSA manages the Federal Acquisition Regulation (FAR) which governs acquisition by executive branch agencies. However, many believe that state generally have their own power in setting procurement policy. Federal funds in some instances can entice states to follow federal guidelines but making state funding conditional on this policy might face legal or constitutional challenges. As such, statutory authorization through appropriate legislation is likely preferable. While it is difficult to imagine all state governments institution this policy concurrently, it seems plausible that were this enacted as a statute on a federal basis and in a handful of states it might influence others to mimic the legislation.

This policy is an improvement to the status quo described above. More examples of and information about predictive algorithms will help create faster and better algorithm development in the future. This might be most true proximately: looking at examples of algorithms involving criminal justice will be most helpful for developing new algorithms in that sphere. But widespread publication of algorithmic processes can also serve as a public knowledge base that helps development in many areas. The cross-pollination of ideas allows for quick stride into new spaces as people use lessons learned and techniques from past predictive algorithms to develop effective algorithms elsewhere. Some government actors have already acknowledged such a market dynamic would be helpful in facilitating innovation. In Chicago, the city's Department of Innovation and Technology created an algorithm to identify restaurants at greatest risk of food

safety violations. The department has made the underlying data for this algorithmic available through a portal, as well as made the algorithm itself available online for others to use or evaluate. The department has noted that they are open to critiques and may incorporate suggested improvements into an improved model.[5] The benefits made possible through public dissemination of algorithm development in this one instance would be just as applicable, and at a much larger scale, if this was required of privately created algorithms as well.

Lastly, this policy ought to largely mitigate the misaligned incentives around transparency from the current development system. Easy accessibility to algorithmic underpinnings would allow for greater accountability for faulty algorithms. Anybody from a citizen-sleuth to a government watchdog to an investigative reporter could dissect how an algorithm makes its decisions. Right now, investigative reporters occasionally can note disturbing trends based on visible disparities in society. For example, a reporter from *Wired* noted how patients with crippling pain were being turned away from doctors based on the assessment of Narxcare, an algorithm now frequently suggested by government to predict risk of opioid addiction. The reporter was able to find several disturbing signs of inaccuracy (the algorithm did not factor in that some patients used opioids for pet needs and incorporated questionable studies suggesting childhood sexual abuse tripled risk of addiction). However, they were unable to completely what information Narxcare based its recommendations on because Appris (Narxcare's developer) claims its algorithms are proprietary.[6] Requirements that this information be publicized would ameliorate issues like this. While uncovering every problem with predictive algorithms would probably take years, there would be immediate benefits as algorithms with blatantly pernicious effects could be retired or rapidly improved.

---

[5] Means, Andrew. "When Algorithms Run the Government." *Digital Impact.* July 6, 2016.
[6] Szalavitz, Maia. "The Pain Was Unbearable. So Why Did Doctors Turn Her Away?" *Wired.* August 11, 2021.

Skeptics of this reform might argue that requirements to publish details of algorithm operation diminish the incentives to create them to begin with. Perhaps to a degree this is true. But this reform has the potential to usher in a market dynamic based around iteration and improvement, rather than siloed creation of algorithms from scratch. Companies may be less focused on creating completely bespoke predictive algorithms that they expect to generate lucrative payments from contracts. But companies instead now have incentives to tinker with existing algorithms tailoring the for uses in certain states or service areas. This is probably healthier than the currently siloed market, where companies labor independently to produce products but rarely can build on lessons from each other's work. Under that dynamic, improvements are slow and incremental. Additionally, even in a market where algorithms are publicly accessible there is still a first mover advantage that incentivizes creation of new-in-kind algorithms.

Secondly, it should be noted that in some instances even if a company publishes all its processes, not all of an algorithm's decisions may be understandable. Some algorithms are built like a "black box": they base decisions around enormous datasets but are not built in such a way that allows even developers to understand the criteria and factors that lead them to these decisions. However, even in these instances, publication of this process is useful to the public. An algorithm that seems to generate arbitrary decisions and is built on "black box" reasoning will face enormous skepticism from the public; there will be heightened pressure to understand why the algorithm reached the decision it did and explain that process. Firms that can clearly trace a path for *how* a predictive algorithm made the decision that it did will have an advantage in this market. Incentives to create algorithms that are comprehendible to humans is beneficial to society.

While not a panacea, government requirements to publicize all datasets and processes related to predictive algorithms from developers would be tremendously useful. Through this reform, government would be facilitating a much more innovative market related to algorithms even as they lower the likelihood that inaccuracies in predictive algorithms disadvantage sectors of the American population.