2018

Algorithmic Accountability And Individual Rights: A Case Study Analysis

Karina Jaimes
University of Alabama at Birmingham

ALGORITHMIC ACCOUNTABILITY AND INDIVIDUAL RIGHTS: A CASE STUDY ANALYSIS

by

KARINA JAIMES

ERIN L. BORRY, COMMITTEE CHAIR
AKHLAQUE HAQUE
PETER A. JONES

A THESIS

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Master of Public Administration

BIRMINGHAM, ALABAMA

2018

ALGORITHMIC ACCOUNTABILITY AND INDIVIDUAL RIGHTS:
A CASE STUDY ANALYSIS

KARINA JAIMES

MASTER OF PUBLIC ADMINISTRATION

ABSTRACT

Algorithms are becoming widespread and their use is set to expand, but it must be emphasized that, although constructive and transformative, they can also be destructive and counterproductive. Various types of organizations utilize algorithms, including government. The consequences of poorly constructed algorithms could lead to the infringement of citizens' individual rights. Exploring two algorithms as case studies allows us to understand government use of algorithms and the ethics associated with them, especially in relation to the potential for the violation of rights. The application of privacy and anti-discrimination as lenses through which algorithms can be analyzed allows government officials to realize what the implications of unregulated algorithms can have on society. The more widespread the use of these algorithms, the higher the risk of infringing rights protected by law and the Constitution. In order to reach the conclusion that indeed algorithms must be regulated in whichever way, government must enforce accountability for the consequences imposed by biasedly constructed algorithms.

Keywords: algorithmic accountability, privacy, anti-discrimination, graffiti, recidivism

ACKNOWLEDGMENTS

First and foremost, I would like to sincerely thank the chair of my thesis committee, Dr. Erin L. Borry for all of her help, patience, and understanding throughout the process. Her incredible knowledge and dedication are greatly appreciated, and I cannot imagine having finished this successfully without her. I could not thank her enough for all of the advice and guidance she has shared with me.

I would also like to thank the rest of my thesis committee: Dr. Akhlaque Haque and Dr. Peter Jones. Dr. Haque was the subject matter expert of my committee and offered invaluable insight for my final edits. The idea for this thesis came about a conversation with Dr. Jones so nothing would have happened without his encouragement to pursue this research topic.

My sincere gratitude also extends to the rest of my undergraduate and graduate professors at the University of Alabama at Birmingham. In the Department of Political Science and Public Administration: Dr. William M. Beale, Dr. Nevbahar Ertas, Dr. Wendy Gunther-Canada, Dr. Angela K. Lewis, Dr. Lisa Sharlach Dr. Nikolaos Zahariadis, and my graduate advisor, Mrs. Carin Mayo. Also, to my undergraduate professors in the Department of Foreign Languages and Literatures, especially Dr. John K. Moore and Dr. Carlos Orihuela, thank you for sharing your passion for languages and cultures with me. I've acquired astounding knowledge from all of you that no one will ever be able to take from

me and that has greatly shaped who I am today and for that I am blessed and forever in-debted.

A special thanks to Dr. Sarah Abroms Kunin for her unfaltering dedication. I definitely could not have seen this through without her help, advice, and attention.

To all of my family and friends, thank you for putting up with my difficult attitude during this process, and I apologize for missing out on so much these past two semesters. Sadi, DeAndre, Arsh, Hafez, thank you for all of your encouragement and for always offering a shoulder to lean on. Angelica and Brian, thank you for always keeping me afloat even though there have been thousands and thousands of miles between us during this time. Abe, Gus, Brooke, I am lucky to call you guys family.

My parents, I would be nothing without your tremendous sacrifices. Thank you for your unfaltering support, spiritually, financially, and everything in between. A lifetime will not be enough to ever repay you both.

And to my Lord, thank You for gifting me with this life, the people in my life, and all the opportunities laid before me. I hope to use the gifts You have granted me to serve others always in Your Likeness.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

Soon after Instagram announced that co-founders Kevin Systrom and Mike Krieger were stepping down, Adam Mosseri quickly decided to change certain aspects of the widely used social media platform (Pardes, 2018). One of the first issues addressed by Mosseri was bullying. The technology team created an algorithm that can filter out any sort of offensive content by finding "instances of bullying" to be sent to people for review. The algorithm detects "attacks on a person's appearance or character, as well as threats to a person's well-being or health" (Pardes, 2018). Additionally, this algorithm will soon be applied to comments left on Instagram's live feature.

Algorithms run various applications that people use regularly. While these may seem to perform relatively simple and progressive tasks, it is still important that users are aware of the technological advancements being made. Algorithmic machines are meant to facilitate tasks and minimize human interactions, and thus, the possibility of error. This minimization of human error also intends to eliminate bias. Algorithms are embedded in everything from deciding the probability of a patient receiving an organ transplant (Diakopoulos et. al., 2018), to who will qualify for a loan or social services (New and Castro, 2018, p. 6), and even to assessing the risk of recidivism among prisoners (Angwin et. al., 2016). Social services such as food stamps and Medicaid are allocated by programs like the Indiana Client Eligibility System which is a system made up of "networked main-

frame computer server, software, and desktop terminals" that seek "to assume the decision making and distribution of welfare resources" (Dennis, 2006, pps. 553-554). Additionally, communication methods such as social media platforms like Instagram, Facebook, Snapchat, etc., all rely on algorithmic systems to function properly. Any form of daily tasks that use software consist of algorithms, and when put into perspective, it is obvious that everyday life is, in some manner, directly impacted by algorithms.

Algorithms are becoming widespread and their use is set to expand, but it must be emphasized that, although constructive and transformative, they can also be destructive and counterproductive. Therefore, it is important that we are aware "about how algorithms exercise their power over us" (Diakopoulos, 2013, p. 1). In addition, instead of mere awareness of the presence of algorithms, knowing how algorithms are formulated is equally, if not more, important.

What Are Algorithms?

Algorithms, according to Gillespie (2014, p. 1), are "encoded procedures for transforming input data into a desired output, based on specific calculations." Although algorithms in essence do perform automatically, the manner in which they function is premeditated. This means that an algorithm is not a creation of its own, but that of a human. That is, people must program algorithms to do what they are intended to do. Moreover, even though an algorithm is triggered without human intervention, the data fed into the algorithm can only be formalized once primed by a human programmer. Algorithms are simply tools that help an individual or an organization make better decisions (House of Commons, 2018).

Algorithms are made up of data that is extrapolated from public activity, particularly through communication networks (Gillespie, 2018, p. 4). These networks are designed to pick up on all digital traces which are then turned into databases. Gillespie (2018, p. 4) provides the following example: "Google…crawls the web indexing websites and their metadata. It digitizes real world information, from library collections to satellite images to comprehensive photo records of city streets. It invites users to provide personal and social details as part of their Google+ profile. It keeps exhaustive logs of every search query entered and every result clicked. It adds local information based on each user's computer's data. It stores the traces of web surfing practices gathered through their massive advertising networks." The collected data is then placed within databases which is then categorized by programmers.

This data has become multifaceted, so one piece of data can now end up having multiple associations with other pieces of data. This associative characteristic is vital for the categorization aspect of algorithms. When accessed by programmers, data can be organized in a multitude of ways, and in turn, can be likened loosely connected to other sets of data. Characterization in turn becomes both a "powerful semantic" and a sort of "political intervention" because data must be organized into categories left at the discretion of whomever is constructing them (Gillespie, 2014, p. 5). Questions such as "what the categories are, what belongs in a category, and who decides how to implement the categories in practice" must be taken into consideration when constructing categories (Gillespie, 2014, p. 5). Gillespie states that once the category of data is determined that category becomes the point of reference that will be treated with "reverence" by all approaching algorithms (2014, p.5).

For example, “in 2009, more than fifty-seven thousand gay-friendly books disappeared in an instant from Amazon's sales lists, because they had been accidentally categorized as "adult." Naturally, complex information systems are prone to error, but this particular error also revealed that Amazon's algorithm calculating "sales rank" is instructed to ignore books designated as adult. Even when mistakes are not made, whatever criteria Amazon uses to determine adult-ness are being applied and reified -- apparent only in the unexplained absence of some books and the presence of others” (Gillespie, 2014, p.5). This incident came to be dubbed #amazonfail. This formulation of categories, constructed by human subjects, although automated, still leaves room for biases and error.

Various types of organizations utilize algorithms, including government. Government “transcends all sectors in a society” and as such, not only provides the foundations for all things legal, economic, and political, but it also “exerts significant influence on the social factors that contribute to [societies’] development” (Elmagarmid and McIver, 2001, p. 1). One of those developments has been automation and the facilitation that automated systems provide for the fast-paced world we live in today. In effect, government institutions have also turned digital (via e-government) in order to facilitate services and encounters with their constituents. According to Elmagarmid and McIver (2001, p. 2), e-government systems are intended to provide one of four of the following levels of service:

1. First-level services provide one-way communication for displaying information about a given agency or aspect of government.
2. Second-level services provide simple two-way communication capabilities, usually for uncomplicated types of data collection such as registering comments.

