Administrative Intelligence: Exploring Balanced Human-AI Decision-Making Relationships in Canadian Administrative Contexts

By

Michael Dockstator

A thesis submitted in conformity with the requirements for the degree of Master of Information

Faculty of Information
University of Toronto

Administrative Intelligence: Exploring Balanced Human-AI Decision-Making Relationships in Canadian Administrative Contexts

Michael Dockstator

Master of Information

Faculty of Information
University of Toronto

2019

Abstract

One of the many promising potentials of artificial intelligence (AI) is its ability to improve public decision-making. But, until there is enough confidence for AI systems to act independently, humans will still play a prominent role in making decisions. However, these human-AI decision-making relationships create a new dynamic that, if not properly structured, can end up working against the principles and goals of administrative law. This thesis will focus on the human element in human-AI decision-making relationships and how these dynamics can be structured to promote the goals and principles of public decision-making. From the perspective of Canadian administrative law, the focus of this thesis will be a qualitative case study on three prominent AI decision-making systems – COMPAS, iBorderCtrl, and PredPol – that are already implemented in jurisdictions with analogous administrative law requirements.

Acknowledgments

I would like to take this opportunity to thank all of the family, friends, professors, and mentors for all of their support, thoughts, perspectives, and guidance during this thesis project and throughout my education. First and foremost, there are the members of my thesis committee: Professor Matt Ratto who acted as my supervisor, Professor Chun Wei Choo who was my second reader, and Professor Amy Salyzyn from the University of Ottawa who was my external examiner. I am also deeply appreciative of all the support from my family: my parents, Jennifer and Mark Dockstator, as well as my brother Jacob Dockstator. There are also the many friends that have supported me throughout this project, with special thanks to Jesse Beatson, Calvin Tennakoon, Taylor Robertson, Chelsea Russell, Alex DeSantis-Smith, Adrian Beattie, and Jamie Duncan. And finally, thank you to all of the influential professors and mentors that I have been fortunate enough to work with and learn from: Selma Purac, Leslie Shade, Tim Blackmore, Lesli Bisgould, Camille Labchuck, Aneurin (Nye) Thomas, Ryan Fritsch, Benjamin Alarie, and Andrew Green.

Table of Contents

List of Tables

List of Appendices

1. Introduction

As Artificial Intelligence (AI) programs become increasingly capable, they will hold a corresponding increasing potential for public decision-making. This potential is reflected in the many governments that have expressed an interest in learning about how they can take advantage of AI capabilities. In June 2018, the Canadian Federal government's Treasury Board Secretariat released a Request for Information to learn about where and how it can responsibly use AI programs (Treasury Board Secretariat of Canada, 2018). At the 2018 G7 summit in In February 2019, the United States president signed an executive order entitled, "Maintaining American Leadership in Artificial Intelligence" to provide leadership for "shaping the global evolution of AI in a manner consistent with out Nation's values, policies, and priorities" (Trump, 2019, sec. 1). And, in April of 2018, twenty five countries in the European union signed a "Declaration of cooperation on Artificial Intelligence" to ensure Europe's competitiveness in relation to AI as well as deal with the social, economic, ethical, and legal questions that surround it (European Commission, 2018a). As governments increasingly look to adopt AI technology, they give rise to the corresponding question of how humans and AI programs work together.

This thesis will explore human-AI decision-making relationships in Canadian administrative contexts and how these dynamics can be structured to enhance their effectiveness and promote consistency with the obligations surrounding public decision-making. This thesis will proceed in eight major sections. Following this introduction, the second section will provide background information on Canadian administrative bodies, artificial intelligence, and how these administrative bodies can benefit from AI. The third section will discuss the tensions between Canadian administrative law and the current state of AI technology and how human-AI decision-making relationships can help alleviate this tension while section four will introduce the new considerations that arise when humans interact with AI programs. The fifth section will introduce this paper's methodology and the analytic framework that will be applied to the cases in section six. The seventh section reflects on the case studies and will discuss their common themes, relevance to Canadian administrative contexts, and introduce some examples of how human-AI decision-making relationships can be differently structured. The eighth section will conclude this thesis. Before beginning section two, the remaining subsections will introduce some definitions, establish the scope of this thesis, and briefly introduce the methodology.

1.1 Definitions

Decision-Maker:

As administrative bodies are created through legislation, there is a lot of flexibility in how they perform their functions. Relevant to this thesis, this flexibility means that there is a lot of variation in who is making the decision, from individuals to groups of people. The term, "decision-maker" will be used to refer to all of the different individuals and groups that make decisions in administrative contexts.

Artificial Intelligence (AI):

While "artificial intelligence" is a very broad term, it will have a more restricted meaning within the context of this paper as it will refer to computer programs developed through machine learning techniques. This narrower definition is used because of the unique issues of transparency and bias that follow from machine learning in comparison to explicitly coded programs, as will be discussed in more detail in sections 2(b) and 3(b).

Harmful Decision-Making Practices:

There are many ways in which decision-making can lead to harmful outcomes. For example, irrelevant factors can be considered as relevant (or vice versa), factors can be weighted improperly and protected characteristics can be used to negatively impact the decision subject's interests. This term captures a broader range of decision-making practices that can lead to harmful outcomes than terms like "bias" and "discrimination". Further, this term is more neutral and escapes the often negative connotations surrounding "bias" and "discrimination" which can have more neutral or positive meanings. For example, decision-making processes can be biased favourably towards disadvantaged groups while "discrimination" can be understood more neutrally as an exercise in classification.

Decision Subject:

The term "decision subject" will be used to broadly refer to an individual, group, organization, or other entity that is affected by an administrative decision. Because administrative decisions can take a wide variety of forms, this term will be used to refer to the correspondingly wide variety of subjects that can be subject to administrative decisions.

1.2 Scope

There are three main restrictions to the scope of this paper that will be introduced here. First, this thesis focuses on the human elements of human-AI decision-making relationships rather than working to solve the concerns within AI decision-making itself. As will be discussed in section 3(c), there are prominent concerns about bias and transparency in AI decisions. This thesis will not be working to solve these concerns through better software development practices and techniques, but rather will look at how these concerns are addressed through relationships with human actors. As will be discussed in section 3(b), this paper will construct AI programs as having a sufficiently robust development process so that there are no immediate or obvious issues of bias, but that the possibility of harmful decision-making practices cannot be fully eliminated.

Second, this thesis will be written from the perspective of Canadian administrative law. In contrast to other public decision-making contexts, administrative law has been selected because administrative bodies have a significant degree of freedom when determining how each of their processes balances the needs of "fairness, efficiency, and predictability of outcome" (*Knight v. Indian Head School Division No. 19*, 1990, p. 685). This flexibility allows for a greater range of possibilities when considering how decision-making processes address the unique challenges of AI decision-making. However, it is the hope that the principles and foundations that contribute to the analytic framework and case study hold enough commonality with broader values of public decision-making so that the findings and insights are generalizable to other contexts.

Third, this thesis only aims to map out early implementations of human-AI decision-making relationships. While the discussion section introduces a few decision-making structures that can help compensate for harmful decision-making practices, these should be thought of more as tools that may be appropriate in certain contexts, rather than a comprehensive framework that can help guide the development of administrative processes. Consequently, a full solution to any problematic dynamics or structures is outside of this paper's scope, especially given the diversity of administrative bodies and their functions.

1.3 Methodology

This thesis will be following Merriam and Tisdell's methodology for a "qualitative case study" to analyze the selected cases of human-AI decision-making relationships (Merriam & Tisdell, 2015). This methodology will be introduced in more detail in section 5 before its application in section 6.

2. Background

In order to properly discuss the tensions between AI and the obligations of public decision-makers, this section will introduce the role of administrative bodies in Canadian governance, key characteristics of Artificial Intelligence, and how AI can benefit administrative bodies.

2.1 Administrative Bodies

While they are not an original part of Canadian governance, administrative bodies developed in response to a need for government decision-making in contexts which neither the court, executive, or legislature were suited for (Flood & Sossin, 2013, pp. 4–5). Against a backdrop of intense debate of whether or not such bodies should exist, the first "non-elected full-time body outside any departmental structure" began with the *Railway Act* of 1903 which created Board of Railway Commissioners (Law Reform Commission of Canada, 1980, p. 23). This reflected a significant shift where experts were appointed to decide issues rather than to merely advise politicians which was iterated again with the *Canada Grain Act* of 1912 which established an administrative agency to generally oversee "all matters related to the inspection, weighing, trading and storage of grain" (Law Reform Commission of Canada, 1980, p. 24). From these beginnings, the growth of administrative bodies rapidly grew as government intervention greatly expanded to meet the war effort for World War I (Law Reform Commission of Canada, 1980, p. 24). These administrative bodies proliferated because they held a set of advantages that made them better suited to deal with certain issues than the executive, legislature, or judiciary.

There are three main characteristics of administrative bodies which define the niche that they fulfill: expertise, throughput, and political insulation. First, administrative bodies provide a level of domain expertise that may not be found in elected representatives or the judiciary (Law Reform Commission of Canada, 1980, p. 35). Second, administrative bodies deal with a large quantity of issues that would take up a disproportionate amount of time if they were to be resolved through parliamentary or judicial processes (Law Reform Commission of Canada, 1980, p. 35). And finally, administrative bodies are not directly affected by the political considerations in the same way that elected representatives must always consider the electorate and political optics which can negatively influence decisions (Law Reform Commission of Canada, 1980, p. 35). As the scope of government function grew, administrative bodies were created to fulfill a need for arm's-length decision-making with specialized expertise and greater decision-making capacity (Flood & Sossin, 2013, p. 8).

2.2 Artificial Intelligence (AI)

Often, the terms "artificial intelligence" and "machine learning" are used interchangeably, but the differences between them are important for this paper. The modern field of "Artificial Intelligence" started in the 1950s and refers to "understanding, modeling, and replicating intelligence and cognitive processes by invoking various computational, mathematical, logical, mechanical, and even biological principles and devices" (Frankish & Ramsey, 2014, pp. 1, 17). Machine learning represents a subfield which, in contrast to previous computer programs that can "only [execute] the algorithm it was programmed to run", focuses on programs that have “the capacity to define or modify decision-making rules autonomously ” (Alpaydin, 2016, p. 16; Frankish & Ramsey, 2014, p. 18; Mittelstadt et al., 2016, p. 3). This difference means that programs can be considered as AI if they perform human-like tasks but can only be considered as ML when they “learn” how to perform tasks from analyzing data rather than explicit programming from humans. While ML research has its origins in a checker playing program designed by Arthur Samuel in 1959, ML has only recently become competent enough for widespread implementation because of the availability of powerful and inexpensive computing resources and vast datasets (Frankish & Ramsey, 2014, p. 18; O’Leary, 2013, p. 96).

At its foundation, ML-based AI programs use large amounts of data to generate models that provide a prediction about how future cases should be understood. In a very simple summary, ML programs generate models through “training” where the program searches for patterns and probabilities in a dataset (Broussard, 2018, p. 32). While there are many different techniques which ML programs use to search for patterns and probabilities – such as deep learning, neural networks, supervised learning, and unsupervised learning – a common feature is that ML programs use the dataset to determine how the features of the input data relate to the outcome (Alpaydin, 2016, pp. 24–25). Once the training process is complete, the AI program will have defined the weight of each data point’s influence over the outcome to create a “model” which can be used to predict the outcome for similarly structured data that it has not encountered before (Broussard, 2018, p. 32).

