THOUGHT PROVOKING IDEAS OF THE GLOBAL ESSAY COMPETITION 2021

Rebuilding Trust in Algorithm-driven Public Services

Janz Irvin Chiang is one of the top three contributors to this year's Global Essay Competition Award. He studied at Beijing University and attended the 50th St. Gallen Symposium as a Leader of Tomorrow.

Introduction

We live in a world of algorithms. Each day we entrust more and more of our lives to these bundles of 1's and 0's, and it makes sense that we would do so. Algorithms promise to be more efficient at analyzing troves of data and generate results free from human error. They rely on billions of data points and defined parameters. And we trust them because we believe that algorithms arrive at decisions not influenced by stereotypes or whichever stars were present at the time of one's birth. Nonetheless, algorithms "are built upon socially derived perceptions and under-standings, not fixed universal, physical laws." (Eischen, 2002). They are not absolute, and they are not without error.

As the proliferation in the adoption of algorithms for determining resource alloca-tion increases, so too has the cases of adoption by public institutions, which has given algorithms more power than ever before. Whereas users can choose to avoid using biased searched algorithms, citizens do not have the luxury of alternatives when it comes to algorithm-driven public services. Mistrust in algorithm-driven public services due to cases of prejudice generates negative views of public institutions and reduces the efficiency of public resource allocation.

Algorithm-driven public services are growing in popularity and poised to become the norm moving forward. Personal experiences or reports of biased algorithm-driven resource allocation are valid reasons for mistrust, but public institutions can change that. Public institutions need to reassess how algorithms operate to ensure that it works inclusively and equitably. Therefore, **public institutions can rebuild trust by implementing solutions that promote transparency and inclusivity to address biases and imperfections in algorithms**. Algorithms are not perfect and may never be perfect, so the process through which institutions and various direct stake-holders address such defects is crucial. Public institutions need to rebuild our trust in algorithms lest we miss out on the benefits derived from the appropriate and effective deployment of algorithms.

Algorithm-driven public services

Data-driven algorithms promise significant benefits over how governments and their institutions perform their functions.

Algorithms can harness massive amounts of data to process them in ways that promise to deliver objective results, thereby taking out the "human" factor in decisions that can commonly introduce biases. Whether it is the use of algorithms to determine conditions for bail in the justice system based on the likelihood of the defendant skipping bail or using detailed demographic information to determine optimal resource allocation in a region, algorithms will only continue to play an increasingly important role in the delivery of public services.

The Congressional Research Service of the United States of America conducted a study on the use of risk and needs assessment instruments to determine an offender's risk of recidivism. Researchers noted that even when the algorithm has no inherent racial bias, an instrument's application could have a disparate racial effect (James, 2018). A risk assessment algorithm was found to be biased against African Americans when it assigned a higher score to one compared to a white seasoned-criminal whose theft was valued at a similar amount (Angwin and Larson, 2016). A study of a predictive policing algorithm found that it relied on crime reports driven in part by racially-biased policing, which created a loop of over-policing in minority neighborhoods (Chammah, 2016). An algorithm used for visa application screening has been found to speed up the process for white people

This is not to say that all algorithms are wrong. Algorithms used to detect money laundering are objectively good for the public because they can significantly reduce the amount of human intervention required to process terabytes of data on financial transactions. Using algorithms that can learn and be fine-tuned by engineers when required, institutions can help prevent the illicit use of capital.

To ensure fairness, algorithms must be open for audit and examination in some way. To say that algorithms must be 100% objective before the government uses them is an incredibly high bar to set and, if anything, will stifle innovation. Instead, the goal should be to create ways for the concerned public to audit algorithms so that they can be corrected if there are biases. Recognizing the imperfection of algorithms create an opportunity to improve them, thereby beginning the process of rebuilding our trust in algorithms.

Implications of mistrust

The use of algorithms by public institutions is one of many forms of contact that citizens have with the government. So, in many ways, the quality and consequences of those reactions contribute to citizens' trust or mistrust in public institutions. Algorithms can exacerbate existing systemic biases or be used as tools by bad actors to disenfranchise oppressed groups. Rebuilding public trust in algorithms used by public institutions helps create positive and meaningful interactions.

Promoting transparency and inclusivity rebuild trust in algorithm-driven public services

Transparency and inclusivity are the fundamental principles behind rebuilding trust in algorithm-driven public services.

Here, transparency forms the foundation for driving inclusive solutions to provide public services more equitably. The secretive and incredibly technical nature of algorithms makes it difficult to rebuild our trust in them, so transparency is vital. Here, I propose three solutions that promote transparency: assigned audit responsibility, benchmark dataset submission, and improved communication.

Assigned audit responsibility

This solution seeks to address the need for a thorough understanding of algorithms being used by public institutions and the protection of intellectual property and trade secrets. Rather than allowing the public to inspect the source code to understand how data is

processed, this solution asks that the public assign the exercise of their right to information to expert working groups. These expert working groups will comprise relevant experts, such as social workers, mathematicians, and policy experts, who will be subject to non-disclosure agreements and other applicable contracts and laws. A more straightforward way to understand this would be to compare it to a third-party audit of security software providers where external parties are invited to inspect the code and give their unbiased assessment in a way that is both objective and protects trade secrets.