3. Third-level services facilitate complex transactions that may involve intragovernmental workflows and legally binding procedures. Examples of such services include voter and motor vehicle registration.

4. Fourth-level services seek to integrate a wide range of services across a whole government administration, as characterized by the many emerging government portals.

Algorithms are a tool that governments can use for e-government purposes because algorithms "stand in for calculations and processing that no other human could ever do on their own," and because "critically, algorithms do not make mistakes" (Caplan et.al., 2018, p. 3). Public entities' intended use of algorithms should be to facilitate institutions' line of work in order to more efficiently and effectively serve constituents while still protecting the individual rights of citizens.

Individuals' rights can be protected against potential impacts of algorithms by way of algorithmic accountability. Algorithmic accountability is "the principle that an algorithm should employ a variety of controls to ensure the operator can verify it acts in accordance with its intentions, as well as identify and rectify harmful outcomes" (New and Castro, 2018, p. 1). It is suggested by experts that by applying algorithmic accountability, government institutions would be able to both "promote the vast benefits of algorithmic decision-making and minimize harmful outcomes, while also ensuring laws that apply to human decisions can be effectively applied to algorithmic decisions" (New and Castro, 2018, p. 2). Institutions would be able to exploit the efficiency and efficacy that algorithms can provide while still protecting the basic rights of the individuals affected by the algorithm.

Research Questions

Scholars have begun questioning the ethical considerations of big data and automation (Jurkiewicz, 2018; Borry and Getha-Taylor, 2018), but little is known about ethical considerations of algorithms. Thus, this research seeks to answer the following questions:

1. How do governments in the United States currently use algorithms?
2. What individual rights might algorithms infringe upon?
3. What role does algorithmic accountability play in the United States?

To answer these questions, I first present a literature review that explores what is currently known about government use of algorithms. I then introduce the ways in which algorithms have the potential to lead to unintended consequences, such as violating individual rights. Two specific rights under investigation are privacy and anti-discrimination. To explore the potential for algorithms to infringe on these rights, I evaluate two case studies: one about a recidivism algorithm and the other about a gang graffiti recognition algorithm. I explain both cases and apply the lenses of privacy and anti-discrimination to understand the potential for those algorithms to negatively impact individuals and their rights. Given these cases, I show that regulation may be an answer to this lack of regulation which would provide algorithmic accountability.

Contributions to Literature

This case study analysis will help further understand the vast use of algorithms and their impact on daily lives, particularly by public entities. The research questions posed in this thesis seek to understand the ways in which algorithms can, if at all, violate

individual rights and to explore the potentially critical role of algorithmic accountability. Considering that algorithms are indeed a new technology that has grown exponentially, it is difficult to know what some of the long-term consequences that algorithms may have on society. This research provides an opportunity to explore those consequences, especially if algorithms can have negative impacts on people and their rights.

Exploring two algorithms as case studies allows us to understand government use of algorithms and the ethics associated with them, especially in relation to the potential for the violation of rights. The application of privacy and anti-discrimination as lenses through which algorithms can be analyzed allows government officials to realize what the implications of unregulated algorithms can have on society. Analyzing algorithms through the values upheld by the Constitution, provides a compass by which institutions, both private and public, can be held accountable. The application of individual rights as lenses for determining whether certain algorithms provides a new perspective that seeks to highlight the manner in which algorithms could violate the law.

What's to Come

The rest of this thesis is organized as follows. Chapter two provides a literature review and explores what an algorithm is and how they are currently being used by our government. The concept of algorithmic accountability is introduced. Additionally, accountability and government accountability are defined in this chapter. The chapter also introduces the potential for the violation of individual rights—privacy and anti-discrimination— by way of unregulated algorithmic practices. Chapter three presents the data and methods used for this case study analysis research as well as describing the case studies

being analyzed through the lens of privacy and anti-discrimination. Finally, chapter four includes the conclusion and a discussion which presents the idea of regulation as a mean to hold institutions accountable for their algorithms.

CHAPTER 2

LITERATURE REVIEW

Considering the ever-changing nature of technology in society, it is important to consider the consequences that said technology could have on everyday functions. Algorithms, or "encoded procedures," are the building blocks of the data that make up many of the systems intended to automatically facilitate the tasks of government institutions (Gillespie 2014, p. 1). As much as these institutions' intentions for using automated technology is simply to perform more efficiently and effectively, algorithms that are being used within the automated technologies can have unintended consequences. One such unintended consequence is bias, especially for those algorithms processing data about people. Although the biases embedded within algorithms may not be intentionally or immediately known or used by government institutions, the consequences of poorly constructed algorithms could lead to the infringement of citizens' individual rights. Given the severity of such infringement, it is important that governments are held accountable for algorithmic development. Algorithmic accountability can take the form of regulation amongst the industries responsible for algorithmic formulation by governing institutions.

Efficient algorithms have become necessary for government institutions because they are able to analyze vast amounts of information. Not only is a good algorithm expected to be efficient, it must always give correct answers. So how is efficiency, in terms of algorithms, measured? (Khan Academy, 2018). Considering resource usage, algorith-

mic efficiency is measured by the time it takes to produce a correct answer. The algorithm itself is meaningless unless it is placed within a database. This is because the algorithm itself is limited. It does not make decisions but rather serves as a tool to provide information to then help make decisions. The information that is excluded or included from a database is a human construct. The data that makes up an algorithm is managed in a manner that, according to Gillespie (2014, p. 6), is "reminiscent of the 20th century debates about the ways choices made by commercial media about who is systematically left out and what categories of speech simply don't qualify can shape the diversity and character of public discourse." Although algorithms are thought as being computer-generated, they are in essence influenced by human application.

For example, Mike Ananny revealed that upon installing Grindr, an application for gay, bisexual, and curious men looking for relationships, friendships, and socialization, he immediately noticed an application called "Sex Offender Search" to help "find sex offenders near you…so you can keep your family safe" in the list of "related" applications (2011). Was this a simple coincidence? According to Ananny, it is Android Marketplace that designs and maintains the application store, which puts "related applications" under their purview. But since their algorithms are private property, the public cannot know how Android Marketplace makes such relationships among data. He questions whether "this was an editorial decision made by a human curator of the Marketplace who thought the two applications were somehow related?" (2011). He concludes by saying that the choice to relate these two applications says more about the curator and the algorithm, and less about the applications (Ananny, 2011). While this is an example of pri-

vate company use of technology, it does illustrate how human-constructed patterns of inclusion can predetermine the output of algorithms. Thus, these same patterns of inclusion that could potentially violate individual rights if biases are not addressed or avoided in the creation of algorithms used by government.

This literature review seeks to explain how algorithms are currently being utilized by American governing bodies. It will provide an overview of the current use of algorithms by government and will define algorithmic accountability, an important concept because it allows governments to be able to "promote the vast benefits of algorithmic decision-making" while still being able to ensure individuals' rights are not violated (New and Castro, 2018, p. 2). The importance of accountability in American government will also be referenced in order to further understand the importance of algorithmic accountability. Following this, two important individual rights are introduced as those that algorithms, if unchecked, could potentially infringe upon: privacy and anti-discrimination. Privacy is protected by the 4th Amendment, while anti-discrimination is protected by various important precedents, including the Constitution's Equal Protections Clause, the Civil Rights Act of 1964, and the Equal Employment Opportunity Act of 1972.

Government Use of Algorithms

How do governments currently use algorithms? Algorithms Tip, a project at Northwestern University Computational Journalism Lab, is a compilation of algorithms that have been found on government websites. The purpose of this project is to "predict how algorithms are being referred to across the government agencies" (Diakopoulos et

al., 2018). The algorithms listed on the project website include only government algorithms that "are either actively being used in government operations or are being enforced by the government to assist third-party actions" (Diakopoulos et al., 2018). Moreover, the algorithms included are classified into "newsworthiness," meaning those on the list are perceived as being of great public interest. Most of the algorithms listed belong to the federal government. For algorithms provided, the Project includes the following information:

- the name of the algorithm
- a description of what the algorithm intends
- the possible repressions the algorithm may have on the general population
- the government field or jurisdiction the algorithm belongs to
- the government level
- the government agency it belongs to
- whether the algorithm was developed by a contractor or is owned by a certain company
- who created the algorithm
- the date it was launched or updated
- if the algorithm is being actively used, endorsed for use, or could potentially be used
- whether the algorithm is computational or uses non-software calculations, and
- the URL of the algorithm documentation.

See table 1 for some examples.