The contemporary advancements in AI are enabled through foundations in Big Data. “Big Data” refers to a trend of decreasing storage costs alongside increasing computing power to the point where it was feasible for companies to collect and store vast amounts of data (de Laat, 2017, p. 2; Lazer & Radford, 2017, p. 20). These developments made it economically feasible to collect minute details that can be leveraged to make decisions in large areas of society, such as credit worthiness, employment and hiring, higher education, and criminal justice (Executive Office of the President, 2016, p. 2). Since the era of Big Data began, data collection techniques have only become more refined as storage has only become less expensive, computers more powerful, and sensors more ubiquitous (Executive Office of the President, 2016, p. 4). While AI predates Big Data, the data collection, storage, and processing capabilities that define Big Data have created the datasets necessary to fuel growth and capability in AI systems.

2.3 Benefits of AI in Administrative Contexts

While there is a lot discussion about how AI will fundamentally transform society, AI programs can provide many benefits for administrative decision-making as many AI capabilities align with the rationale for administrative bodies. First, and perhaps most obviously, AI programs can improve the efficiency of administrative bodies by analyzing information and producing a decision faster with less cost than a human or group of humans would (Gaon & Stedman, 2019, p. 19). Second, AI programs can also improve access to justice as more efficient decision-making can lower costs, provide faster decisions and lower barriers related to the time and cost of travelling to an administrative body (Beatson, 2018, p. 320). Further, AI programs can enhance equality under the law as all decisions would be made by the same program instead of a diverse array of human decision-makers (Corbett-Davies, et al., 2017, p. 1). Even though the administrative body's previous decisions are not binding on itself, it is still desirable that similar cases are still decided similarly (*Domtar Inc. V. Quebec (Commission d'appel en matière de lésions professionnelles)*, 1993, p. 799). And finally, AI decision-making can promote the equal application of the law as all decisions would be decided by the same program instead of humans which are very diverse and can be influenced by a wide array of conscious and unconscious processes (Beatson, 2018, pp. 323–324; Jones, 2013, p. 121; Rehaag, 2017, p. 53). These benefits strongly map onto the rationale of administrative bodies which, as mentioned previously, were created to provide expert, efficient, and arms-length decision-making on behalf of parliament.

3. Tensions between Administrative Law and AI

While AI programs can significantly contribute to administrative decision-making, these benefits are contingent on their proper design, implementation, and operation. There are many points where AI programs can fail to uphold the duties of public decision-makers and, instead, promote harmful decision-making practices. This section will first review the duties and obligations imposed on public decision-makers under Canadian administrative law, review potential sources of conflict between AI decision-making and the duties of public decision-makers, and conclude with a discussion of how human involvement can help alleviate these conflicts.

3.1 Duties and Obligations in Administrative Decision-Making

Consistent with the wide variety administrative bodies and their functions, administrative law has also developed a complex flexibility to accommodate the diverse processes and decisions that administrative bodies deal with. This subsection will provide a brief overview of the requirements of procedural fairness, impartial decision-making, and substantive review to provide a framework of duties and responsibilities for the following discussions of human-AI decision-making relationships.

3.1.1 Procedural Fairness

While administrative bodies generally have the freedom to determine their own decision-making process, administrative law requires that their procedures meet the requirements of procedural fairness. Procedural fairness embodies the principle that those affected by an administrative decision "have the opportunity to present their case fully and fairly, and have decisions affecting their rights, interests, or privileges made using a fair, impartial, and open process" (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 28). Yet, as the diversity of administrative bodies requires a more flexible approach, the requirements of procedural fairness are "eminently variable and… to be decided in the specific context of each case" in relation to the "statutory, institutional, and social context" of the decision being made" (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 28; *Dunsmuir v. New Brunswick*, 2008, para. 79).

The demands of procedural fairness in any particular administrative context are determined through a two-part test. First, procedural fairness will only apply if the administrative decision affects "the rights, privileges or interest of an individual" (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 20). If this threshold is met, then the court will proceed to consider five separate factors:

1. **"[T]he nature of the decision being made and the process followed in making it",** where the subject's procedural entitlements will be more demanding if the decision-maker's process is more similar to judicial processes (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 23).
2. **"[T]he nature of the statutory scheme and 'the terms of the statute pursuant to which the body operates'"** which considers how the decision made relates to the larger

statutory scheme. For example, if the decision cannot be appealed, then the demands of procedural fairness are more strict or, if the decision is about an exception to broader statutory rules, then procedural fairness will be less demanding (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, paras. 24, 31).

3. **The importance of the decision to the individual or individuals affected** where more important or significant decisions support a more rigorous process (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 25).
4. **The legitimate expectations of the person challenging the decision,** where procedural fairness will be more strict if the claimant had "a legitimate expectation that a certain result will be reached" but was not met (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 26).
5. **The choices about procedure made by the agency itself,** where the court must respect the ADM's established procedures, especially when the statute grants the power for the administrative body to make its own procedures or the administrative body has expertise to determine appropriate procedures (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 27).

After considering these factors, the court will decide the level of procedural fairness that is appropriate to the circumstances and whether this standard is met by the administrative body's existing procedures (Flood & Sossin, 2013, p. 27). If the existing procedures are insufficient, the court may require additional procedural elements such as an oral hearing, the opportunity for written arguments to be submitted, or requiring the decision-maker to provide reasons for its decision (Flood & Sossin, 2013).

3.1.2 Impartial Decision-Making

Another tenet of procedural fairness is that decisions should "be made free from a reasonable apprehension of bias by an impartial decision-maker" (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 45). A reasonable apprehension of bias is determined under the following test:

> *what would an informed person, viewing the matter realistically and practically - and having thought the matter through - conclude. Would he think that it is more likely than not that [the decision-maker], whether consciously or unconsciously, would not decide fairly* (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 46).

Importantly, the standards for a reasonable apprehension of bias can vary depending on the "context and the type of function performed by the administrative decision-maker involved" (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, para. 47). For example, in *Baker*, the decision-maker was an immigration officer and, within this context and function, the Supreme Court of Canada found that the test for bias must account for "a recognition of diversity, an understanding of others, and an openness to difference" which was not demonstrated (*Baker v. Canada (Minister of Citizenship and Immigration)*, 1999, pp. 47–48). Here, the Court provides guidance that bias is defined relative to the decision being made, rather than being a static element to look for.

3.1.3 Substantive Judicial Review

In contrast to procedural fairness which is concerned with the transparency of and ability for affected individuals to participate in the decision-making process, substantive judicial review ("substantive review") evaluates the merits of the administrative decision itself (*Dunsmuir v.*

New Brunswick, 2008, para. 130). However, similar to procedural fairness, the requirement of "acceptable" is context-dependent which is captured in the question of what "standard of review" the decision will be reviewed under.

There are two standards of review which determine how strictly the court will review the ADM's decision: reasonableness and correctness. (*Dunsmuir v. New Brunswick*, 2008, paras. 34, 45). Under the reasonableness standard, the court is concerned with checking for "justification, transparency and intelligibility" and, ultimately, whether "the decision falls within a range of possible, acceptable outcomes which are defensible in respect of the facts of law" (*Dunsmuir v. New Brunswick*, 2008, para. 47). Reasonableness is a deferential standard which requires courts to hold "a respectful attention to the reasons offered or which could be offered in support of a decision" (*Dunsmuir v. New Brunswick*, 2008, p. 48). In contrast, the standard of correctness involves no deference as the court undergoes its own analysis and uses its own conclusion instead of the ADM's (*Dunsmuir v. New Brunswick*, 2008, p. 50).

The correct standard of review is determined through a two-step process. First, the court must look to previous decisions to see if the appropriate standard of review for the current context has been determined before (*Dunsmuir v. New Brunswick*, 2008, paras. 57–58). If the standard of review for a particular context has not been determined before, then the appropriate standard of review is selected by the "pragmatic and functional" test outlined below:

1. **The Existence of a privative clause** will strongly support a reasonableness standard as it indicates that the legislature intended to give the decision-maker greater deference (*Dunsmuir v. New Brunswick*, 2008, para. 52). A "privative clause" is a statutory provision that "limit[s] or preclude[s] judicial review of a defined category of administrative action or decision (Mann & Blunden, 2010).

2. **A question of fact, discretion or policy** will normally "automatically" apply a reasonableness standard (*Dunsmuir v. New Brunswick*, 2008, para. 53).
3. **A tribunal interpreting its enabling statute, closely connected statutes, or other legal areas it has gained expertise in** will support reasonableness because of the tribunal's familiarity and expertise with the law (*Dunsmuir v. New Brunswick*, 2008, para. 54). An "enabling statute" in this context refers to the legislation that gives the administrative body its powers (Law, 2018).
4. **The nature of the question of law**, where a question of jurisdiction or "central importance to the legal system" and outside of the ADM's "specialized area of expertise" will support a correctness standard (*Dunsmuir v. New Brunswick*, 2008, paras. 55, 59–60). Here, a question of jurisdiction is when the tribunal determined if its enabling statute granted it the power to decide a particular matter (*Dunsmuir v. New Brunswick*, 2008, para. 59).

3.1.4 Key Points

From this overview of administrative law, there are a few key elements that will be important to later discussions of artificial intelligence. These key elements will be outlined below:

1. Administrative law must accommodate a very diverse range of decision-making structures and processes.
2. To accommodate this diversity, administrative law has developed tests for procedural fairness, bias, and substantive judicial review that are defined relative to the subject matter and legal context of the decision.

3. Courts are generally reluctant to intervene in the substance of a decision unless it is an issue of jurisdiction or "central importance to the legal system" outside of the ADM's expertise.

Collectively, the requirements of procedural fairness, impartial decision-making, and substantive judicial review form the obligations and responsibilities of public decision-makers within the context of this paper. While there are many more sources of obligations and responsibilities in administrative law, the ones introduced in the previous subsections are highly relevant to human-AI decision-making relationships and will form the foundation for how these relationships are evaluated. The significance of these obligations and responsibilities will be discussed in section 3(c) after section 3(b) establishes how they are in tension with AI decision-making.

3.2 Sources of Conflict in in AI Decision-Making

The current state of AI development can conflict with the duties and obligations of public decision-making established in the previous subsection. This subsection will discuss bias and transparency concerns in contemporary AI programs to support the following discussion of how implementing AI programs in administrative decision-making contexts can be problematic.

3.2.1 Bias in AI Decision-Making

Concerns about bias surround the possibility that AI programs create harmful correlations between a protected characteristic and a negative outcome (Barocas, Bradley, Honavar, & Provost, 2017, p. 2; BC Human Rights Tribunal, n.d.). Bias is most explicitly found when protected characteristics such as race, gender, and age are used to make decisions (Barocas et al., 2017, p. 2). While it is fairly easy to prevent ML programs from directly considering protected characteristics, the task of eliminating bias is much more difficult and complicated (Barocas et al., 2017, p. 2). At its foundation, AI programs can develop biases because, during the training process, their task is to identify whatever features of the data are relevant to the outcome to create an accurate prediction. This process creates various opportunities for bias to emerge.

