Self-Driving Black Box: Algorithms and Automation

From just 145,000 lines of code to land humans on the moon in 1969 to the more than 2 billion lines of code that run Google's search engine, today's technological systems are labyrinths of data structures (Leetaru, 2016). In half a century, the demand for complex problem-solving has increased exponentially, forcing the algorithms that underline these technologies to become more intricate in tandem. Yet, technology is now entering uncharted territory where algorithms act beyond comprehension and prediction: the black box. Technology that involves algorithms so advanced and sophisticated that they result in a lack of human understanding are deemed "black box" systems because they can only be observed in terms of their inputs and outputs, while their internal workings remain mystery. But this disengagement comes at a cost; with progressive technology comes progressively consequential societal implications. One looming algorithm-based technology is autonomous vehicles, the great promiser of the transportation revolution. Over the last few years, leading developers like Uber have been navigating challenges posed by the intricacy of driverless cars to bring algorithms up to reliable standards. However, this same intricacy also provides a veil and excuse for deceptive activity and exploitation of consumer data. Extreme lack of understanding among both consumers and developers regarding black box algorithms in autonomous vehicles ultimately results in safety concerns, privacy infringement, and social inequalities.

Black boxes embody a paradox of the Information Age: data is growing in breadth and depth, yet the information most important to us is often out of reach. When it comes to technological products, complex problems are simplified by hiding all but the relevant data

through abstraction. This results in an efficient system but no clear idea of how data is used, who it is used by, and to what ends. When algorithms are structured by billions of lines of code and guarded by developers, consumers are often incapable of understanding how platforms produce the information presented to them or use the data that they input into a system. Conversely, now that algorithms are able to restructure their internal workings and "think" for themselves, developers can lose control over the decisions that their algorithms make. In a technological society where knowledge is power, the implications of being barred from how knowledge is created are potent and far-reaching. Companies can hide malpractice. Developers can distort the image of their technologies. Consumers' privacy can be breached and their data abused. Algorithms can inflict inadvertent harm (Matheson, 2019). When understanding of black box algorithms is lacking, there are two primary consequences: (1) lack of understanding among developers and manufacturers may result in the unleashing of technology that inadvertently impacts society and (2) lack of understanding among consumers allows developers and manufacturers to exploit users and their data.

Are these real concerns? The impervious nature of black box algorithms implies that we are often unconscious of their very existence. That is, until they *do* wreak consequence. Black box scandals have recently claimed national headlines. In 2018, data-analysis firm Cambridge Analytica improperly obtained and used the data of some 50 million Facebook users to serve pro-Trump ads for the 2016 election. It was later revealed that Facebook's terms of service secretly granted third parties the ability to access users' profiles and sensitive data, making it available for Cambridge Analytica to then influence those very data constituents (Confessore, 2018). In similar black box fashion, two of Boeing's 737 Max airplanes crashed within a five

month span in 2019 as a result of a flawed flight control system that forced the planes into uncontrollable nose dives. The company's egregious algorithmic miscalculations evaded pilots, regulators, developers, and consumers before 346 people onboard lost their lives (Schaper, 2019). The black box algorithms behind Facebook's data collection and airplanes' control systems are relatively fundamental implementations of administrative software, yet slight deceptions and bugs have had drastic political and fatal repercussions.

With a technology as algorithm-dependent as self-driving cars, fraud and errors will only bring about more severe ramifications. Autonomous vehicle developers rely on algorithms mainly for artificial perception and data collection. Uber, the technology giant at the forefront of driverless ventures, has been experimenting with these two lucrative algorithms in their push towards the autonomous future. However, the intrinsic black box nature of artificial perception and extrinsic black box nature of data collection and analysis have proven dangerous to society's urban, economic, and social constructs.

Safety is jeopardized when autonomous car developers lack understanding of how their black box algorithms make decisions. With manual cars, most users have little conception of the internal connections between the gas pedal and exhaust pipe, let alone of the electronics systems that control GPS or FM radio. Abstraction is further exacerbated in driverless cars, where the algorithms behind automated machine-learning technology are deeply entrenched black boxes. One of the leading benefits that autonomous vehicles promise is better safety standards than manual vehicles by mitigating human perception errors at the wheel. Deep learning, a recent algorithmic breakthrough that allows cars to recognize objects at superhuman capacity, is the crux of the black box phenomenon. By consuming massive sums of data, algorithms can "learn"

from prior experiences at an exponential rate, similar to the neural processes of the human brain. Deep learning algorithms involve millions of artificial neurons which form and modify their connections with new data, disengaging developers from the algorithm's structure. Certainly, the average consumer knows little about these algorithms, but in using an autonomous vehicle assumes their validity and places trust in the developers. Yet, there is no way for developers to predict or program what a deep learning algorithm will correctly or incorrectly recognize, leaving fateful decision-making on the roadway up to code and probabilities. Since deep learning algorithms designate perception between people, animals, street signs, buildings, and the like, errors can be fatal (Lipson et al., 2017).