Table 1

Examples of Government Algorithms

Name of algorithm	Description	Importance	Topic	Agency	Creator/Vendor	Date	Adoption Stage	Computational
Automated detection of improper data	Detects inappropriately collected data	Violates privacy	Cybersecurity/privacy	NSA	NSA	Oct 2014	Active use	Yes
Automated Targeting System	Target, identify, prevent terrorist entry	Inappropriate flagging	Safety	DHS	NA	NA	NA	NA
Federal Merit Promotion Program	Promotion	Inappropriate inclusion/ exclusion	Personnel	MSPB	MSPB	Mar 2013	Active use	No
Framingham Risk Equation	Determine cardiovascular risk	Cause poor treatment/ no treatment	Health	DHHS	NA	NA	NA	NA
Gang Graffiti Automatic Recognition and Interpretation	Identify/ Interpret gang Graffiti/tattoos	Miss signals/ Mistakenly identify	Safety	DHS	Purdue University	Jan 2014	Endorsed for use	Yes
Major Hazard Risk Assessment	Determine safety in mines	Failure to flag dangers	Safety	CDC	CDC	Oct 2008	Active use	No
National Flood Insurance Program Community Rating System	Provide incentives for flood protection development	Unfairly exclude/benefit certain groups	Safety/ Management	FEMA	FEMA	Aug 2011	Active use	Yes

Information in this table is drawn from Algorithm Tips (Diakopoulos et al., 2018)

Table 1 includes seven algorithms presented in Diakopoulos' Algorithm Tips project. Most algorithms presented in Algorithms Tips, and all seven algorithms presented in the table, are overseen by the federal U.S. government. An algorithm, according to Algorithms Tip is considered computational if it uses software to draw up calculations (Diakopoulos et al., 2018). Diakopoulos and his team (2018) consider the importance of an algorithm to be its potential positive and negative effects on the population; the importance is included for all seven examples in this table. For example, Automated Targeting System, as its title suggests, is an algorithm meant to target and identify potential terrorists and prevent their entry into the United States, but this algorithm risks inappropriate flagging. This means that the algorithm could mistakenly, because of a person's name, nationality, or religion, discriminate against innocent individuals. Moreover, privacy could be violated given the information being collected about the person because prediction, when using such algorithms, replaces the need for proof.

Privacy, as well as discrimination, is an issue that faces the new age of big data, but as Kerr and Earle (2013) suggest, it is less about the data itself but more about "big data's power to enable a dangerous new philosophy of preemption." There is no accepted definition for big data, but Keith Gordon considers that there are five characteristics that, when combined, can determine whether data is indeed 'big data':

1. Volume – where the amount of data to be stored and analyzed is sufficiently large so as to require special considerations.
2. Variety – where the data consists of multiple types of data potentially from multiple sources; here we need to consider structured data held in tables or objects for which the metadata is well defined, semi-structured data held as documents or similar where the metadata is contained internally (for example XML documents), or unstructured data which can be photographs, video, or any other form of binary data.
3. Velocity – where the data is produced at high rates and operating on 'stale' data is not valuable.

4. Value – where the data has perceived or quantifiable benefit to the enterprise or organization using it.
5. Veracity – where the correctness of the data can be assessed (Gordon, 2013, p. 12)

In order to better understand the purpose that big data serves, understanding its predictive nature is important. Most of big data's predictive characteristics fall under three categories: consequential, preferential, and preemptive. Predictions include the formulaic use of zetabytes of data to anticipate everything from consumer preferences and customer creditworthiness to fraud detection, health risks, and crime prevention. Through the predictive power of these algorithms, big data promises opportunities like never before to anticipate future needs and concerns, plan strategically, avoid loss, and manage risk" (Kerr and Earle, 2018). Consequential predictions "attempt to anticipate the likely consequences of a person's action," and "to allow individuals to eschew risk by choosing future courses of action that best align with their own self-interest, forestalling unfavorable outcomes (Kerr and Earle, 2018). A consequential prediction could be made by a doctor when he makes a diagnosis, a lawyer when she predicts what a client's verdict may be, or other professionals for a profit. Preferential predictions are systems that are focused on predicting what clients find interesting in order to sells goods or services. An example of this is the predictive nature of Netflix or Amazon's "suggested for you" information. Most of big data nowadays is focused on preferential prediction. Preemptive predictions, such as analytical systems that determine the likelihood of recidivism of a detainee, as Kerr and Earle (2018) suggest, aim to "assess the likely consequences of allowing or disallowing a person to act in a certain way." Being able to understand what each type of prediction consists of allows for the location of the potential threats of big data.

Avoiding risks is the most important factor of big data and as such, corporations, governments, and others use big data to stall activities, but "often, this will be done with little or no transparency or accountability" (Kerr and Earle, 2018). Big data in turn becomes a potential violation against privacy, due process, and discrimination because there are decisions being made about individuals without their knowledge. Thus, although big data may provide efficiency, utility, profit, and pleasure, "there is wisdom in setting boundaries around the kinds of assumptions that can and cannot be made about people" (Kerr and Earle, 2018). Although it may be most efficient to use big data to perform certain tasks in this fast-paced world, it is important to also take notice of the potential risks.

Government use of algorithms and big data to facilitate and expedite service to their citizens. Governments are able to use e-government to allow direct contact with their constituents, but there are risks when using this technology. Biased algorithms, if undetected, have the potential to discriminate against certain types of people. As we see on the examples provide by Algorithm Tips (2018), the government's vast use of algorithms could prove problematic and counterproductive if no party takes responsibility or accountability for the negative consequences brought on by faulty algorithms. The use of inequitable algorithms could further strain racial relationships with government entities, thus algorithmic accountability would allow for impartiality as well as fairness when utilizing algorithms.

The Importance of Algorithmic Accountability

The fear that algorithms can potentially create risky situations has led to their widespread scrutiny. Algorithmic "black box" is the idea that algorithmic decision-mak-

ing is composed of "extraordinarily complex underlying decision models involving millions of data points and thousands of lines of codes" that then become impossible to discern because of their vast complexity (New and Castro, 2018, p. 5). This complex relationship among inputs and outputs, without the ability to have tangible knowledge of the construction and connection of their relationships, essentially becomes the issue at the center of the controversy. The issue with algorithms' complexity is that it "creates opportunities for bias to inadvertently influence algorithms" (New and Castro, 2018, p. 5). This bias may occur because "the data that algorithms train on can be flawed" (New and Castro, 2018, p. 5).

A simple example illustrating a potential flaw that can influence an algorithm to perform biasedly against a certain population can be something that may go widely overlooked: diversity among the developers of the algorithm. As New and Castro (2018, p. 5) explain, that "the lack of diversity in the developer community creates the risk of homogenous developer teams failing to consider how their own unconscious biases may influence their work, such as not recognizing their training data as not being representative." They reiterate that although this type of scenario is realistic, it does not necessarily mean that, even if there is no diversity in the world of data development, these kinds of biased outcomes cannot be avoided. Accounting for the potential risks and training algorithms to identify biases beforehand can be achieved regardless of who is behind their creation (New and Castro, 2018, p. 6).

There also exists the possibility that government entities could easily hide behind faulty algorithms and "use algorithmic decision-making as a cover to deliberately exploit, discriminate or otherwise act unethically," whether it be to cut government spending by

manipulating algorithms to erroneously manipulate welfare eligibility or to cover up for negligent judicial outcomes (New and Castro, 2018, p.6). Governments using algorithms for decision-making purposes also sheds light on, as the authors suggest, another important characteristic of algorithms that poses a threat to citizens' individual rights: scalability.

Scalability refers to the idea that algorithms are becoming more widespread. The scalability of algorithms "poses a challenge" because of "[an algorithm's] capacity to make a large number of decisions significantly faster than humans" (New and Castro, 2018, p. 6). Different sectors are using algorithms for their daily activities on a regular basis, which in turn increases the probability of an algorithm being used inappropriately on a large scale. An example that illustrates the potential catastrophic impact that using algorithms may leave on individuals is the private sector's banking industry. The loan industry has seen a dramatic increase in the use of algorithms because "banks could use algorithms to dramatically shorten the time it takes to evaluate applicants while reducing operating costs, and then pass those savings on to borrowers in the form of lower interest rates," but if for some reason those algorithms turn out to be flawed, the potential negative impact is too massive because a poorly formulated algorithm could misevaluate hundreds of loan applications at a bank branch which would cause the bank harm on a large scale (New and Castro, 2018, p. 7). If algorithmic accountability is to be regulated by the government, it is important that regulators consider the complexity and scalability that algorithms entail.

Government Accountability

Algorithmic accountability is important in its own right, but especially related to government, as accountability is an important value in the United States government system. The United States functions as a democracy. The government is "representative because [representatives] are elected: if elections are freely contested, if participation is widespread, and if citizens enjoy political liberties, then governments will act in the best interest of the people" (Przeworski et. al., 1999, p. 29). Thus, citizens living in a democratic society "deserve to hold governments responsible for the results of their past actions" (Przeworski et. al., 1999, p. 29). This definition of accountability implies that once an elected official is placed in a position of public service, it is his or her responsibility to uphold and to serve in the public's best interest.