First, the AI program may develop biases by "learning" from an incomplete dataset. If data collection practices fail to include important information or fail to include certain groups, then the AI program will not "learn" how to properly evaluate missing or underrepresented information during operation (IEEE, 2016, p. 26). A simple example would be a face-detection system that was trained on Caucasian faces which would then have difficulty recognizing non-Caucasian skin tones (IEEE, 2016, p. 27). Similarly, if a dataset generated from the internet is used for training an ML program, populations with lower internet use will likely be underrepresented (Mergel, Rethemeyer, & Isett, n.d., p. 933). Ensuring that that AI programs learn from a representative dataset is a difficult, yet necessary task.

Second, AI programs can revive and perpetuate historical biases that contributed to the dataset. For example, historical biases in police forces could result in more information gathering on minority communities which results in predictive policing programs devoting more police resources towards minority communities (Karppi, 2018, p. 4; Lum & Isaac, 2016, p. 18). Similarly, computer programs may learn gendered word associations, such as associating "receptionist" with "female", when training on natural language datasets (Bolukbasi, Chang, Zou, Saligrama, & Kalai, 2016, p. 1; Caliskan, Bryson, & Narayanan, 2017, pp. 183–184). When working with historical datasets, it is very difficult to identify any biases that went into its development.

Human programmers can also contribute to biased AI programs through the subjective decisions inherent in the development process. Far from being an objective process, there are many choices with many possible outcomes that must be made when developing a ML program which have no clearly correct choice (Mittelstadt et al., 2016, p. 7). For example, AI developers can unintentionally influence the AI program when labelling training data. In a more malicious example, an AI program could intentionally bias the training data which, without sufficient transparency and testing, would be hidden in the AI program's operation (Barocas & Selbst, 2016, pp. 681–682, 692). As a result of these subjective choices, ''the values of the author [of an algorithm], wittingly or not, are frozen into the code, effectively institutionalising those values'' (Macnish, 2012, p. 158).

The issue of bias would not be so problematic if an AI program's "reasoning" could be easily audited for any harmful decision-making practices. If this were the case, then incorrect decisions could be easily identified and remediated. However, the training process creates an inherent layer of opacity that makes it quite difficult to determine how an AI program made any particular decision.

3.2.2 Transparency in AI Decision-Making

Machine learning programs are generally viewed as inherently opaque because of how difficult it is to understand the relationships they "learn" through the training process. As ML programs generate a model from the training data, it becomes difficult to interpret "how each of the many data-points used… contribute to the conclusions it generates" (Mittelstadt et al., 2016, p. 4). While releasing the program's source code for analysis may seem to solve the issue of transparency, it is not completely effective in the ML context as "the decisional rule itself emerges automatically from the specific data under analysis, sometimes in ways that no human can explain" (Kroll et al., 2017, p. 638). The complexity generated through the training process creates a considerable barrier to developing transparent and accountable ML programs.

Even if an AI program can communicate all the factors it considered to make a decision, the possibility of proxy factors still creates significant barriers for effective transparency. A "proxy factor" occurs when an otherwise acceptable factor corresponds to a protected characteristic that should not be considered (Mittelstadt et al., 2016, p. 8). A prominent historical example is 'redlining' where banks would refuse services to people located in certain areas; while location is not a protected characteristic, the specified locations corresponded to race which made the practice discriminatory (d'Alessandro, O'Neil, & LaGatta, 2017, p. 121). Similarly, even if ML programs do not explicitly consider a protected characteristic, which is a fairly easy task to accomplish, it may pick up associations which correspond to a protected characteristic that are much harder to detect (Kilbertus et al., 2017, p. 2; Romei & Ruggieri, 2014, p. 584).

3.3 Tensions Between Administrative Law and AI Decision-Making

There are three major potential conflicts between AI programs and the requirements of administrative law that can create barriers for implementing AI programs in public contexts. Putting the previous two subsections in conversation with each other, this subsection will discuss how the concerns about bias and transparency in AI programs create three significant sources of tension with the interest of impartial, transparent, and fair public decision-making under Canadian administrative law.

First, the potential for AI programs to develop bias can conflict with the requirement for ADMs to make impartial decisions. This source of tension results from sections 3(a)(ii) and

3(b)(i) which discussed how ADMs must make decisions impartially and that AI systems can "learn" harmful decision-making practices from the datasets they train on. If an AI program acts on a prohibited bias, then this would conflict with the duty of impartial decision-making.

Second, concerns about a lack of transparency in AI programs can not only compound the issue of bias, but also conflict with procedural fairness requirements as well. As discussed in section 3(b)(ii), AI programs cannot be fully transparent yet, which makes it difficult to find and address any harmful decision-making practices. In relation to the requirements of procedural fairness establishes in section 3(a)(i), the lack of transparency can also conflict with requirements to give reasons to the decision subject as the decision may not be fully explainable.

And finally, any remedies ordered to correct an AI program can be undermined because the level of efficiency and scale that they can operate on. While AI programs are promising because they can help increase the efficiency of administrative bodies as established in section 2(c), this also means that any faults found in them after their implementation could have already influenced many more decisions than a human decision-maker could have made. This means that relying on *ex-post* review of AI decisions carries more risk as errors can propagate on a much larger scale than previously possible.

3.4 Human-AI Decision-Making Relationships as a Solution

One solution to the potential detriments of AI programs is to create decision-making processes that integrate both human and AI capabilities. This is reflected in the Government of Canada's *Directive on Automated Decision-Making* ("the Directive"). The Directive organizes Automated Decision Systems into four categories, with "Level I" making the least consequential decisions and "Level IV" making the most consequential decisions (Government of Canada, 2019, sec. Appendix B). The Directive further requires that higher level systems being held to proportionally higher standards of development and operation (Government of Canada, 2019, sec. Appendix C). Most immediately relevant are the "human-in-the-loop" requirements for Level III and Level IV systems which state that "[d]ecisions cannot be made without having specific human intervention points during the decision-making process; and the final decision must be made by a human (Government of Canada, 2019, sec. Appendix C). Of central importance to this thesis is the trust placed in human co-determination and oversight of AI programs where humans are always involved and have final discretion over the most significant decisions. As the following section will show, putting humans in decision-making relationships with AI systems creates new dynamics that must be accounted for to promote positive processes and outcomes.

4. New Considerations in Human-AI Decision-Making Relationships

While humans can significantly contribute to monitoring for and preventing harmful decision-making practices in AI programs, this also adds additional considerations found in the relationship between humans an AI programs. This section will discuss how cognitive biases and heuristics, incentives, and trust can upset effective decision-making relationships.

4.1 Biases and Heuristics

While they are often perceived negatively, biases and heuristics are not always harmful and, in many cases, can be quite valuable. Biases and heuristics can be generally understood as strategies and rules that help simplify and process large amounts of information which can lead to “systematic and predictable errors in judgment” (Blumenthal-Barby & Krieger, 2015, p. 539). Even though relying on a bias or heuristic can potentially lead to relevant information being overlooked or skew the weighting of known information, they also act as valuable mechanisms to conserve attentional resources and make reasonable decisions when it is impractical to engage in more extensive information gathering and processing (Blumenthal-Barby & Krieger, 2015, pp. 539–530; Choo, 2005, p. 208).

However, biases and heuristics can be harmful because of their often subtle and persuasive nature in decision-making. As Jones writes, biases and heuristics can operate invisibly as the brain constructs rational justifications for their operation after they have already unconsciously exerted their influence (Jones, 2013, pp. 60–61). Further, even when the individual is consciously aware of activated biases and heuristics, “the mind’s inclination is to support and confirm, rather than to critically analyze and constantly reconsider” (Jones, 2013, p. 61). For these reasons, any ways in which an AI program can activate biases and heuristics in human decision-makers can increase the prominence and impact of harmful decision-making practices.

This subsection will be largely guided by Blumenthal-Barby and Kreiger’s work on biases and heuristics. Blumenthal-Barby and Kreiger’s work consists of a review of 214 empirical studies on biases and heuristics in decision-making; which identified a total of 19 different biases and heuristics that manifest in human decision-making (Blumenthal-Barby & Krieger, 2015, p. 542). As this work was completed in the context of medical decision-making, the survey will be used to identify the biases and heuristics that are relevant to the context of human-AI decision-making relationships (Blumenthal-Barby & Krieger, 2015).

4.1.1 Anchoring and Adjustment

Anchoring and adjustment biases occur where an initial value will often bias responses towards it even if the value itself is arbitrary. (Kahneman, Ritov, & Schkade, 1999, p. 226). Applied to human-AI decision-making, anchoring and adjustment bias can limit the human actor's ability to intervene or contribute to the decision by inhibiting any adjustments or deliberations from straying far from the AI's determination. This can compromise the decision-making process as anchoring biases could prevent the human from making substantial adjustments when necessary. On a larger scale, anchoring and adjustment biases can allow harmful decision-making practices to go unaddressed as the human actor does not sufficiently compensate for them.

4.1.2 Confirmation Bias

Confirmation bias occurs where one believes that "there is more support for [their] beliefs than actually exists in the evidence at hand" (Kahneman, Tversky, & Slovic, 1982, p. 149). Confirmation bias poses a risk to human-AI decision-making relationships as it creates the risk that human actors will only use the AI output when it reinforces their own evaluation and otherwise discard it. The opportunity for confirmation bias to express itself is problematic as it can undermine any intended benefits gained from the consistent application of the law through the AI program.

4.1.3 Default/Status Quo Bias

The default/status quo bias reflects the strong tendency for individuals to accept the status quo and avoid disturbing it. In operation, this bias can lead humans to either over-accepting or over-rejecting AI decisions depending on how they are situated (Kahneman & Tversky, 2000, p.

163). The status quo bias can compromise human-AI decision-making relationships as it can lead human actors to avoid intervening in the AI's decision. For example, if a human must provide reasons to intervene in an AI program's decision but not to agree with it, then this establishes the AI program's decision as a "default".

4.1.4 Sunk Cost Effect

A sunk cost effect favours continuation once "an investment in money, effort, or time has been made" (Arkes & Blumer, 1985, p. 124). In relation to human-AI decision-making, sunk cost effects create the risk that AI decisions will be regarded as an investment of effort which human actors would be hesitant to intervene in. For example, if the human decision-maker also operates an AI program, then they would have invested effort in producing the output which creates space for sunk cost effects to potentially inhibit any review of that output.

4.1.5 Order Effects

The order in which information is presented can impact how persuasive that information is. "Order effects" refer to how information presented at the beginning (primacy effect) or end (recency effect) will be more easily remembered and have a greater impact on the outcome than information presented in the middle (Kahneman & Tversky, 2000, p. 569). Consequently, the timing of when the AI program's decision is presented to the human overseer can influence its persuasiveness despite being an irrelevant factor. For example, if the AI output is presented before a human actor has the opportunity to consider the information, then the human's ability to conduct an independent evaluation may be compromised.