In 2018, an Uber self-driving test vehicle struck and killed 49-year-old Elaine Herzberg in Arizona as she walked her bike across the street. This first pedestrian fatality associated with driverless cars was later revealed to be the result of flaws in Uber's object recognition software, which failed to properly identify the woman as a pedestrian until 3.3 seconds too late. In fact, the system's structure did not include a consideration for jaywalking pedestrians at all. According to the National Transportation Safety Board's report of the incident, the underdeveloped deep-learning algorithm turned out to be additionally responsible for 37 other non-fatal crashes over the prior 18 months (Shepardson, 2019).

While driverless cars promise safer roadways by removing human error from the driving equation, relying on an algorithm to correctly guide a speeding vehicle through hazardous downtowns, constricted highways, and winding sidestreets with erratic pedestrians requires that the algorithm be virtually infallible. As demonstrated by Uber's fatal autonomous accident, deep learning algorithms are a matter of uncontrollable and unprogrammable probabilities.

Self-driving cars can react quicker, more accurately predict, and better navigate than humans, but their permanently rogue algorithms raise ethical concerns (Lipson et al., 2017): Should we put black box algorithms in a position to take human lives? Who is responsible in these situations? The developer? The consumer? Can *anyone* be expected to accept blame when deep learning algorithms make decisions for themselves?

When autonomous car consumers lack understanding of how their data is collected, analyzed, and ultimately exploited by developers, privacy breaches can have severe economic and social implications. As technology grows more complex and vigilant, and increasingly simple interfaces are underlined by increasingly complex algorithms, consumers lose sight of how their data feeds ever-present data-harvesting algorithms. Developers employ "technological propaganda" by making algorithms appear much more difficult than need be, thus isolating data surveillance and "rocket science" in professionalism (Feyerabend, 1993). For instance, according to a 2018 digital understanding survey by the think tank Doteveryone, 83 percent of internet users were unaware that their personal information can be indirectly collected (Miller et al., 2018). Such terms carve out a domain of ignorance in society because many people feel that technology is too complicated for them to understand and is better taken at face value.

Uber's use of black box data surveillance algorithms currently allows them to micromanage and exploit consumers, and sheds light on how these privacy breaches will increase with the advent of autonomous vehicles. In their manual cars, Uber uses black box data-harvesting to manipulate riders through incentives and price discrimination. Because Uber serves as the overseer of drivers and consumers, the intermediary between them, and the leaser of the technology that the two parties rely on, they have monopolized every aspect of the

ridesharing experience. Such authoritarian control allows Uber to access riders' smartphones, track their whereabouts, use their social networks for promotions, and analyze users' actions within the Uber app. While Uber insists that their pricing practices are not discriminatory, they subtly modify pricing distributions without notifying drivers or riders. Concealed by the veil of algorithms, they analyze passenger data and identify those willing to pay higher fees. One Uber driver that I interviewed, Bulent, reported driving a well-off passenger from Grand Central Terminal to Port Washington, Long Island for $350 without any influences of surge pricing or premium models -- a 23 mile trip that should have cost $80.51 according to Uber's claimed pricing index. And, whenever Uber faces criticism for profiteering, they are able to use their algorithms as a scapegoat rather than accepting the blame themselves (Rosenblat, 2019).

Driverless cars will be no exception to these patterns. They will be outfitted with cameras, microphones, and tracking devices which will capitalize on users' habits in the passenger seat by selling sensitive data to advertisers to commercialize the self-driving experience. While a suggested stop for branded coffee may feel routine to oblivious consumers, developers will deceptively profiteer from black box algorithms because users do not understand the role of their data inputs in analytics (Lipson et al., 2017). Data collecting mediums will be covered in shrouds of complexity to move them into a black box and out of the reach of consumers to scrutinize. Even if criticism does arise, developers will skirt responsibility by deflecting blame to an algorithm. These data malpractices can be particularly dangerous when their influence extends outside of a business model and dips into the social sphere.