Accountability, according to Romzek and Dubnick (1987, p. 228), "involves the means by which public agencies and their workers manage the diverse expectations generated within and outside the organization." In terms of matters related to public administration, there are four varieties of public accountability each with varying levels of the following two factors, as suggested by Romzek and Dubnick (1987, p. 228) : "(1) whether the ability to define and control expectation sis held by some specified entity outside the agency and (2) the degree of control that entity is given over defining those agency's expectations." The idea that there must be some sort of management of an agency's means of accountability suggests that there must also be an "authoritative source of control" whether internal or external (Romzek and Dubnick, 1987, p. 228). The second criteria when determining the source of agency control is the degree to which an agency has control (high or low).

Given those dimensions, Romzek and Dubnick deduce that there are four types of accountability systems with the realm of public administration: (1) bureaucratic accountability system which focuses on frequently managing the people at the top, (2) legal accountability which involves the relationship between an external controlling party imposing a constant level of control upon an agency, (3) professional accountability suggests that the placement of control is given to the employees of the agency, and finally (4) political accountability which simply implies a relationship between an elected official and his or her constituents (external (Romzek and Dubnick, 1987, pps. 228-229).) Thus, regarding algorithmic accountability and individual rights and considering the political accountability system, elected officials and administrators should be responsive to the needs of his or her constituents. Accountability, in a democratic society, is ultimately held by the citizens.

Individual Rights

Algorithms, which are intended to facilitate efficiency and effectiveness, can have unintended consequences. These consequences could include the violation of individual rights as protected by the American Constitution and other laws. Two rights in particular that may be impacted by government use of algorithms are privacy and anti-discrimination. Privacy, protected by the 4th amendment of the Constitution and other laws, relates to the potential intrusion into and dissemination of an individual's personal data into a public sphere for an institution's personal gain. Additionally, the outputs created by algorithms or government programs that make use of data obtained and used by algorithms could discriminate against individuals, despite anti-discrimination protections offered by

the Constitution and Civil Rights Law. Below, I address both of these rights and the potential for infringement by algorithms.

Privacy

The Fourth Amendment to the Constitution states that it is "the right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized" (Cornell Law School, 2018). The main provision of the Fourth Amendment is to protect people from unreasonable privacy intrusion by the government alone. The notion of privacy, although protected by the Constitution, has been a point of contention throughout history, undergoing various changes.

In 1961, former Dean of Harvard Law School, Roscoe Pound, wrote about the history to a right to privacy as we know it today. He noted that Samuel Warren and Louis Brandeis, both attorneys, were the first to seriously propose the idea of privacy as a right in *The Right to Privacy* published in 1890 by the Harvard Law Review (Pound, 1961, p. 36). Up until this proposal, invasion of privacy was only deemed as such if a physical attack had existed by way of "swords, knives, and stones," on a person or their property without legitimate reason of threat, so when the two attorneys suggested that privacy was something intangible that could not physically be seen, it was considered a radical proposition (Pound, 1961, p. 36). Previously, if force and arms were not utilized, privacy had not been infringed upon. Warren and Brandeis advocated for the "right of one who has remained a private individual, to prevent his public portraiture… from pen portraiture,

from a discussion by the press of one's private affairs…" (Warren and Brandeis, 1890, p. 36). Although a literal interpretation of this statement may seem to solely refer to the protection of privacy rights against the media, the idea itself was revolutionary for the times. Even though this new idea was not well received, Warren and Brandeis's article was cited about a year later in a case of a British doctor whose name was used as a reference in an advertisement recommending a remedy for catarrh (Pounds, 1961, p. 37).

About ten years after its publication, *The Right to Privacy* was yet again referenced in a case brought to the highest court of New York, in which a complaint was filed against Franklin Mills Company for allegedly using the portrait of the plaintiff on over 25,000 prints of an advertisement without her consent (Pound, 1961, pps. 37-38). According to Pound, two questions were posed: "(1) Did the complaint state a cause of action at law? (2) Did it state a cause of action in equity?" (1961, p. 38). The Court, in a 4 to 3 decision decided that neither of those stipulations were met. In addition, Chief Judge Parker, upon finding no support for privacy as a right in previous cases, decided the idea was invalid. (Pound, 1961, p. 38). Chief Judge Parker referenced the following words from Justice Lumpkin of Georgia to support his final decision:

> The law protects the person and the purse…The body, reputation, and property of the citizen are not to be invaded without responsibility in damages to the sufferer… There are too many moral obligations, too delicate and subtle to be enforced in the rude way of giving money compensation for the violation. Perhaps the feelings find as full protection as it is possible to given in moral law and a responsive public opinion. The civil law is a potential business system, dealing with

> what is tangible, and does not undertake to redress psychological injuries (Pound, 1961, p. 38).

Dissenting Judge Gray argued the following:

> This position is, to me, an inconceivable one that these defendants may, unauthorized, use the likeness of this young woman upon their advertisement, as a method of attracting widespread public attention to their wares, and that she must submit to the mortifying notoriety, without right to invoke the exercise of the preventative power of a court of equity (Pound, 1961, p. 39)

Upon reading both sides of the divided court, it is evident that, although the decision still favored privacy as a tangible entity, the courts were slowly moving towards adopting the idea of privacy as Warren and Brandeis had previously defined it.

As Pound continues formulating his historical timeline, he notes the real change occurred during *Pavesich v. New England Life Insurance Company* in 1904 (1961, p. 39). Much like the New York case previously mentioned, the plaintiff's photo was used without permission on an advertisement that also included statements falsely attributed to him. The plaintiff, an artist by trade, found the publication to be "peculiarly offensive" (Pound, 1961, p. 39). The court in the state of Georgia used the New York case as a reference but reversed its ruling. In thirty pages, the court agreed with Judge Gray's dissent and disagreed with Chief Judge Parker's argument meaning that the court concluded that privacy did not need to be tangible in order to be violated (Pound, 1961, p. 39).

Although the Fourth Amendment, when read literally, solely protects property, it was not until the ruling in *Katz vs United States* (1967) that the Supreme Court agreed

that privacy could be protected under said amendment (Little, 1981, p. 313). The ruling on Katz suggested a two-point test to determine whether a privacy claim is protected under the Fourth Amendment. The first is that "the expectation must be an 'actual' one, subjectively held, by the person affected by the search," and the second expectation is "one that society is prepared to recognize as reasonable" (Little, 1981, p. 313). Although an exact definition for privacy has not been agreed upon, it is understood that privacy consists of two criteria: "the ability to keep personal information unknown to others and to keep oneself separate from interaction with others" (Little, 1981, p. 329). Thus, upon reading what the Fourth Amendment states, "secrecy and solitude" can be interpreted using the definition of privacy provided, and consequently, courts must uphold government institutions to this idea of privacy (Little, 1981, p. 315). Moreover, there still exists the idea of a "lesser or a greater" importance in terms of privacy interests. This practice would in turn contradict the idea of protecting privacy, and it would seem that a "privacy hierarchy" would be "unfair and unworkable in practice" (Little, 1981, p. 331). This type of ranking would be subject to an individual's personal perception of privacy.

In addition to the Fourth Amendment, the Privacy Act of 1974 further protected the rights of individuals against the unlawful publication of a person's privacy without proper permission. According to the Department of Justice (DOJ) Office of Privacy and Civil Liberties, the Privacy Act of 1974, "establishes a code of fair information practices that governs the collection maintenance, use, and dissemination of information about individuals that is maintained in systems of records by federal agencies" (DOJ, 2015) Moreover, it requires institutions to give the Federal Registrar a public notice of records.

The Privacy Act also "prohibits the disclosure of a record about an individual from a system of records absent the written consent of the individual" (DOJ, 2015). The right to privacy, as protected by the Privacy Act, protects both the tangible and the intangible aspects of privacy.

The right to privacy remains an open-ended concept that can be argued or utilized for or against individuals. Many of the same issues that were applicable to the time during which Pound wrote about privacy still apply today. This excerpt from the article written by Pound almost sixty years ago very accurately applies to the potential privacy issues that the widespread use of algorithms may cause:

> The right of privacy is a modern demand growing out of the condition of life in the crowded communities of today. Publicity as to private matters of purely personal concern is an injury to personality in a time when modern means of gathering and transmitting news makes everyone a next-door neighbor of everyone else whether he will or not. Indeed, impairment of the peace and comfort of the individual may produce suffering more acute than that brought about by a mere bodily injury. But since the injury is mental and subjective there are difficulties in securing the interest of the person whose privacy is invaded…But the aggression…to sacrifice private feelings to their individual pain have been calling upon the law to do more by way to securing the individual interest than merely take incidental account of infringement of it (Pound, 1961, p. 37).