4.2 Trust and Suspicion

To the extent that trust is the use of prior attitudes and beliefs to guide one's judgment, it acts as a heuristic as well (Cummings, 2014, p. 7). However, in contrast to the previously discussed heuristics and biases, trust is unique as it represents the more social dimensions of human-automation interaction (Lee & See, 2004, p. 50). Structuring human-AI decision-making relationships presents the interesting challenge of maintaining the appropriate degree of trust and suspicion to support appropriate human reliance on AI decision-making. This subsection will largely rely on the influential work of Lee and See (2004), Parasuraman and Riley (1997), and Parasuman, Sheridan, and Wickens (2000).

In human-AI decision-making relationships, the amount of trust which the human element holds towards the AI program can problematically affect the decision to agree or intervene. Defined as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability", trust in an AI program plays a significant role in whether the human overseer decides to agree or intervene (Lee & See, 2004, p. 51; Parasuraman & Riley, 1997, p. 234). If humans trust automation too much, then it leads to the issue of "complacency" where "the operator may not monitor the automation and its information sources and hence fail to detect the occasional times when the automation fails" (Parasuraman, Sheridan, & Wickens, 2000, p. 291). The challenge in designing human-AI decision-making relationships is to maintain an appropriate amount of trust so that the human element is still likely to catch potential errors, but not too much or too little trust so that AI decisions are accepted or intervened in too often.

4.3 Significance

While a comprehensive overview of how human biases and heuristics contour human-AI decision-making relationships is beyond the scope of this thesis, the provided overview reflects the more prominent risks that they pose to proper decision-making. Alongside the previous section on the obligations and responsibilities of public decision-makers under Canadian administrative law, this section will inform the analytic framework that will be applied to the selected cases under the qualitative case study.

5. Methodology

Using Merriam and Tisdell's qualitative case study methodology, this section will establish the framework that will be used to study human-AI decision-making relationships in public contexts to determine how they have been implemented and address – or fail to address – the many concerns surrounding human-AI interactions. After providing an overview of the methodology and how it applies to human-AI decision-making relationships, the following subsections will establish the selection criteria and the relevant dimensions that each case will be compared across.

5.1 Qualitative Case Study

Under Merriam and Tisdell's work, a qualitative case study is "an in-depth description and analysis of a bounded system" (Merriam & Tisdell, 2015, p. 39). Within this thesis, AI programs operating in public contexts will constitute the "bounded system"; while it is difficult to disentangle public AI systems from their overarching administrative and legal contexts, the bounded system will be defined as the decision-making process that AI systems are deployed in.

This will span the inputs into the decision-making process – such as information, applications, and people – and end with the process' output and its surrounding context.

This case study will largely rely on mining data from documents to analyze the selected cases (Merriam & Tisdell, 2015, p. 162). Because of their use in government decisions and their impacts on individual rights and interests, there are a lot of public documents and academic research to support inquiries into human-AI decision-making relationships (Merriam & Tisdell, 2015, p. 163). Further, the software vendors often have robust information about their products which can also assist into this investigation. Here, "documents" is used to broadly capture "a wide range of written, visual, digital, and physical material relevant to the study" (Merriam & Tisdell, 2015, p. 162).

5.2 Selection Criteria

Because of the relatively few uses of AI in public contexts, cases will be selected through purposive sampling (Merriam & Tisdell, 2015, p. 96). Purposive sampling has been selected over random sampling because it supports this thesis' goal to discover "what occurs, the implications of what occurs, and the relationships linking occurrences" in human-AI decision-making relationships whereas random sampling is more suited for questions of quantity and frequency (Merriam & Tisdell, 2015, p. 96). Because of the relatively few uses of AI in public contexts, convenience sampling will be used to examine the relatively few implementations most relevant to Canada through the selection criteria outlined below (Merriam & Tisdell, 2015, p. 98).

First, and most fundamentally, the AI program must be implemented in public contexts meaning that it is used by a government to make legal determinations. This requires that the AI program's decisions have some legal significance by affecting the rights or interests of the subject.

Second, the government system must have legal protections for the subject that are similar to the protections required by Canadian law. Broadly, this means that the selected cases must be held to similar protections found in Canadian administrative law and, by extension, the Canadian Charter of Rights and Freedoms.

Third, and related to the scope of this paper, the AI program must be developed through machine learning techniques. While the definition of "artificial intelligence" is broader, the scope will be restricted as machine learning programs carry with them the unique issues of bias and transparency established in section 3(b).

Together, these criteria will maintain relevance between the examined cases and human-AI decision-making relationships in Canadian administrative contexts. The first two criteria ensure that the AI system is subject to public decision-making obligations and responsibilities that are similar to those found under Canadian administrative law. The third requirement ensures that each case holds concerns about bias and transparency that cannot be fully resolved through technical means, and consequently require human and legal counterbalances to the trust placed in them.

5.3 Analytic Framework

The following cases will be reviewed under a novel analytic framework that, building on the previously established administrative law requirements and biases in human-automation interaction, will help analyze the quality of human-AI decision-making relationships. Broadly, this analytic framework will ask, "How is the AI program situated in relation to human elements in the decision-making process as well as the broader legal framework?" and will consist of two major dimensions. This question breaks down into two separate dimensions which the rest of the section will elaborate on.

5.3.1 Inquiring into Human-AI Dynamics

The first dimension inquires about how the AI program is situated in relation to human aspects of the decision-making process. This dimension is largely rooted in the earlier discussion about human decision-making biases as it looks at the dynamics between humans and AI programs and how they can help amplify or compensate harmful decision-making practices in the other.

1. The **independence** factor considers how closely humans and AI programs work together as a reflection of concerns about how much of an opportunity each element has to influence the other, such as through confirmation bias, trust, or suspicion (sections 4(a)(ii) and 4(b)). This factor would consider questions such as whether the AI program operates autonomously through automated data collection, requires a human operator, or if it relies on many humans for data input. Similarly, there would also be a consideration of whether a human operator is also the decision-maker. It is also important to consider if humans and AI programs consider the same factors or if there is a division of labour where each considers separate factors.
2. Looking at the **sequence** of AI programs examines when AI outputs are presented to human elements in the decision-making process. This factor will mainly address concerns about anchoring and adjustment, status quo bias, sunk cost effects, and order effects (sections 4(a)(i), 4(a)(iii), 4(a)(iv) and 4(a)(v)). The most prominent question that follows

from this factor asks about whether the human has an opportunity to independently consider the case before being presented with the AI's output.

3. **Transparency** looks at how the AI program's output is communicated to humans as a product of concerns about trust and complacency, trust and suspicion, anchoring, and status quo bias (sections 4(a)(i), 4(a)(iii), 4(a)(v) and 4(b)). This factor will produce questions surrounding how much information an AI program gives about its decision for humans to engage with, such as how much the AI program's "reasoning" is presented alongside its conclusions and whether a confidence level in the decision is calculated as well.
4. The factor of **instruction** looks at how human elements in the decision-making process are trained and direct to interact with AI programs. First, this factor looks at whether any training was provided and, if so, its frequency (one-time or ongoing), what organization provided that training, and whether the training properly equips humans to interact with the AI program. Beyond any training provided, this factor also looks at any directions that the decision-maker must, should, or can consider.

5.3.2 Inquiring into the Legal Context Surrounding AI

The second dimension of the analytic framework examines the legal context surrounding human-AI decision-making relationships. This dimension reflects the earlier discussion about the tensions between administrative law and AI. These factors reflect how much trust is placed in the AI program's determinations and how the surrounding legal context addresses concerns about harmful decision-making practices in AI programs. This dimension reflects the larger trade-off between efficiency and process where more counterbalances to AI determinations improves the chances that a decision will be properly made but, in turn, detract from the benefits of AI decision-making discussed in section 2(c) (efficiency, consistency, access to justice).

1. Looking at the **weight and significance** of the AI program's output examines how strong of a basis the output is for further action. For example, if the AI program only contributes a small part of a larger decision, if it is only a part of a multi-step process, or if its determination is completely subject to human consideration and discretion, then these would reflect a lower amount of trust in the AI program. In contrast, if the AI's determination is all that is needed to support action, or if there is no human co-determination, then there is a very large amount of trust placed in the AI program.
2. Examining the **appeal processes** surrounding AI programs analyzes how easy it is for the decision subject to appeal an AI determination. For example, applying human-oriented appeal structures to AI programs would reflect a very high level of trust while creating an automatic right to appeal an AI determination reflects a lower level of trust. Example questions will look at the opportunities which humans have to contest AI decisions as well as the costs associated with contesting that decision.

5.4 Methodology Summary

Under Merriam and Tisdell's qualitative case study methodology, the following section will apply the analytic framework to three cases of human-AI decision-making relationships. By examining the relationship between human actors and AI programs, and how these relationships are situated in the surrounding legal framework, this thesis will hope to map out how humans and AI programs work together to make decisions, how these structures compensate for concerns about harmful decision-making practices in AI, and what factors and considerations these structures do not fully address.

How is the AI program situated in relation to human elements in the decision-making process as well as the broader legal framework?		
Category	**Dimension**	**Example Questions**
Situated in Relation to Human Decision-Makers	Distance/Independence	• How closely is the AI program working with human actors? E.g. is the human actor operating the AI program or is there a division of factors that each considers? • How autonomous is the AI program? Is there a human operator? Is data entry done by one, a few, or many people? • Is the human DM operating the AI program or is the operator and human DM different people? • How static or dynamic is the required input? Does the operator have a lot of freedom in how the input is shaped or is the form of input more straightforward?
	Sequence	• When is the AI output presented to a human - before or after they consider the facts? • Is there an opportunity for human actors to independently consider the factors?
	Transparency	• How is the AI program's decision communicated to humans? • For example, is it just the conclusion? Does it provide an explanation? Does it include its confidence? • Has an effort been made to communicate the AI program's "reasoning" to the human DM?
	Instruction	• Is there training to prepare humans to interact with the AI program? • Is the decision-maker presented with any instructions on how to engage with the output?
		•
Situated in Relation to the Broader Legal Framework	Weight/Significance	• Is the AI output the basis for legal action or is it subject to human consideration and discretion? • Is there a meaningful capacity for humans to disagree with the AI program? • Can the AI output independently support further action?
	Appeal	• How can a human contest the AI decision? • What are the costs associated with appeal?

Table 1: Summary of the Analytic Framework

6. Case Study

There are three cases that meet the selection criteria and will be used for the qualitative case study: Northpointe's COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) risk assessment system, the European Union's iBorderCtrl program, and the PredPol predictive policing system. Each case study will first begin with a description of the AI program's use, the process for generating an output, and the jurisdictions which it is used in. These introductory sections will be followed by the application of the analytic framework which will be broken down into two sections: the AI program's relationship to human decision-makers and its relation to the broader legal context. Following the analytic framework, two subsections will discuss how the human-AI decision-making relationship is structured to compensate for harmful decision-making practices in AI programs and what factors or concerns are not fully addressed. Before applying the analytic framework, this section will establish how each case meets the selection criteria established in section 5(b).