Benchmark dataset submission

Separate from assigned audit responsibility, this solution would allow the public to submit benchmark datasets to public institutions for processing, and the results would be provided to everyone (Diakopoulos, 2016). Here, the outcomes can be compared by data types to determine whether biases or prejudices exist. There are two benefits to this solution. The first benefit is that datasets would not violate intellectual property and trade secrets since the source code and processing would not be exposed. Only the results from the submission of test datasets would be made public. The second benefit is that were there to be a benchmark dataset, the same dataset can be submitted to be processed by similar algorithms to compare their performance. The results can then be used for comparative and improvement purposes. However, certain assumptions will have to be made on the processing of information based on the output since the source code will not be inspected.

Improved communication:

The lack of information invites speculation and, in some cases, conspiracy thinking. Even if the public can audit the algorithm itself, the lack of technical knowledge is a significant barrier to rebuilding trust. Public institutions need to improve how they communicate their

use of algorithms in delivering public services. They should clearly communicate their data sources, how each piece of data is being used to generate resource allocation decisions, and the data quality control mechanism in place.

The implementation of the above solutions creates the foundation for inclusive solutions. Transparency allows the public to identify biases in algorithms, while inclusivity will enable stakeholders to address these biases, leading to more equitable algorithm-driven public services. Here, I propose three solutions that promote inclusivity: creating inclusive development teams, questioning the data, and fostering ongoing conversations.

Inclusive development teams

Biases are often subconscious and invisible until deeply examined, which is why this solution would reduce that by introducing more diverse voices in the development of algorithms. The inclusion of developers who come from similar backgrounds as the people they intend to help will expand the conversation to include diverse experiences crucial to empathizing with the public. Developers who know of the challenges in accessing welfare through direct or indirect experiences are likely to be more cognizant of biases perpetuated by welfare systems than someone from an affluent background.

Question the data

As with anything, it is garbage in, garbage out when it comes to algorithms. With sufficient transparency, stakeholders can question how data is being collected and what types of data are being fed into the algorithm. Certain types of data reflect and can potentially magnify prejudices. For example, police resource allocation algorithms that rely on crime reports sometimes reinforce over-policing in neighborhoods dominated by minorities (Buranyi, 2017). Without questioning the data being used, public institutions risk

exacerbating inequalities through algorithm-driven public resource allocation.

Ongoing conversation

Ongoing conversation drives continuous feedback, which feeds into improving algorithms. Just as companies talk to their customers to determine their needs and pain points, so too should public institutions and algorithm developers. It is essential to listen to the people being impacted. The continuous involvement of direct stakeholders is crucial to continuing a journey towards more inclusive outcomes.

Rebuilding trust for inclusive outcomes

Algorithm-driven public services are here to stay, but it does not mean that we need to accept them as they are now. Public mistrust in algorithms due to personal experiences or reports of biased resource allocation exists and must be urgently addressed. Continued mistrust runs the risk of eroding broader trust in public institutions themselves.

Algorithms are not perfect, so the public must know when they are biased so that corrections can be made. Therefore, rebuilding trust in algorithm-driven public services through solutions that promote transparency and inclusivity are vital to equitable algorithm-driven public services and creating more touchpoints to generate trust in public institutions.

References

Angwin, Julia, and Jeff Larson. "Machine Bias." Text/html. ProPublica, May 23, 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Buranyi, Stephen. "Rise of the Racist Robots – How AI Is Learning All Our Worst Impulses." the Guardian, August 8, 2017. http://www.theguardian.com/inequality/2017/aug/08/rise-of-the-racist-robots-how-ai-is-learning-all-our-worst-impulses.

Chammah, Maurice. "Policing the Future." The Marshall Project, February 3, 2016. https://www.themarshallproject.org/2016/02/03/policing-the-future.

Diakopoulos, Nicholas. "Accountability in Algorithmic Decision Making." ACM, February, 2016. https://cacm.acm.org/magazines/2016/2/197421-accountability-in-algorithmic-decision-making/fulltext.

Eischen, Kyle. "The Social Impact of Informational Production: Software Development as an Informational Practice." UC Santa Cruz: Center for Global, International and Regional Studies, 2002. https://escholarship.org/uc/item/9rp0c3w4.

James, Nathan. "Risk and Needs Assessment in the Federal Prison System." Washington DC: Congressional Research Service, July 10, 2018. https://fas.org/sgp/crs/misc/R44087.pdf.

Marsh, Sarah. "Councils Scrapping Use of Algorithms in Benefit and Welfare Decisions." the Guardian, August 24, 2020. http://www.theguardian.com/society/2020/aug/24/councils-scrapping-algorithms-benefit-welfare-decisions-concerns-bias.