Although technology since the 1960s has evolved dramatically, the messages relayed by the statements above remain imperative today. The use of algorithms by public and pri-

vate entities will now affect citizens' everyday life at alarmingly fast rates due to the efficiency that cab ne obtained by using algorithms. Thus, it is in the interest of governments to be able to adapt current technologies and their implications to the long-established ideas of privacy in order to avoid rampant infringement on both of the aforementioned individual rights.

Anti-Discrimination

Along with the potential for algorithms' use by the government to violate the privacy of citizens, another individual right that could be infringed by the misuse of algorithms is the right to fair, anti-discriminatory treatment by government institutions. This individual right is protected under several federal policies and constitutional amendments that protect citizens, including due process and equal protection under the Fourteenth Amendment, the Civil Rights Act of 1964 (specifically, Title VII), and the Equal Employment Opportunity Act of 1972.

Constitutional right to equal treatment. The Fourteenth Amendment mandates the protection of individuals' rights under the Equal Protection Clause. The first section of this amendment states:

> All persons born or naturalized in the United States and subject to the jurisdiction thereof, are citizens of the United States and of the state wherein the reside. No state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any state deprive any person of life,

> liberty, or property, without due process of law, nor deny to any person within its jurisdiction the equal protection of the laws (US Court. Amend. XIV, sec. 3).

According to Cornell Law School, equal protection means that governments cannot deny people protection under their laws and that each person should be treated equally under the already-established conditions (Cornell Law School, 2018). Equal protection ensures that individuals of a different race, nationality, or religious affiliation are given the same treatment and the same protection under the laws of the land. If a government institution were to use an automated system or algorithms that produced discriminatory outputs, that institution may be violating the Equal Protection Clause of the Fourteenth Amendment. A way in which the government could avoid such instances is by holding the institutions creating or using said algorithms accountable for the damages precipitated.

Some scholars argue that the Constitution of the United States is outdated and that it would be beneficial to update it in accordance to the sociopolitical culture of the country, but researchers such as Chiang would argue that the true issue is what is "empirically verifiable" and that which has truly remained unchanged: racism. She states, "first, that African Americans are disproportionately affected by laws that burden the poor or socially disadvantaged because they are disproportionately poor and socially disadvantaged; second…the racism that remains…is of a subconscious variety…; and third, that racial inequalities persist in America…" (2014, p. 842). The racism of a 'subconscious variety' is the racism that can potentially slip into algorithms that provide services to minorities, and although privacy and anti-discrimination policies offer some constitutional protection, the idea of algorithmic accountability should be considered in order to avoid infringing individuals' rights.

Title VII of the Civil Rights Act 1964. The Civil Rights Act of 1964 is the dominant anti-discrimination law in the United States and outlawed segregation in businesses and in public accommodations such as swimming pools, libraries, and schools, banned discrimination in the workplace, addressed voting rights, and nondiscrimination within in federally assisted programs on the basis of ethnicity, color, religion, sex, or national origin (National Archive and Records Administration). Additionally, Title VII of the Civil Rights Act of 1964 prohibits discrimination in the workplace based on race, color, nationality, sex, or religion. Sex in this context also includes pregnancy, childbirth, or any related medical issues, and according to Cornell Law School (2018), "makes it illegal for employers to discriminate in relation to hiring, discharging, compensating, or providing terms, conditions, and privileges of employment."

In case of suspected violation on the basis of Title VII of the Civil Rights Act of 1964, according to Richard Posner in the University of Pennsylvania Law Review (1987, p. 517), a plaintiff could make two cases. The first is the "disparate treatment" approach. Under this approach a person must prove intentional discrimination, but this is difficult to prove because no logically-thinking employer "will admit or lave a paper record showing that it has refused to hire or has fired a worker because of a worker's race" (Posner, 1987, p.518). Usually, this approach will not result in nothing more than a believable reason from the employer for not hiring or firing. The second approach is "disparate impact" (Posner, 1987, p. 518). Under this approach, "if a firm uses a screening device such as an aptitude test…that has the effect of excluding a disproportionate

number of blacks, the device is unlawful unless the frim can show a strong business justification for it…the crux of the problem is identifying disproportionate exclusion" (Posner, 1987, p. 518). Taking Posner's conclusions to algorithms, it could be argued that disparate impact could occur, and even though discriminatory intent may not exist in the creation of algorithms, they may produce discriminatory outcomes.

Equal Employment Opportunity Act of 1972. The Equal Employment Opportunity Act of 1972 amended Title VII of the Civil Rights Act of 1964 to give the Equal Employment Opportunity Commission the authority to "conduct its own enforcement litigation" (Cornell Law School, 2018). This was the fourth attempt by Congress improve the provisions of Title VII and it resulted in the following:

- EEOC has litigation authority. If the agency cannot secure an acceptable conciliation agreement, it has the option of suing nongovernment respondents -- employers, unions, and employment agencies.
- Educational institutions are subject to Title VII. Congress found that discrimination against minorities and women in the field of education was just as pervasive as discrimination in any other area of employment.
- State and local governments are no longer exempt from Title VII. Removal of this exemption results in 10 million more employees being immediately added to Title VII's coverage.
- The Federal Government is subject to Title VII. Federal executive agencies and defined units of the other branches must make all personnel actions free from discrimination based on race, color, sex, religion or national origin.
- The number of employers covered by Title VII is increased by reducing the number of employees (from 25 to 15) needed for an employer to be covered by the Act.
- Charging parties have a longer period of time to file their charges, 180 or 300 days rather than 90 or 210 days. Additionally, charging parties now have 90 days rather than 30 days to file a lawsuit after EEOC has informed them that it is no longer working on their charge. This extension of time affords charging parties a better chance to find a lawyer if they wish to pursue their charges in court. (EECO "Milestones:1972").

The Equal Employment Opportunity Act of 1972 as put by Congress itself after its enactment was meant “to correct the defects in its own legislation. The promises of equal job opportunity made in 1964 must [now] be made realities…" (EECO “Milestones: 1972”).

Government institutions have the responsibility of making sure that the new technologies being used throughout their systems are efficient, effective, and fair. Although anti-discrimination is protected under various laws, the ever-changing nature and consequences of technologies may not be accounted for by laws written before the introduction of said technologies, so it is in the best interest is in the best interest of government institutions to protect and advocate for nondiscriminatory practices within the use of technology.

Algorithmic Accountability, Privacy, and Anti-Discrimination

Privacy and anti-discrimination are individual rights protected by the Constitution and other federal laws, but as technology advances, particularly the usage of algorithms, the lines of privacy and anti-discrimination become more and more blurred. Developers of algorithms may be creating unintended consequences, such as violating privacy or discriminating against individuals, either through implicit bias or poorly designed programs. Moreover, the use of algorithms in everyday tasks is vast. Government agencies are particularly keen of the use of algorithms to facilitate their tasks. By utilizing algorithms, government institutions intend to more efficiently and effectively serve their constituents. An example of this is the automation, through the use of algorithms, of personnel management.

Upon typing "personnel" in the search bar on the Algorithm Tips website, 19 different algorithms result. Of these, many algorithms are used by the federal government for managing personnel whether it be hiring, personnel evaluations, appraisals, or promotions. The Office of Personnel Management, Internal Revenue Service, Central Intelligence Agency, Department of Labor, Department of the Air Force, Merit Systems Protection Board, United States Postal Service, Department of Agriculture, Nuclear Regulatory Commission, Department of the Interior, Department of Defense all use algorithms in which personnel could potentially be negatively impacted by a poorly constructed algorithm. The importance, or "the possible repressions the algorithm may have on the general population," of the algorithm is listed as "might be unfair in excluding appropriate candidates or generate negative impact by including inappropriate candidates for government employment" by Algorithm Tips (Diakopoulos et. al., 2018). Although these are only 19 instances of algorithms the government uses, it is worth noting that the more the government uses algorithms, the more likely it is for something to go awry. Thus, the natural response for avoiding the transgression of rights, such as privacy and antidiscrimination, is to hold algorithm creators and vendors accountable for creating appropriate algorithms before publishing their products.

In personnel management, where privacy and anti-discrimination is highly protected, consequences that result from algorithms are important. Borry and Getha-Taylor, who studied the impact of automation on the public sector workforce, write that technology for the purposes of efficiency is "particularly problematic when utilized in the public sector, where commitment to equal employment and nondiscrimination is essential" (Borry and Getha-Taylor, 2018, p. 2). This speaks also to the use of algorithms: the more

widespread the use of these algorithms, the higher the risk of infringing rights protected by law and the Constitution.

CHAPTER 3

CASE STUDY ANALYSIS

To address the three research questions presented in this thesis, case study analysis is used because it allows for "extensive and in-depth description of…social phenomena" (Yin, 2017, p. 4). Additionally, case study research has five components: a case study's questions; its propositions, if available; its cases; a common thread linking the data to the propositions; and an explanation for how conclusions or interpretations were made (Yin, 2017, p. 27). The questions of this research are stated in Chapter 1, while theoretical linkages are explored in Chapter 2. While not exactly "propositions," these linkages set the framework for analyzing the cases presented in this study.