All consumer data is not created equal; existing social inequalities are exacerbated and new inequalities are introduced by targeting marginalized groups for data collection with black

box solutionism. As technology grows more complex and vigilant, developers and urban planners seek to solve more problems using data surveillance and management. At the same time, consumers believe that innovation and progress stem directly from technology, and are willing to instill trust in algorithm developers to make effective, socially mindful decisions. Ben Green, author of The Smart Enough City, labels this tendency as "tech goggles." "At their core, tech goggles are grounded in two beliefs: first, that technology provides neutral and optimal solutions to social problems, and second, that technology is the primary mechanism of social change" (Green, 2019, 4). However, when rigid algorithms are used to manage nuanced social dynamics, tech goggles can result in inequality. While data collection touches all consumers, diminishing privacy primarily impacts the poor and minorities. Most marginalized individuals are more concerned about privacy than society's elites, but traditionally have narrower access to technology and lower degrees of education. With an incomplete knowledge of privacy and policy algorithms, these individuals are unable to reduce the extent to which they are monitored. In turn, their data is used to manage and profile marginalized groups instead of to fix lacking education and access to technology, which require a more socially and politically involved effort than raw data analysis (Eubanks, 2017).

Since data-harvesting is a fundamental feature of driverless cars, marginalized groups will face exacerbated dicscrimination when the black-box-filled transportation revolution arrives. Marginalized individuals will have no conception of how their data is being extracted or what is being done with it, but will see the effects that it has on their social isolation. Developers will take advantage of their incapacity for complex technology by increasing the already high levels of data collection that marginalized groups face. Governments will use this data for management:

they may assess and analyze the tendencies of marginalized individuals, resulting in punitive public policy but doing nothing to attack systematic root causes. Autonomous vehicle companies will use this data for prejudice: they may identify lower class individuals and refuse rides, saving the passenger seat for riders willing to pay higher fees.

Black box algorithms are an inevitable feature of driverless cars' software, but pose serious safety concerns beyond developers' programming and expectations. Black box algorithms are also a tactical tool for developers to capitalize on consumers' incomplete technical knowledge, jeopardizing society's prized democratic values of privacy and equality. While autonomous vehicles are not yet commonplace, fatal faults in early testing and implications of current data surveillance reveal that we must be skeptical of autonomous ubiquity.

The key to redirecting the dangerous trajectory of black box algorithms in driverless cars is transparency. Technology should be more easily understood and open about its priorities so that consumers, developers, and society as a whole are not exploited by the very technology designed for betterment. Currently, governments regulate industry because companies perform public functions along with private profit-seeking ones. Instead, citizens should be given a voice in this regulation so that companies are held more accountable for their technology. To this end, developers should be required to make their software open-source, meaning that it is open for public regulation and improvement. With a democratic system fully open to the scrutiny of consumers and experts invested in technology's success, deep learning algorithms for autonomous cars would be much more likely to be error- or omission-free. If developers want to retain private control over their code, then they should be required to be candid about what their

code does and how it does what they claim. This would allow consumers to be more conscious of privacy breaches and give marginalized groups the ability to understand in simple terms the algorithms being used against them so that these individuals can repress. While the goal of black boxes is to increase speed and efficiency, these reforms would slow developers and cost time (Pasquale, 2016). Yet, we have seen the instabilities that black box algorithms pose to our urban, economic, and social constructs. We must ask ourselves: What kind of society do we really want?

Works Cited

Confessore, N. (2018, April 4). Cambridge Analytica and Facebook: The Scandal and the Fallout So Far. Retrieved from https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

Eubanks, V. (2017). *Automating Inequality*. New York, NY: St. Martin's Press.

Feyerabend, P. (1993). Against method. London: Verso.

Green, B. (2019). *The Smart Enough City*. Cambridge, MA: The MIT Press.

Leetaru, K. (2016, January 4). In Machines We Trust: Algorithms Are Getting Too Complex To Understand. Retrieved from https://www.forbes.com/sites/kalevleetaru/2016/01/04/in-machines-we-trust-algorithms-are-getting-too-complex-to-understand/#60bf7d9a33a5

Lipson, H., & Kurman, M. (2017). *Driverless: Intelligent Cars and the Road Ahead*. Cambridge, MA: The MIT Press.

Matheson, R. (2019, May 31). Cracking open the black box of automated machine learning. Retrieved from http://news.mit.edu/2019/atmseer-machine-learning-black-box-0531

Miller C, Coldicutt R and Kitcher H. (2018) People, Power and Technology: The 2018 Digital Understanding Report. London: Doteveryone.

Pasquale, F. (2016). *Black box society: the secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.

Rosenblat, A. (2019). *Uberland: how algorithms are rewriting the rules of work*. S.l.: University of California Press.

Schaper, D. (2019, November 27). 737 Max Scandal Cuts Boeing's Once Rock-Solid Image. Retrieved from https://www.npr.org/2019/11/26/783197253/737-max-scandal-cuts-boeings-once-rock-solid-image

Shepardson, D. (2019, November 6). In review of fatal Arizona crash, U.S. agency says Uber software had flaws. Retrieved from https://www.reuters.com/article/us-uber-crash/in-review-of-fatal-arizona-crash-us-agency-says-uber-software-had-flaws-idUSKBN1XF2HA