First, the AI system must be implemented in a public context. COMPAS operates publicly as it is implemented in many places throughout the American criminal justice system and will be specifically analyzed in its use in sentencing decisions (*State v. Loomis*, 2016, paras. 18–19). The PredPol system operates in a public context as it assists law enforcement organizations with allocating policing resources (PredPol, 2018). And finally, the iBorderCtrl program is implemented in a public context as it is designed to control how third country nationals cross land borders between EU Member States (iBorderCtrl, 2016c). It is important to note that the iBorderCtrl project is only a research project and not being developed for direct implementation (iBorderCtrl, 2016a). Despite this slight dissonance, the iBorderCtrl project will still meet this criteria as it has been designed to account for the "ethical principle and legal safeguards relating to human-machine interaction, privacy, personal data protection and informed consent, etc.," that it would have to abide by if officially implemented (iBorderCtrl, 2016b).

Second, the AI system must be operating in a jurisdiction that has similar protections to those found under Canadian law. PredPol and COMPAS, both of which are examined in the American context, are subject to the US constitution's fourth and fifteenth amendments which guarantee due process and equal protection under the law and are analogous to sections 7 through 15 of the *Canadian Charter of Rights and Freedoms* (Bender, 1983, pp. 836, 841, 847). Similar protections are applied to the iBorderCtrl project under the European Union's *Charter of Fundamental Rights* which also ensure that an affected individual has an opportunity to present their arguments for an outcome, that the administrative agency provides reasons for its decision, and that the decision-maker is unbiased (Shipley, 2008).

And finally, the system must be developed through machine-learning techniques. This requirement is met by COMPAS which has used regression modelling and machine learning to develop its risk scales (Northpointe, 2015, p. 14). PredPol also meets this criteria as it uses a machine learning algorithm to translate a law enforcement organization's records into predictions (PredPol, 2018). And finally, the iBorderCtrl program uses machine learning in its "Automated Deception Detection System" which "quantifies the probability of deceit in interviews by analysing interviewees' non-verbal micro expressions" (iBorderCtrl, 2016b).

6.1 COMPAS Risk Assessment System

While the COMPAS risk assessment system forms part of a broader software suite developed by Equivant (formerly Northpointe), this case study will focus on its use to inform

Wisconsin sentencing decisions (Northpointe, 2015, p. 2). The Wisconsin context has been selected because, alongside the official documentation, the Wisconsin Supreme Court's recent decision in *State v. Loomis* provides valuable details about how COMPAS is used in practice. Angwin et al.'s influential 2016 study also contains insights into the forms and outputs surrounding the COMPAS program that will further support the analysis. These three sources constitute the primary materials that will be used to establish how humans interact with the COMPAS program to make sentencing decisions.

When working with other components of the software suite, COMPAS is used to support placement, management, and treatment decisions for people convicted of crimes (*State v. Loomis*, 2016, para. 13). In the sentencing context, COMPAS risk assessments ("COMPAS reports") are also used in the United States to inform sentencing decisions to provide an indication of the convicted person's risk of violence and recidivism (Northpointe, 2015, p. 1; *State v. Loomis*, 2016, para. 18). A COMPAS risk assessment is based off information gathered from the defendant's criminal history as well as an interview with the defendant (*State v. Loomis*, 2016, para. 13). The Wisconsin risk assessment consists of 137 questions organized across 15 categories: current charges, criminal history, non-compliance, family criminality, peers, substance abuse, residence/stability, social environment, education, vocation (work), leisure/recreation, social isolation, criminal personality, anger, criminal attitudes (Angwin, 2016a). The questionnaire is partially reproduced in Appendix A.

Once completed, the information is then processed through the COMPAS risk assessment program which then produces the risk report. First, the COMPAS report presents the individual's "Overall Risk Potential" through measurements of the individual's likelihood for (1) violence, (2) recidivism, (3) failure to appear, and (4) community non-compliance in red bar graphs

(Angwin, 2016b). The following section presents the individual's "Criminogenic and Needs Profile" that, through green bar graphs, reflects "areas in a defendant's life that should be addressed in order to prevent a person from committing a new crime"; this section includes the headings of "criminal history factors", "needs assessment", "criminal attitudes", "social environment", and "higher order factors" (Angwin, 2016b). Each bar graph is placed on a scale of 0-10 where lower scores represent a lower risk or need (Angwin, 2016b). A sample COMPAS risk assessment can be found in Appendix B.

These COMPAS reports can then be attached to a Presentence Investigation Report (PSI) which informs the defendant's sentencing hearing. A PSI supports sentencing decisions by presenting information about (but not limited to) the defendant's family history, community ties, education, employment history, physical health, mental and emotional health, history of substance abuse, financial condition, and willingness to accept responsibility for their offense(s) (United States Probation and Pretrial Services, n.d.). This information is used to assess the defendant's risks and needs when under state supervision (United States Probation and Pretrial Services, n.d.).

It is important to note that a COMPAS report does not constitute the entire PSI, but is presented alongside other relevant information prepared by a probation officer (*State v. Loomis*, 2016, para. 12; United States Probation and Pretrial Services, n.d.). In a sample PSI released by the Wisconsin State Public Defenders, the document depicts many sections hold text generated by the COMPAS program alongside text created by the parole officer as well (Wisconsin State Public Defenders, 2014). Information contributed by COMPAS is marked by a small icon while the parole's officers comments that appear in the same section appear in a separate text box labelled, “agent comments” (Wisconsin State Public Defenders, 2014). The sample PSI is partially reproduced in Appendix C.

The implementation of COMPAS risks assessments has been met with some controversy surrounding the possibility that the risk assessments compromise the defendant’s due process rights. Angwin et al.’s study attracted a lot of attention when it found significant racial disparities which disadvantaged black defendants relative to white defendants; a finding that Equivant disputes (Angwin et al., 2016; Dieterich, Mendoza, & Brennan, 2016, p. 1). Angwin’s study and Northpointe’s response are indicative of the contested discourse surrounding the proper role of technology in the legal context where many other authors have discussed how COMPAS and other artificial intelligence programs should be properly situated within the courtroom (Freeman, 2016, pp. 75–78; Washington, 2018, pp. 133–136; Wisser, 2019, pp. 1811–1812). With the significant rights and interests at stake in criminal hearings, it is unlikely that this debate will be resolved in the near future.

Within the use of COMPAS risk assessments in the Wisconsin criminal justice system, there are four distinct human actors surrounding it. First, there is the person who operates the COMPAS program by providing the required information (“the operator”). Second, the defense and prosecution – each acting as a separate party - contribute to the interpretation of the COMPAS report. Here, the “defense” refers to the defendant and their legal representation (if present) because the latter will hold the same interests and arguments as the defendant. And finally, there is the judge who determines the significance of the COMPAS report to the sentencing decision and holds the position of “decision-maker”. With the context established for this case, the remainder of this section will apply the analytic framework to COMPAS.

6.1.1 Relation to Human Decision-Makers

Distance:

The use of COMPAS reports in sentencing decisions has a substantial degree of distance between the operator and the decision-maker. First, there is a division of functions where the COMPAS operator is a different individual from the decision-maker. Second, the presentation of the COMPAS report in the PSI is determined by the assigned probation officer. And third, the defence and prosecution both have the opportunity to present arguments about the interpretation of the COMPAS report for the consideration of the judge. Through these intermediary steps, there is a significant amount of space between the COMPAS report and the judge.

Further distance is found in the COMPAS system's static input as the human operator does not exercise a significant amount of control over how information is submitted. As the Supreme Court of Wisconsin writes, COMPAS relies heavily on static information with "limited use of some dynamic variables" such as the defendant's number of criminal associates, or history of substance abuse (*State v. Loomis*, 2016, para. 54). This means that most of the required information is quantification with little room for the operator's interpretation.

Sequence:

The COMPAS report does not enjoy special prominence in the sentencing hearing in relation to the timing of its presentation. As the COMPAS risk assessment is presented within the larger PSI, the judge receives it alongside other relevant information. This presentation happens at the beginning of the sentencing hearing as well, which gives a substantial opportunity for the defense and prosecution to argue about the risk assessment's interpretation, contextualization, and reception.

Transparency:

The COMPAS report does not take many steps to be transparent. While the report itself is quite easy to understand and any information generated by COMPAS is clearly labelled, it does not explain how it arrived at the results from the inputs. This is because Northpointe, the company which developed COMPAS, considers it a "proprietary instrument" and therefore "does not disclose how the risk scores are determined or how the factors are weighted" (*State v. Loomis*, 2016, para. 51). This leaves little opportunity for the prosecution, defense, or judge to engage with the reasons behind the risk assessment.

Instruction:

Judges have not received special training to engage with COMPAS reports, but instead, are subject to numerous instructions on how to properly use them. In *State v. Loomis,* the COMPAS report came with instructions about how it should be used to "identify offenders who could benefit from interventions and to target risk factors that should be addressed during supervision" as well as a warning that it "should not be used to determine the severity of a sentence or whether an offender is incarcerated" (*State v. Loomis*, 2016, paras. 16–17). However, the Wisconsin Supreme Court provides more guidelines on how courts can and cannot use risk assessments in sentencing decisions.

In response to the risk that COMPAS reports pose towards due process rights, the Wisconsin Supreme Court puts forward several cautions, restrictions, and requirements that judges must following when considering one. After acknowledging that the possibility of bias and inaccuracy cannot be fully eliminated, the Wisconsin Supreme Court states that "the use of a COMPAS risk assessment must be subject to certain cautions" and limitations so that courts are equipped to "better assess the accuracy of the assessment" and give the appropriate weight to the risk scores (*State v. Loomis*, 2016, paras. 65–66). Any PSI containing a COMPAS report must present the following cautions:

1. "[T]he proprietary nature of COMPAS has been invoked to prevent disclosure of information relating to how factors are weighed or how risk scores are to be determined;
2. [R]isk assessment compares defendants to a national sample, but no cross-validation study for a Wisconsin population has yet been completed;
3. [S]ome studies of COMPAS risk assessment scores have raised questions about whether they disproportionately classify minority offenders as having a higher risk of recidivism; and
4. [R]isk assessment tools must be constantly monitored and re-normed for accuracy due to changing populations and subpopulations" (*State v. Loomis*, 2016, paras. 66, 100).

The PSI must also state that the COMPAS report cannot be used…

1. "to determine whether an offender is incarcerated";
2. "to determine the severity of the sentence"; or
3. "as the determinative factor in deciding whether an offender can be supervised safely and effectively in the community" (*State v. Loomis*, 2016, para. 98).

The Court also specifies a list of permissible uses for COMPAS reports at sentencing hearings, which can be considered when

1. "diverting low-risk prison-bound offenders to a non-prison alternative";
2. "assessing whether an offender can be supervised safely and effectively in the community"; and

3. “imposing terms and conditions of probation, supervision, and responses to violations” (*State v. Loomis*, 2016, para. 88).

And finally, the Wisconsin Supreme Court states that a court must explain its use of the COMPAS risk assessment. Specifically, it must explain “…the factors in addition to a COMPAS risk assessment that independently support the sentence imposed” (*State v. Loomis*, 2016, para. 99). This requirement helps maintain the integrity of the process by ensuring that the sentence can be supported without relying on the risk assessment.