The two cases analyzed here are "Recidivism" and the second case analyzed is titled "Graffiti." The case for the recidivism algorithm is explored using secondary sources that report on an algorithm COMPAS that is endorsed by the Department of Homeland Security. The second case study was found after a search on Algorithm Tips. This particular algorithm, Gang Graffiti Automatic Recognition and Interpretation, deals with the recognition of graffiti and is used by various police departments throughout the US particularly Indiana. Secondary sources found on the ProPublica publication inform this case. Both cases are explained and analyzed through the lenses of privacy and antidiscrimination, as well as algorithmic accountability.

The two cases presented in this paper were chosen because both deal with algorithms that could have unintended consequences that could lead to an infringement upon

the two individual rights that are defined in the literature review: privacy and equal protection. Both cases, and the application of these individual rights, emphasize the need for algorithmic accountability by way of regulation.

In the remainder of this chapter, I present the two cases. For each case, I will describe the algorithm, explain the algorithm's intended outcomes, and its untended consequences, potential or real. Subsequently, I will explain how each algorithm has or could potentially transgress the individual right to privacy and the right to anti-discriminatory treatment. Finally, for each algorithm, I will suggest how algorithmic accountability could help avoid such violations.

Case Study 1: Recidivism

According to "Machine Bias," a study published in ProPublica, in the year 2014, 18-year-old Brisha Borden was arrested for the petty theft of a boy's bicycle in Fort Lauderdale, Florida. The authors, Julia Angwin and her colleagues, compared Borden's crime and criminal record to that of 41-year-old Vernon Prater, which was extensive in comparison. Prater had been convicted of armed robbery and attempted armed robbery prior to being arrested for stealing from a local Home Depot. He had served time for his prior armed robbery charge, but when both Prater and Borden were booked into jail, the computer system predicted a score that placed Borden at a higher likelihood than Prater of committing a crime again. Borden, who is black, was rated at a higher risk, while Prater, white, was rated low-risk (Angwin et. al., 2016).

Two years after the incidents, the system was found to be wrong. The algorithm which was designed to predict the possibility that a detainee would recommit a crime if

let out earlier than necessary had given the likelihood rate of recidivism was incorrect. Two years later and out of jail, Borden had not committed any more crimes; conversely, Prater is currently serving an eight-year sentence for breaking into a warehouse and stealing thousands of dollars' worth of electronics (Angwin et. al., 2016). So, what are these risk assessments and what do they measure?

Courtrooms across the nation are increasingly utilizing risk assessments to determine who can be freed from jail, how much bonds should be set at, and other general decisions about a suspect or criminal's freedom. States including Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virginia, Washington, and Wisconsin give these assessments to judges at the time of sentencing (Angwin et. al., 2016). Moreover, this risk assessment is made in conjunction to whatever a criminal's future rehabilitation necessities may be. Although the Justice Department's National Institute of Corrections suggested that risk assessments be made at every stage of the criminal process, it was former U.S. Attorney General Eric Holder that suggested that such assessments "might be injecting bias into the courts" (Angwin et. al., 2016). Holder stated that "although these measures were crafted with the best intentions…they inadvertently undermine our efforts to ensure individualized and equal justice," and he concluded that "they may exacerbate unwarranted and unjust disparities that are already too common in our criminal justice system and in our society" (Angwin et. al., 2016).

According to the Angwin and her team (2016), the algorithm that is used to formulate these risk scores was created by a for-profit company that derives its scores by asking 137 questions. Some of the answers to these questions are pulled directly from criminal records while others are the answers provided by the defendants. Some of the

questions that are asked include the following: “Was one of your parents ever sent to jail?” “How many of your friends/acquaintances are taking drugs illegally?” “How often did you get in fights while at school?” (Angwin et. al., 2016). Moreover, defendants must also provide answers either agreeing or disagreeing to whether “a hungry person has the right to steal,” and to whether “if people make me angry or lose my temper, I can be dangerous” (Angwin et. al., 2016).

The risk assessments included in Angwin’s study are part of a database made up of information provided by the Federal Post Conviction Risk Assessment, which according to the Administrative Office of the United States Courts (“Administrative Office”; 2018, p. 2), is “a scientifically-based instrument…to improve the effectiveness and efficiency of post-conviction supervision.” Through “evidence-based” practices, the Administrative Office makes a “conscientious use of the best evidence currently available to inform decisions about the supervision of individuals, as well as the design and delivery of policies and practices, to achieve the maximum, measurable reduction in recidivism” (“Administrative Office”; 2018, p. 2). The government institutions of the US have taken the efficiency and effectiveness of these technological advancements and may be neglecting other potential considerations as a result.

Since the sentencing commission refused to conduct any sort of study for the implications of such risk assessments, ProPublica conducted extensive research to examine the effects that biased algorithms may have on the justice system. The authors of this study obtained the risk assessments of people arrested in Broward County, Florida. These 7,000 assessments were then reviewed to determine, out of the individuals scored, who recidivated within the two-year mark that was used as the benchmark by the creators of

the algorithm. As suspected, the study showed that "only 20 percent of the people predicted to commit violent crimes actually went on to do so" (Angwin et. al., 2016). It also indicated that when all crimes were considered, including misdemeanors, the algorithm for determining recidivism "was somewhat more accurate than a coin flip" (Angwin et. al., 2016).

More alarming than only a 50 percent accuracy rate for predicting recidivism, is the fact that racial disparities were highly apparent. According to ProPublica, "the formula was particularly likely to falsely flag black defendants as future criminals, wrongly labeling them this way at almost twice the rate as white defendants…white defendants were mislabeled as low risk more often than black defendants" (Angwin et. al., 2016). Angwin and her team determined that a defendant's prior crimes or the types of crimes that had been committed had nothing to do with the numbers that the algorithm produced. According to the findings of the study, when isolating race from criminal history, recidivism, age, and gender, the group found that black defendants were "77 percent more likely to be pegged as at higher risk of committing a future violent crime and forty-five percent more likely to be predicted to commit a future crime of any kind" (Angwin et. al., 2016). The numbers presented by Angwin and her team expose a racially-biased algorithm.

Violation of Rights

Upon analyzing the case built by Angwin and her team, it can be said that this algorithm, dealing with recidivism and its less than accurate outcome, could violate both the privacy and the right to anti-discrimination of black inmates. The algorithm placed

Borden at a higher risk of recommitting a crime seemingly because of the color of her skin, when in reality it was Prater, a reoffender, who committed the crimes after being released again. This output by the algorithm was likely an adverse impact.

When examining this algorithm and the resulting events through a lens of privacy as protected by the Constitution, the violations could be many. Because the data the makes up the algorithm is not publicly disclosed information because Northpointe is a private company, it is impossible to know what information is being used to deduce risk scores. According to Angwin et. al. (2018), "defendants rarely have an opportunity to challenge their assessments. The results are usually shared with the defendants' attorney, but the calculations that transformed the underlying data into a score are rarely revealed." Defendants in turn have no access to the information being used against them, and they are not made aware who is sharing their personal information to private companies such as Northpointe. "Risk assessments should be made impermissible unless both parties get to see all the data that go into them…it should be an open, full-court adversarial proceeding" (Angwin et. al., 2018). Moreover, in terms of data collected by government, it is important to point out that much of the information being collected results from the size of digital footprints left on communication networks by individuals and organizations, so individuals who have been in the system in the past will likely have more information stored about them, which in turn can skew data and create an issue of inequity. The individuals involved in these risk assessments should be allowed to access their assessment and know what personal information was utilized and where the information was sourced from.

Although judges are not supposed to reference risk scores to decide length of sentencing, there have been instances when judges have done just that. The Due Process Clause of the Fourteenth Amendment calls for fair and equal treatment when applying the law regardless of race, sexuality, religion, or nationality. Giving the findings of Angwin et. al. (2018) about the recidivism algorithm, it is possible that because of the racially-biased risk scores that the algorithm creates, black defendants are not being granted their right to due process. "Risk scores alone should not determine the sentence of an offender…we don't want to say, this person is a 10…as far as risk, and therefore I'm going to give him the maximum sentence" (Angwin et. al., 2018) The risk scores are placing the black defendants at a disadvantage because judges use the risk scores to determine various aspects of their sentencing process.

Taking Responsibility

The algorithm itself can be said to be discriminatory, either intentionally or by consequence, and by association the government institutions using this algorithm or providing information for this algorithm, can also be contributing to discrimination of black defendants. Anti-discrimination is a right granted by multiple federal policies and constitutional amendments. The Fourteenth Amendment as well as the Civil Rights Act of 1964 and its amendments protect all citizens from discriminatory treatment. Considering the violations brought about by these types of errors in recidivism algorithms, it would be helpful to establish some type of regulatory body to ensure that such mistakes are not made and rights are not violated. Not only is it impactful on the individuals' lives

but keeping inmates for unnecessarily long periods of time can prove costly to governments if it is occurring in large volumes. In addition, regulation may reduce the potential for civil cases brought against the government as a result of these violations by algorithms.

Case Study 2: Graffiti

According to Algorithm Tips (2018), Gang Graffiti Automatic Recognition and Interpretation (GARI) is an algorithm which "helps law enforcement and gang task force officers identify and interpret gang graffiti or tattoo images" by comparing "the image against an image database and provides details about the identity and meaning of the graffiti or tattoo." It is an algorithm endorsed and funded by the Department of Homeland Security (DHS) Science and Technology Directorate as part of the Visual Analytics for Command, control and Interoperability Environments Center of Excellence (VACCINE) at Purdue University, and according to DHS, the algorithm helps users with the following: to determine when a new gang moves into an area, to identify what gangs are active in an area, to target youth who are at risk of gang recruitment, and to prepare for potential outbreaks of gang violence (DHS, n.d.).

The algorithm works by user submissions; users are typically law enforcement officials. A user can take a picture with a cell phone or tablet and submit it to the database linked to GARI by uploading the photo. The application then feeds back information about the graffiti or tattoo image, such as gang affiliation, what the image means, historical information, geographical locations of other similar images. This information is in-

tended to allow law enforcement to "track gang movement, growth, membership, and activity" (DHS, n.d.). Once the picture has been taken, GARI then adds the image to an image database, and along with the picture, it records the date, time, and coordinates where the image was acquired. If a picture is taken and a similar image already exists in the database, GARI returns information to the device on which the image was taken. Since it was instituted in 2012, the algorithm has collected over 8,000 graffiti images.

History Between Graffiti and Law Enforcement

The relationship between law enforcement and graffiti in American has always been contentious (Dickinson, 2008, p. 27). Graffiti, regardless of its subject, has always been criminalized (Dickinson, 2008, p. 27). Even if graffiti is an expression of art, the historical crackdown on graffiti, created a tight-knit subculture that only allowed for graffiti to flourish (Dickinson, 2008, p. 28). For example, in the 1970s, the New York Police Department (NYPD) cracked down on graffiti culture by criminalizing graffiti. Graffiti writers were convicted of crimes, even though there were no city ordinances or laws that condemned the writing of graffiti. Maggie Dickinson writes that "graffiti speaks to…one's race, class gender, ethnicity, nationality, religion…which can determine how one's own creative labor is interpreted" (2008, p. 27). She claims that the manner in which the city responded and continues to respond to graffiti has everything to do with "restructuring the city to save the needs of capital accumulation" (2008, p. p. 28). Because graffiti did not and does not assimilate to the city's restructuring project, it became problematic. In order to cater to the business community, government officials along with the business community and media outlets, came together in order to legislate for legal

penalty in October of 1972. By early 1973, 1,562 arrests had already been made, and although punishment was minimal, graffiti writers were now portrayed as vandals and thugs. This "rhetoric of war" according to Dickinson, "made it virtually impossible for graffiti writers to get across to a wider public the many positive aspects of the subculture" (2008, p. 31).

What positive aspects did these graffiti writers intend? Writers believe that their graffiti was one of the only ways they could communicate. As Dickinson explains (2008, p. 31), "a writer would write someone's name whose style they admired with her own name next to it and he or she would write back," called "third rail mail." The writers, especially the younger ones, felt that their "spatially segregated" communities came together through the use of graffiti because the community of writers had transformed into an open, welcoming community. Despite the city's works, the effort to eliminate graffiti was abandoned, but again in the 1980s, under supervision of Mayor Koch, the mission was continued with more force (Dickinson, 2008, pp. 32-33).

Koch's mayoral run was an unfortunate time for blacks and Latinos in New York (Dickinson, 2008, pp. 32-33). His policies were racially divisive. Under his leadership, four hospitals found within minority neighborhoods were closed down, while all hospitals in predominately white neighborhoods remained open. He also took to cutting the city employees' services and pay. Koch's campaign against graffiti included attack dogs and razor-wired fences around the train yards where writers painted. As Dickinson points out, although the money to campaign against graffiti could have been used to control the issue with muggers on the trains, which posed a real threat, by the 1980s, New York residents had accepted the idea of gentrification of the city, and nothing was done to protect the

lives or rights of the writers (Dickinson, 2008, p. 34). As understood from Dickinson's piece, the issue with graffiti writers at the time was more or less just one of the attacks on people of color living in New York City at a time when gentrification was once again being masked as economic progress. The fight to eliminate all graffiti has continued on through the leadership of Mayor Giuliani, and as of 2002, the NYPD decided to conglomerate the anti-graffiti unit to the Transit Bureau Vandal's unit, creating the Citywide Vandals Task Force because according to the NYPD the view that graffiti can be an art form is "not only puerile, it is misguided as well" (NYPD, n.d.).

Violation of Rights

Since no information regarding the success or failures of the algorithm were provided upon reaching out to the creators of the algorithm, inferences must be drawn. Given the tense history between graffiti and law enforcement, the GARI algorithm can potentially be problematic because of the biases that historically exist about graffiti. The rights of individuals can be infringed by GARI if the individuals collecting information are implicitly biased against graffiti because of the historical misunderstanding that all graffiti is criminal in nature. The privacy of an individual, in terms of the inappropriate dissemination of incorrectly portrayed images can lead to the invasion of privacy of an individual. Law enforcement officials could infer that someone's tattoo or certain graffiti has links or meaning in relation to gangs. The official would take the picture, and the information about the picture would be placed into the database. The information, now in the database and available to everyone who uses it, may depict the person in the wrong light and their image and privacy is now available to all users.

The Privacy Act of 1974 claims that every individual has the right to "fair information practices that governs the collection, maintenance, use and dissemination of information about individuals that is maintained in systems of records by federal agencies" (DOJ, 2015). Taking this portion of the Privacy Act of 1974 into account, every individual has the right "fair information practices" if the information is stored within the records of the federal government. Since GARI is endorsed and funded by the Department of Homeland Security, the improper dissemination of a person's information by GARI would fall under the jurisdiction of the federal government.

Historically, graffiti, as mentioned above by Dickinson (2008), is an ode to an individual's race, class, gender, ethnicity, nationality, religion, and because graffiti has always been criticized regardless of subject, it can be inferred that graffiti and the graffiti writer can become easy targets for discrimination. An aspect about GARI that can be easily scrutinized, does not lie within the manner in which data is added to the database or any of the technological aspects of the algorithm, but the fact that the data that is entered into the system lies on the opinion of a law enforcement official. If a law enforcement official has implicitly biased tendencies against the practice of graffiti, it can be inferred that that law enforcement official may feel the need to strictly target graffiti in a negative manner, whether it be graffiti or tattoos, and only send data that could potentially hurt that particular group of people.

Taking responsibility

If a situation like the one referred to above were to continually occur by multiple persons in the law enforcement field, it is possible that the database that feeds back information about images could be greatly biased. Images that are uploaded with the assumption of criminality could plague the algorithm's database and potentially create greatly discriminatory situations. Invasion of privacy, as Pound (1961) explains, is not simply just a physical, tangible act anymore like it was centuries ago, invasion of privacy now entails any sort of infringement upon an individual's peace that may result in mental stress. Moreover, personal biases could greatly affect the manner in which GARI functions. If the database from which the algorithm is drawing its information from is historically biased then the outputs that the algorithm produces will inevitably also be biased, thus it is important to hold every aspect of the algorithms' processes clear of faulty input. The creators of GARI and the government institutions utilizing it should be held accountable for upholding the rights protected under the Constitution and the laws of the United States.

CHAPTER 4

DISCUSSION AND RECOMMENDATIONS

The purpose of this thesis is to better understand how algorithms influence our everyday lives. It is important to understand how they are being used by the government of the United States. After establishing what types of agencies use algorithms, knowing if algorithms are being used appropriately is equally as important. Understanding the degree to which the government holds institutions accountable for potentially biased algorithms that could infringe individual rights is next in question regarding algorithmic use. The following research questions were posed:

1. How do governments in the United States currently use algorithms?
2. What individual rights might algorithms infringe upon?
3. What role does algorithmic accountability play in the United States?

The first question was addressed by searching for algorithms used within government systems. Although many of the algorithms created are considered private property, Algorithm Tips (2018) was able to compile an extensive list of algorithms used or endorsed by the federal government. Upon choosing one algorithm from the list and choosing another algorithm from a study published by ProPublica (Angwin et. al., 2016), two individual rights that could possibly be infringed upon were determined. Upon this determination, it was concluded that indeed the government is responsible for establishing algorithmic accountability in order to avoid the violation of rights.