6.1.2 Relation to Broader Legal Context

Weight:

The decision in *State v. Loomis* works to dilute the weight of a COMPAS report so that presiding judges can make their own evaluation. The latticework of restrictions, limitations, and cautions – as introduced under the “instruction” factor – completely subjugate the risk assessment to the discretion of the judge. This subjugation is reflected in the Court’s statement, “[p]roviding information to sentencing courts on the limitations and cautions attendant with the use of COMPAS risk assessments will enable courts to better assess the accuracy of the assessment and the appropriate weight to be given to the risk score”(*State v. Loomis*, 2016, p. 66). This sentiment is echoed against later as the Court writes, “[j]ust as corrections staff should disregard risk scores that are inconsistent with other factors, we expect that circuit courts will exercise discretion when assessing a COMPAS risk score with respect to each individual defendant” (*State v. Loomis*, 2016, para. 71). While these statements diminish the weight of COMPAS risk reports, they also reflect a meaningful opportunity for judges to disagree with the results and put forward their own evaluation of its accuracy. Ultimately, The Wisconsin Supreme

Court's treatment of COMAPS risk assessments render them subordinate to judicial discretion and unable to independently sustain legal action.

Appeal

The *State v. Loomis* decision makes few changes to the existing appeal structure surrounding sentencing hearings. Just as before, the defendant can appeal the sentencing decision to a higher court, which includes appealing the judge's treatment of the COMPAS report. The most concrete changes that *State v. Loomis* creates surround the restrictions on how COMPAS reports can be used as defendants have more information on what grounds they can appeal the court's treatment of a risk assessment (e.g. used to determine severity of punishment, outcome not independently supported by other evidence).

By not substantially changing the appeal structure, the Court indicates an increased reliance on the adversarial nature of judicial proceedings to deal with the complexity introduced by evolving risk assessment technologies. Here, the "adversarial nature" of the U.S. court system refers to the belief that the best outcome will be found when each opposing party has the opportunity to put forward the arguments that best support their preferred outcome within a system of formal rules (Calhoun, 2002; Kagan, 2009, p. 9). From this perspective, the court delegates the proper interpretation of the risk report to the defense and prosecution as it writes, "[a]lthough Loomis cannot review and challenge how the COMPAS algorithm calculates risk, he can at least review and challenge the resulting risk scores in the PSI" (*State v. Loomis*, 2016, para. 53). As there is no way to directly appeal COMPAS' results, the only way for the defense or prosecution to immediately challenge the results is by putting forward arguments for its interpretation during the sentencing hearing itself.

6.1.3 Structures within the COMPAS-Court Dynamic

The treatment of COMPAS reports in sentencing hearings by the Wisconsin Supreme Court holds a few insights into how the human-AI decision-making relationship is structured to counteract the potential harmful decision-making practices that the program can contain. First, the Court deals with the novel considerations surrounding AI input into judicial decision-making by subordinating the COMPAS report to professional judgment. Instead of focusing on the process or substantive calculations behind a COMPAS report, the Court made a decision to allow the reports in their entirety and trust in the adversarial process and discretion of the judge to identify and address any issues with the report as they manifest themselves in the case. This decision reflects confidence in pre-existing structures of due process to adapt to the novel and somewhat alien decision-making practices of AI.

However, this trust in professional judgment and judicial discretion comes with qualifications which are meant to limit the impact any harmful decision-making practices hold. The Court specified many conditions under which COMPAS reports can be used through restrictions, guidelines, and permissible uses. Cumulatively, these qualifications position the risk assessments as a merely supportive role and prohibit their use as a determinative factor. These qualifications are meant to restrict or prevent any harmful decision-making practices behind a COMPAS report from having too strong of an impact on sentencing decisions by limiting how much the judge can rely on it. By fully subordinating the risk assessment to human discretion, the Wisconsin Supreme Court indicates a trust that judges will be able to properly compartmentalize the COMPAS report's findings so that they only support the permitted uses.

Third, subjugating the COMPAS reports to human discretion allows the existing appeal structure to be maintained. While it can be argued that the Court's decision created slight modifications to appealing sentencing decisions by clarifying on what grounds the defence or prosecution can contest the improper use of a risk assessment, the overarching appeal structures remain intact. That is, there are no structures introduced that are specifically designed to deal with the novel considerations surrounding AI decision-making. Rather, appealing a sentencing decision is still reduced to exercises of human discretion.

Fourth, the prominence of COMPAS reports are somewhat diluted as they are presented alongside human determinations. As the PSI contains the parole officer's evaluation of the defendant and the COMPAS report, the risk assessment is presented alongside other information which can provide more context, perspective, and information for the sentencing hearing. Consequently, the impact of the COMPAS report is somewhat muted in comparison to being presented independently. However, this consideration does not compensate for the impact that COMPAS' clear and easily readable bar graphs have on its prominence relative to the text-based information found in the rest of the PSI.

And finally, the COMPAS report is introduced at the beginning of a substantial process that involves further exercises of human discretion. Specifically, the PSI is presented at the beginning of the sentencing hearing where the prosecution and defense have many opportunities to argue about how the COMPAS report should be received. This also takes place in the criminal justice context which support more robust procedural requirements because of how significantly the decision affects the defendant's interests. While there is no opportunity to appeal the COMPAS report's determinations or reasoning, there is an immediate opportunity to immediately challenge them throughout the sentencing hearing.

6.1.4 Unaddressed Factors in the COMPAS-Court Dynamic

The human-AI decision-making relationship established by the Supreme Court of Wisconsin involves various trade-offs that leave some concerns unaddressed. First, the COMPAS report's lack of transparency, combined with the reliance on judicial discretion, creates opportunities for human biases to express themselves. As the human actors cannot engage with the COMPAS program's "reasoning", it is possible that the risk report will become the object of confirmation bias and only agreed with when it supports the judge's conclusion and will otherwise be discarded. This opacity also creates an opportunity for judicial trust or suspicion in AI to play a role as, without any reasons to support the risk scores, their persuasiveness can be heavily determined by how much the judge trusts or distrusts the program that created them.

The sequence in which the COMPAS report is presented can also be problematic. As mentioned previously, the prominence of the risk assessment is diluted because it is presented alongside human-generated information and its potential harmful impacts are controlled through its presentation at the beginning of a robust process. Despite these structures, there are still unaddressed concerns about the sequence as the COMPAS report is presented at the beginning of deliberations which creates an opportunity for the document to anchor the following discussion. In its decision, the Court did not consider the influence which the COMPAS report's clear and quantified measurements of risk can have when presented at the beginning of the decision-making process.

6.2 European Union's iBorderCtrl

The iBorderCtrl project aims to develop a comprehensive solution to "enable faster and thorough border control for third-country nationals crossing the land borders of EU Member States" through the use of technology (iBorderCtrl, 2016c). In this context, "land borders" refers to road, walkway, and train stations (iBorderCtrl, 2016c). As it is only a research project, iBorderCtrl has not been the subject of comprehensive analysis at the level needed for the analytic framework; consequently, this thesis will primarily rely on the official documentation to introduce and establish the project.

The iBorderCtrl system relies on many different technologies to evaluate people crossing the border, not all of which use machine learning technology. There are eight "modules" that constitute the overall solution:

- The **Automated Deception Detection** System "performs, controls and assesses the pre-registration interview by sequencing a series of questions posed to travellers by an Avatar".
- The **Biometrics Module** validates the identity of the traveller through fingerprint and palm vein technologies.
- The **Face Matching Tool** creates a biometric signature for the traveller through video and photo images which is then used to verify the traveller's identity during pre-registration and border crossing.
- The **Document Authenticity Analytics Tool** checks the validity of travel documents (e.g. passports, visas) presented at both pre-registration and border crossing.

- The **Hidden Human Detection Tool** supports border guards in "searching and detecting hidden people inside various vehicles", especially those in closed compartments.
- The **External Legacy and Social Interfaces** system helps the iBorderCtrl system crosscheck traveller information with older systems and databases.
- The **Risk Based Assessment Tool** "aggregates and correlates the risk estimations received by the processing of the travellers' data and documents" into a risk score that is presented to the border guard. This module also identifies "cases that deserve further investigation".
- The **Integrated Border Control Analytics** Tool works to analyze the data generated by the iBorderCtrl system to improve its adaptability by identifying new patterns, create feedback by evaluating effectiveness, improve decision-making by studying false acceptances or rejections, and provide traffic projections from traffic data (iBorderCtrl, 2016b).

The program itself is divided into two stages. The first "pre-screening" stage consists of an online application where the traveller uploads "pictures of their passport, visa and proof of funds" and also informs them of their rights, travel procedure, provides advice, and gives alerts "to discourage illegal activity" (European Commission, 2018b). This pre-screening step also uses a webcam to record the traveller's responses to questions asked by a computer-animated border guard; these responses are then processed by the Automatic Deception Detection System to "[quantify] the probability of deceit…by analysing interviewees non-verbal micro expressions" (European Commission, 2018b; iBorderCtrl, 2016b). At this phase, travellers are alerted to any errors in document preparation and collection to provide an opportunity for correction prior to crossing the border (iBorderCtrl, 2016b). If the iBorderCtrl system identifies a potentially criminal crossing, then this information is withheld from the traveller who is flagged to border guards for further evaluation (iBorderCtrl, 2016b). Once this stage is completed, the traveller receives a QR code that they present to the border guard at the next stage (iBorderCtrl, n.d., p. 2).

The second "border crossing" stage involves further checks conducted by the iBorderCtrl system and the border guard. This stage involves another evaluation of the physical documents that were digitally imaged and uploaded during the pre-registration phase, but is mostly limited to "validating that indeed the originals contain the same information as what was collected at pre-registration" (iBorderCtrl, 2016b). The traveller further undergoes various biometric checks to validate their identity and, if travelling in a vehicle, an additional check for any hidden humans inside the vehicle (iBorderCtrl, 2016b). Once the checks are completed, the iBorderCtrl platform then "calculates an evaluation score, taking also into account the evaluation score calculated at the pre-registration phase, producing a total evaluation score for each traveler" (iBorderCtrl, n.d., p. 3). The total evaluation score is then presented to the border guard who chooses between one of three outcomes for the traveller: pass, no pass, or further control (iBorderCtrl, n.d., p. 3). Importantly, the border guard is only presented with the total evaluation score, and not all of the scores produced by each individual module to prevent any single determination from exerting disproportionate influence over the outcome (iBorderCtrl, 2016b).

As mentioned previously, the iBorderCtrl system is only a pilot project and does not hold any legal weight. Rather, participation is completely voluntary as travellers are invited to cross through a simulated iBorderCtrl border check process (iBorderCtrl, 2016b). This simulated crossing only occurs once the participant has successfully crossed the border under standard procedure and is prohibited from positively or negatively affecting the standard border crossing process (iBorderCtrl, 2016b, 2016c). Despite iBorderCtrl's status as a pilot project, it still holds relevance as it was designed as if it were operating under the actual legal framework, as discussed at the beginning of section 6.