Taking a closer look at the case study analysis of two algorithms, we see that the first algorithm deals with the consequences of recidivism risk scores of inmates in correctional facilities, and the second algorithm considers the dangers of using a gang graffiti and tattoo automatic recognition system to be able to build a database for the use of law enforcement. While presenting the case and explaining how the algorithm functions, the potential biases that arise are investigated through the lens of two individual rights protected by the Constitution: privacy and antidiscrimination. The case study analysis and the literature review for the two individual rights allowed for the answers to the research questions.

Recommendations: Regulation and Algorithmic Accountability

The changing technological environment in this country is revolutionizing the manner in which government institutions interact with their constituents. Government agencies nowadays aim for efficiency and effectiveness, so in order to keep up with these demands, they have turned to automation, but there are consequences. One of these consequences is the possibility of faulty, biased algorithms that can potentially infringe upon the rights of individuals. Algorithmic accountability would allow government agencies to hold creators and vendors accountable for poorly constructed algorithmic systems. These agencies, vendors, and creators could be held accountable for their algorithms through regulatory implementations.

Regulation would allow the government to have some control over what algorithms are being introduced into the market, and laws could also be set in place to hold

any institution accountable for infringement of individual rights. Regulation on algorithms should not be completely restrictive because it is a new form of technology that is just starting to develop, and not all algorithms or new forms of technology are inherently evil. Algorithms help facilitate everyday lives, so complete regulation would be counterproductive. Algorithmic regulation. Although, regulation can be a form for government to hold public and private entities responsible for the consequences (good or bad) of their algorithms, it is unclear what type of regulation would work best. Thus, future research could explore the appropriate methods for regulating algorithms and their use. Some experts believe a "light-touch" approach would work best, but little research has been done about what exactly that type of regulation would entail (New and Castro, 2018, p. 2). Finally, because life now revolves around technology and because technological advancements will only continue to multiply and proliferate, government institutions must learn to stay up to date. Agarwal (2018, p. 5) suggests that "the pace of change outside our public policy institutions is faster than the pace of change within," and she points out that leaders of the artificial intelligence world "are calling for the governments to be more engaged and more active in shaping policy around change." Algorithms are part of that change and policy must allow these new technologies to breathe and grow, but also help set guidelines that will protect the rights of citizens.

Limitations

This study is not without limitations. Considering that only two case studies were investigated through the lens of only two individual rights, the argument that algorithmic accountability is necessary to avoid the infringement of rights could have been further

supported if more case studies were analyzed. Future research on this subject could address other ethical considerations or other individual rights. Moreover, deciding how algorithms should be regulated or formulating a regulatory system that can protect individual rights without hindering technological advancements would be another subject for future research.

Conclusion

Understanding algorithms and how they influence everyday lives allows citizens to be wary of how institutions, both private and public, utilize their information in order to efficiently and effectively serve them. Although the intentions of these institutions may not be purposefully biased, the risk of running on unregulated algorithms enables the possibility of infringing individual rights such as privacy and anti-discrimination. Thus, in order to reach the conclusion that indeed algorithms must be regulated in whichever way, government must enforce accountability for the consequences imposed by biasedly constructed algorithms. Setting guidelines for the appropriate use of algorithms will allow this new technology to prosper and grow while still protecting the individual rights awarded by the Constitution.

REFERENCES

Administrative Office of the United States Courts, A. O. o. t. U. (2011). An overview of the federal post-conviction risk assessment.

Agarwal, P. K. (2018). Public Administration Challenges in the World of AI and Bots. *Public administration review, 78*(6), 917-921.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *ProPublica, May 23*.

Ananny, M. (2011, April 15). The Curious Connection between Apps for Gay Men and Sex Offenders. Retrieved October 22, 2018, from https://www.theatlantic.com/technology/archive/2011/04/the-curious-connection-between-apps-for-gay-men-and-sex-offenders/237340/

Borry, E. L., & Getha-Taylor, H. (2018). Automation in the Public Sector: Efficiency at the Expense of Equity? *Public Integrity*, 1-16.

Caplan, R., Donovan, J., Hanson, L., & Matthews, J. (2018). *Algorithmic accountability: A primer*.

Chiang, E. (2013). The New Racial Justice: Moving Beyond the Equal Protection Clause to Achieve Equal Protection. *Fla. St. UL Rev., 41*, 835.

Cornell Law School. (2014, December 04). Civil Rights Act of 1964. Retrieved October 5, 2018, from https://www.law.cornell.edu/wex/civil_rights_act_of_1964

Cornell Law School. (2017, December 23). Equal Employment Opportunity Commission. Retrieved September 25, 2018, from https://www.law.cornell.edu/wex/equal_employment_opportunity_commission

Cornell Law School. (2018, September 19). Equal Protection. Retrieved September 9, 2018, from https://www.law.cornell.edu/wex/equal_protection

Cornell Law School. (2017, June 05). Fourth Amendment. Retrieved October 19, 2018, from https://www.law.cornell.edu/wex/fourth_amendment

Dennis, M. R. (2006). Proletarian or Promethean? Impacts of automation and program integration on social service workers and their clients. *Journal of contemporary ethnography, 35*(5), 552-582.

Department of Homeland Security. (2018). Enhancing Community Safety: Gang Graffiti Automatic Recognition and Interpretation System. DHS Science and Technology Center of Excellence. Retrieved October 3, 2018, from https://www.dhs.gov/sites/default/files/publications/Enhancing%20Commnity%20SafetyGang%20Graffiti%20Automatic%20Recognition%20and%20Interpretation%20System-GARIJan2014_2.pdf

Department of Justice. (2015, July 17). Privacy Act of 1974. Retrieved August 29, 2018, from https://www.justice.gov/opcl/privacy-act-1974

Diakopoulos, N. (2013). Algorithmic accountability reporting: On the investigation of black boxes. A Tow/Knight Brief. Tow Center for Digital Journalism, Columbia Journalism School. Retrieved August 29, 2018, from http://towcenter.org/algorithmic-accountability-2/

Diakopoulos, N., Trielli, D., & Stark, J. (2018). *Algorithm Tips – Find tips for stories on algorithms*. Retrieved September 12, 2018, from http://algorithmtips.org/

Dickinson, M. (2008). The Making of Space, Race and Place: New York City's War on Graffiti, 1970—the Present. *Critique of Anthropology, 28*(1), 27-45.

EECO. (2018). Milestones: 1972. Retrieved October 30, 2018, from https://www.eeoc.gov/eeoc/history/35th/milestones/1972.html

Elmagarmid, A. K., & McIver, W. J. (2001). The ongoing march toward digital government. *Computer,34*(2), 32-38.

Gillespie, T. (2014). The relevance of algorithms. *Media technologies: Essays on communication materiality, and society, 167*.

Gordon, K. (2013). What is Big Data? *Itnow, 55*(3), 12-13.

House of Commons. (2018, May 15). Algorithms in decision-making. Retrieved November 26, 2018. https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/351.pdf

Jurkiewicz, C. L. (2018). Big Data, Big Concerns: Ethics in the Digital Age. *Public Integrity*, 1-14.

Kerr, I., & Earle, J. (2013). Prediction, preemption, presumption: How big data threatens big picture privacy. *Stan. L. Rev. Online, 66*, 65.

Khan Academy. (2018). What is an algorithm and why should you care? Retrieved October 4, 2018, from https://www.khanacademy.org/computing/computer science/algorithms/intro-toalgorithms/v/what-are-algorithms

Little, R. (1981). Protecting Privacy under the Fourth Amendment. *Yale Law Journal, 91*, 313.

National Archive and Records Administration. (2018, March 21). Civil Rights Act (1964). Retrieved September 7, 2018, from https://www.ourdoc uments.gov/doc.php?flash=false&doc=97

New, J., & Castro, D. (2018, May 21). How Policymakers Can Foster Algorithmic Accountability. Retrieved September 11, 2018, from http://www2.datainnovation.org/2018-algorithmic-accountability.pdf

NYPD (New York Police Department) (n.d.). Retrieved October 15, 2018 www.nyc.gov/html/nypd/html/transportation/vandals

Pardes, A. (2018, October 09). Hey Bullies, Instagram's Niceness Cops Are Comin' For You. *Wired*. Retrieved October 15, 2018, from https://www.wired.com/story/instagram-anti- bullying-algorithm/

Posner, R. A. (1987). The efficiency and the efficacy of Title VII. *University of Penn-Sylvania Law Review, 136*(2), 513-521.

Pound, R. (1961). The Fourteenth Amendment and the Right of Privacy. *W. Res. L. Rev., 13*, 34.

Przeworski, A., Stokes, S. C., & Manin, B. (1999). *Democracy, accountability, and re-presentation* (Vol. 2): Cambridge University Press.

Romzek, B. S., & Dubnick, M. J. (1987). Accountability in the public sector: Lessons from the Challenger tragedy. *Public administration review*, 227-238.

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review*, 193-220.

Yin, R. K. (2017). *Case study research and applications: Design and methods*: Sage publications.

US Court. Amend. XIV, sec. 3.