The iBorderCtrl system will be examined in relation to two human actors: the traveller and the border guard. The traveller is the individual who participates in the iBorderCtrl project and is positioned as the decision subject. The border guard acts as the decision-makers as they are presented with the total evaluation score and are responsible for making the determination about whether the traveller will proceed, be barred from entry, or require additional control. Each stage involves a different operator, with the traveller acting as the operator in the first stage and the border guard in the second. With an overview of iBorderCtrl and its related human actors established, the following subsection will apply the analytical framework to it.

6.2.1 Relation to Human Decision-Makers

Distance:

iBorderCtrl is quite distant from human decision-making as it operates independently from human intervention once it receives the required input. Most immediately, the distance is represented in the highly automatic nature of the many modules that constitute iBorderCtrl. Many of the modules rely on static inputs – such as scans of documents, basic traveller information, or facial images – that the operator has little control over the presentation of. Even the Automated Deception Detection System, which gives the traveller a high degree of freedom in how they answer the questions, is not analyzing the substance of the responses but rather searching for involuntary facial movements. This creates a substantial degree of distance as the human operator has very little control over the form and presentation of the input.

The distance between human and AI decision-making is also reflected in the division between iBorderCtrl and the border guard's analysis. First, a lot of the traveller's evaluation takes place during the pre-screening stage which occurs before they encounter the border guard. Second, the first steps at the border crossing phase involve further static inputs that are facilitated by the border guard. Despite this interaction between the two parties, the border guard and iBorderCtrl remain quite distant as there is no opportunity for substantive co-determination of results because the border guard has very little capacity to influence the provided input.

Sequence:

The iBorderCtrl system seems to present its determinations at the beginning of the border crossing phase. As stated in the project's technical description, "…all information of the traveller gathered during the pre-registration phase is now available to the Border Guard with iBorderCtrl bringing all analytic results from each technology together to identify risks to the agent that support him in both an overall evaluation of the traveller" at the border crossing stage. This quote indicates that the determinations of the iBorderCtrl system's first stage are made immediately available to direct the border guard's evaluation of the traveller. While the guard will have the opportunity to form their own evaluation of the traveller, this takes place after they receive the results from the pre-registration phase.

Transparency:

The iBorderCtrl system does not have a high degree of transparency towards the border guard. As mentioned in the description, all the outputs of each different module are not presented individually, but rather amalgamated into a single evaluation score that represents the traveller's risk through the Risk Based Assessment Tool. This amalgamated score is designed so that, in the case of an erroneous reading, no single output will exert a disproportionate influence over the border guard's decision (iBorderCtrl, 2016b). If the system was more transparent by presenting each module's score, then there would be a risk that the border guard would focus on one outlier score even if all of the others indicated an opposite level of risk. Amalgamating all of the scores through the Risk Based Assessment Tool allows each module's output to be given a consistent weight and translate any individual high risk scores into directions and guidance for the border guard's investigation. However, this high degree of opacity means that the border guard cannot critically engage with the reasoning behind the evaluation scores.

Further opacity is found in the traveller's lack of access to iBorderCtrl's results. While some basic information is provided to the traveller, such as documentation errors that they can correct before crossing the border, the traveller is unaware of what risk score iBorderCtrl assigns them. Without information about iBorderCtrl's conclusions and reasoning, there is very little opportunity for the traveller to understand how their evaluation score affects the border guard's decision.

Instruction:

While there is no mention of it in the documentation provided by the iBorderCtrl project, it is reasonable to expect that border guards will be given enough training to become familiar with the program and learn how it operates. However, as no specifics are provided, the instructions given to human decision-makers cannot be examined in-depth.

6.2.2 Relation to Broader Legal Context

Weight:

The weight and significance of the iBorderCtrl system's determinations is multi-faceted and somewhat internally contradictory. Formally, the output of the iBorderCtrl system is completely subject to the discretion of the border guard, who holds the ultimate decision over whether the traveller passes, is refused entry, or proceeds to further inspection. This is reflected in the statement that "iBorderCtrl is a human in the loop system and the Border Guard will use his/her experience in making the final decision" (iBorderCtrl, 2016b). This sentiment is further reinforced as the project website writes, "iBorderCtrl provides key technology to the border guards both integrated to existing static installations, as well as a portable hardware platform that empowers -through technology- the border guard" (emphasis removed) (iBorderCtrl, 2016b). A lot of confidence is placed in the border guard's ability to properly consider iBorderCtrl's output and disagree if necessary

However, these statements of trust in human decision-making and empowerment are contradicted in the specific objectives of the iBorderCtrl project. Specifically, one objective of the project is "[t]o **reduce the subjective control and workload of human agents and to increase the objective control** with automated means that are non-invasive and do not add to the time the traveller has to spend at the border" (emphasis in the original) (iBorderCtrl, 2016c). In contrast to the previous statements, this objective indicates that the project is being designed to detract from subjective evaluations by border guards and promote more

This leaves the weight of iBorderCtrl's determinations in a vague position. Formally, the border guard retains full decision-making authority to choose between the three possible outcomes of the border crossing process. However, the system itself puts forward very strong determinations without many opportunities for the border guard to interrogate them. Immediately, it would seem difficult for a border guard to significantly depart from iBorderCtrl's determinations; if the system finds that a traveller is a high risk, it doesn't seem likely that the border guard will be able to let them pass without facing some sort of negative consequence, such as being questioned by supervisors or increased scrutiny. However, beyond a statement of likelihood, the issue of how a border guard considers and negotiates with the system's determinations is too nuanced for this paper and will be better served through a more focused study.

Appeal:

The iBorderCtrl project would make it difficult for the traveller to appeal the system's determinations about them. As the traveller is not given information about how iBorderCtrl has evaluated them, the traveller does not have access to any substantive conclusions or reasons that would affect their passage across the border. As a result, the traveller is not well-positioned to identify and challenge any harmful decision-making practices made by iBorderCtrl.

However, the iBorderCtrl project website does not make any mention of modified appeal procedures. This is likely the result of the system's status as a pilot project that cannot be appealed because it carries no legal weight. As appeal systems are not specified in the project description, factor will be excluded from later consideration to avoid speculation.

6.2.3 Structures Within the iBorderCtrl-Border Guard Dynamic

In its design, the iBorderCtrl system compensates for potential harmful outcomes through three structures in the human-AI decision-making relationship. Having reviewed the iBorderCtrl program, this subsection will now review the significance of the paramountcy of human discretion, parallel determinations, and deliberate opacity.

Similar to COMPAS, the results of the iBorderCtrl program are made completely subject to human discretion. As emphasized at multiple points in its documentation, the iBorderCtrl project makes the human border guard responsible for the ultimate decision of whether the traveller proceeds, requires further control, or does not pass. While somewhat contradicted in the documentation, iBorderCtrl is positioned as supporting human deliberation by providing more information, rather than as taking away decision-making power from the border guard. This reflects an implicit trust in human discretion to identify and compensate for any harmful decision-making practices.

The iBorderCtrl program also involves parallel determinations as both the human and program are tasked with making the same evaluation of the traveller's risk. These determinations are not completely isolated from each other as the iBorderCtrl program's results are presented to the border guard at the beginning of their evaluation. However, the iBorderCtrl project's use of parallel decision-making, where both the border guard and program make the same decision, allows the human to reconcile their own conclusions with the determinations made by the program.

And finally, iBorderCtrl program uses deliberate opacity to control for any potential harmful decision-making practices. Rather than present the border guard with all the risk scores generated from each separate module, the program only presents an amalgamated score. While transparency is valuable, iBorderCtrl demonstrates how opacity can be deliberately used to dilute the influence any harmful decision-making practices can have over the final decision.

6.2.4 Unaddressed Factors in the iBorderCtrl-Border Guard Dynamic

First, reducing the legal weight of iBorderCtrl's findings by subordinating them to human discretion creates space for confirmation bias and the border guard's trust and suspicion of automation to influence the outcome. As the border guard is trusted with properly accounting for iBorderCtrl's evaluation scores in their final decision, there is a lot of space for the border guard to use the evaluation scores as evidence to support their own conclusions and discard them when the scores indicate differently. The degree to which the iBorderCtrl evaluation score is relied on can also greatly vary between travellers as border guards may hold different levels of trust or suspicion towards iBorderCtrl. The unaccounted factors of trust and confirmation bias both undermine iBorderCtrl's goal of increasing objective evaluations and reducing the influence of subjectivity.

The sequence of the iBorderCtrl process can also undermine the trust placed in human discretion to control for harmful outcomes. As the border guard receives the iBorderCtrl evaluation score at the beginning of their interaction with the traveller, there is a possibility that the evaluation score will act as an anchor or default that can influence the border guard's evaluation of the traveller. Together, there is a risk that the border guard would be hesitant to exercise their discretion against the iBorderCtrl program in a meaningful capacity. These concerns are amplified through the deliberate opacity that the amalgamated evaluation score holds as the border guard is prevented from understanding how the total score was generated and the reasons behind it.

6.3 The Los Angeles Police Department's PredPol Predictive Policing System

PredPol is a "predictive policing" technology which predicts where crime is likely to happen in the future. With "predictive policing" defined as "the practice of identifying the times and locations where specific crimes are most likely to occur", PredPol claims that it can help police departments more efficiently allocate police resources as well as prevent crime by enabling the pre-emptive deployment of officers (PredPol, 2018). The system operates by identifying "hotspots", which are 500 square foot areas that indicate where crime is likely to occur (Smith, 2019, p. 25). These hotspot maps are valid for twelve hours and, with two maps generated a day, PredPol provides predictions for the entire day. At the beginning of their shifts, officers are given a map of ten to twenty hotspots that they are encouraged to visit when not otherwise occupied with radio calls or other police-related duties (CBC News, 2018, t. 0:00:30; Smith, 2019, p. 25; WIRED, 2018, t. 0:05:27). A sample hotpot map can be found in Appendix D.

In order to create crime predictions, PredPol uses the police department's records to train their machine-learning algorithm for the jurisdiction it will be applied to. Working with the police agency's records management system, PredPol uses five data points from each recorded incident to train its machine learning algorithm:

1. **Incident Identifier:** a unique identifier that the department uses for each recorded crime.
2. **Crime or Event Type:** "the violation code and/or crime description assigned to a particular incident type" as used in the records management system.

3. **Location of Incident:** The latitude and longitude at which the incident occurred or, if unavailable, then the complete address of the incident (street number, street name, city, state/region).
4. **Timestamps with Start and End Date/Time for Incident:** The time range in which an incident occurred. As the exact time which an incident occurred may not be known, PredPol accepts a range of times (e.g. "between midnight and 8 AM") with time spans of more than 72 hours being excluded.
5. **Record Modified Date/Time for Incident:** an optional field which is used in case a record has been changed because, for example, the crime code has been reclassified (PredPol, 2018).

While PredPol accounts uses a total of five data inputs, it claims that only three are focused on. As the incident identifier is only used to uniquely identify an event, and the record of a modified incident is optional, PredPol focuses on the crime type, location and date and time to train its machine learning algorithm (PredPol, 2018).

Since its implementation, PredPol has come under considerable scrutiny and protest. There are prominent concerns that PredPol has been trained on historical crime data that is the product of biased police practices, which are then revived through the program's operation (Degeling & Berendt, 2018, p. 353). After analyzing PredPol, Lum and Isaac concluded that it disproportionately directs officers towards poor black neighbourhoods and holds the risk of creating a "feedback loop" where continued policing of an area leads to more crime data generation, which then directs increased police patrols to that area (Lum & Isaac, 2016, p. 19). Further, there is the risk that "using a computer to allocate police attention shifts accountability from departmental decision-makers to black-box machinery that purports to be scientific, evidence-based and race-neutral" (Lum & Isaac, 2016, p. 19). In response to these concerns, some studies have found that while arrests increased, there was "no significant difference in the arrest proportions of minority individuals between treatment and control conditions" (Brantingham, Valasik, & Mohler, 2018, p. 5). In other words, the amount of "minority individual" arrests as a percentage of overall arrests was the same for areas within and outside those identified by predictive policing (Smith, 2019, p. 27). However, these issues are far from resolved and concerns about the potential biases that predictive policing can promote continues to be prominent elements of public concern and debate (Puente, 2019; Smith, 2018; S. Thompson, 2018).

In operation, PredPol holds the police officer in a decision-making relationship. As mentioned previously, PredPol generates a map of hotspots that are valid for twelve hours. These maps are given to officers at the beginning of their shifts so they know where to go when not otherwise occupied. As the officer interacts directly with PredPol's output – the hotspots maps – they are in the position to decide whether they follow the predictions made.

The decision-subject is a bit more difficult to identify as the PredPol system does not deal with specific individuals. In contrast to the previous two systems where there was an easily identifiable person who was being evaluated, the PredPol system operates with a much broader perspective. Under this broad perspective, PredPol's decision subject will be defined as the communities that PredPol evaluates for crime risk.

6.3.1 Relation to Human Decision-Makers

Distance:

There is a lot of distance between humans and the PredPol system which is found in its largely autonomous operation, dispersed sources of human input, and the static nature of its input. First, PredPol generates its hotspot maps autonomously as it pulls data from the police department's record management system and automatically integrates it into the predictions. Second, the input which PredPol operates focuses on – crime type, location, and date - is fairly static as there is not a lot of room for interpretation in how this information is entered. As PredPol writes, these are "the three most objective data points collected by police departments" (PredPol, 2018). And finally, PredPol operates at a considerable distance from humans through the dispersed sources of human input as there is no single operator, but rather receives input from all of the police agency's officers as they submit incident reports. These three factors create a significant amount of distance between PredPol and the officer that receives the hotspot map.

Sequence:

In the decision-making process of where a police officer patrols, the PredPol program holds prominence as it is presented at the beginning of each officer's shift. As officers receive a map of hotspots at the beginning of their shift, officers receive PredPol's determined areas of interest before they are in a position to exercise their own discretion about where to patrol.

Transparency:

PredPol does not communicate the reasons behind its hotspot determinations to the officer. The officer seems to only receive the map of hotspots without an explanation of why the hotspot was identified by PredPol. Further opacity is found in the lack of public access as hotspot maps are not available to the public as well.

Instruction:

Officers receive some training to properly interact with the PredPol system. As described in the Office of the Inspector General's report, police officers can indicate when they are patrolling a hotspot so that the duration of their patrol can be recorded to monitor the effectiveness of the system (Smith, 2019, pp. 25–26). This indicates that officers have at least a basic orientation to the PredPol software and a broad overview of how it works. However, there is no documentation of more rigorous training on how police officers should interpret PredPol's hotspot maps and when they should exercise their discretion to act against its predictions.

6.3.2 Relation to Broader Legal Context

Weight:

PredPol's decisions holds very little formal weight in officer decision-making. As there is no requirement that officers visit their assigned hotspots, PredPol's output holds no formal weight in officer decision-making as they are completely subject to the officer's discretion. In their study of PredPol, Brantingham, Valasik, and Mohler write that officers are "encouraged to patrol target areas during any available discretionary time" (Brantingham et al., 2018, p. 2). This is echoed in a statement from Jeff Nolte – the Los Angeles Police Department Captain in 2018 – who emphasized that "[PredPol is] not telling us what to do" (S. Thompson, 2018). While the question of how much weight PredPol's hotspot maps hold in practice, this would require additional study that is beyond the scope of this paper.

Appeal:

Challenging a PredPol decision can take place on two levels, both of which would be difficult to succeed in. First, there is the option of challenging the use of PredPol's hotspot map itself. As hotspot maps are subject to officer discretion, this would be quite difficult as "[l]aw enforcement officials have tremendous discretion to determine the amount and style of policing that occurs in their jurisdiction" (Miller, 2015, pp. 521–522). Second, a challenge could be launched against the program itself. However, to be successful, this challenge would have to prove "direct evidence of purposeful discrimination" and cannot succeed based solely on evidence of "a significant risk of racial bias" (Baldus, Woodworth, & Grosso, 2007; D. Thompson, 2019). As the appeal structure has not changed to accommodate the introduction of PredPol, it seems like any legal challenge against its use would be very difficult.

6.3.3 Structures Within the PredPol-Police Officer Dynamic

The LAPD's implementation of PredPol presents three strategies of dealing with potential harmful decision-making practices that the system produces. This subsection will discuss how the system is subordinated to human discretion, is placed at the beginning of a substantial process, and given a low priority in comparison to other police officer responsibilities.

Just like the previously analyzed systems, the input that PredPol provides to patrol decisions is subject to the complete discretion of the police officer. As there is no requirement for officers to patrol hotspots during their shifts, PredPol's hotspot maps are positioned to support, but not detract from, officer decision-making.

Further, the significance of PredPol hotspot maps are also diminished through its placement at the beginning of a process that accommodates many further exercises of human discretion and escalation. While decisions about where to patrol are significant – especially on the systemic level that PredPol operates on – they are decisions that precede more consequential exercises of discretion such as whether to question, detain, or arrest an individual. Each of these following exercises of discretion involve their own set of checks and balances which present further opportunities to compensate for any faulty decision-making that occurred earlier.

And finally, PredPol decisions are given a low priority in relation to other policing duties. As mentioned previously, officers are only directed to patrol their assigned PredPol hotspots when not otherwise occupied and at their own discretion. This limits the impact which PredPol can have on the LAPD's existing police practices as it cannot override the established system surrounding patrol decisions.

6.3.4 Unaddressed Factors in the PredPol-Police Officer Dynamic

In implementing PredPol, the Los Angeles Police Department has not addressed concerns about confirmation bias and anchoring, systemic impacts, and the limited ability for affected individuals to contest PredPol predictions. First, the LAPD does not consider how PredPol can shape officer expectations. Most immediately, hotspot maps create a risk that officers will be more likely to make consequential law enforcement decisions against people within a hotspot as they could be looking for behaviour that justifies PredPol's prediction. Further, hotspot maps can act as an anchor for patrol decisions where, even if the officer finds reason to disagree with the predictions, may not stray far from where they were initially directed to. And finally, the hotspot map can exert a strong influence through order effects as it is introduced at the beginning of the officer's shift, which can consequently undermine the trust placed in officer discretion. PredPol creates the risk for numerous biases and heuristics to express themselves and shape officer behaviour which have not been addressed in the available information.

Second, the emphasis on officer discretion to control for negative outcomes does not fully address the potential for broader and more dispersed harmful decision-making practices that operate at the systemic level. Many of the concerns surrounding PredPol operate at the systemic level, such as the revival of historical biases in police practices and the potential for feedback loops to keep targeting disadvantaged communities. However, individual human discretion is not well-positioned to identify and counteract the more subtle and dispersed channels that these broader concerns can manifest in.

And finally, the diminished legal weight makes it very difficult for citizens to challenge its use. As mentioned previously, it is very difficult for citizens to successfully challenge how police offers exercise their discretion which, without an indication otherwise, would likely apply to how a police officer uses PredPol's predictions. As it remains unchanged, this appeal context does not address the unique ways in which PredPol can shape officer discretion.

7. Discussion

Having finished the case studies, this section will discuss the insights and findings that can inform future human-AI decision-making relationships. The first subsection will examine common themes found across all of the cases while the second subsection discusses the relevance of these themes to the Canadian context. The third subsection will look at possible structures that can address and help mitigate the concerns surrounding human-AI decision-making relationships, while the final subsection will discuss future opportunities for research.

7.1 Common Themes Found Across Case Studies

In all the reviewed cases, there are four shared features in their human-AI decision-making relationship: subordinating AI determinations to human discretion, not addressing the ways in which AI programs can influence human decision-making through biases and heuristics, trust in human discretion to control for harmful decision-making practices that AI programs can introduce at both the immediate and broader context, and unmodified appeal structures. The rest of this subsection will discuss the significance behind each of these common themes.

The most prominent shared feature is that all examined AI systems have been subordinated to human discretion to control for potential harmful outcomes. By making the AI system's decisions non-binding on the decision-maker, each human-AI decision-making relationship has refused to give the AI system's conclusions any legal weight. Immediately, this prevents AI systems from inflicting large-scale harm as their impact is limited by the speed of human decision-making and how the human decision-maker accounts for the AI's determinations. By subordinating the AI program to human discretion, the benefits and risks that the AI program can introduce are both confined to the limits of human decision-making.

Second, all the cases have placed trust in human discretion to identify and address harmful decision-making practices in AI systems without accounting for how the introduction of AI can influence the exercise of human discretion. Throughout the cases, there were many spaces in which biases and heuristics could express themselves to influence how human actors evaluate AI outputs and make decisions. Even though AI has the potential to significantly change human decision-making, the reviewed human-AI decision-making relationships regard the nature of human discretion as unchanged and the nature of AI-generated information as similar to other information that the decision-maker would otherwise consider.

The trust in human discretion also limits the ability to control for systemic harms that manifest at larger scales. Many of the concerns expressed about harmful decision-making practices in AI are directed at negative outcomes that can only be identified at the systemic level, such as feedback loops, reviving historical biases, or disparate treatment based on protected characteristics. While human oversight can play a role in controlling for negative outcomes, this does point to the need to have additional structures of oversight and reflexivity to fill in the gap that individual perspectives cannot fully address.

And finally, the cases did not introduce new ways of appealing the AI system's determinations. Despite the concerns about transparency, bias, and other harmful decision-making practices that surround AI systems, the decision-subject had no way of directly challenging or examining the AI's reasoning. Rather, the focus of appeal decisions remains on the exercise of human discretion and, with the introduction of AI, how the human actor has accounted for the AI's output. As successfully challenging exercises of discretion is quite difficult, the unmodified appeal structures leave little room for decision subjects to appeal the use of AI programs.

In the tension between the benefits and risks of AI decision-making, these early cases have demonstrated a tendency towards prioritizing risk reduction through familiar decision-making structures over the benefits and risk that come with more autonomy for AI systems. By tethering AI decision-making to human judgment, many of the risks become contained within the more familiar problems of human decision-making that existing procedural and substantive review structures are more comfortable with. However, this also means that the benefits of implementing AI – efficiency, consistency, and improved access to justice – are somewhat muted as the decision-making still operates on human discretion.

7.2 Relevance to Canadian Administrative Contexts

The Canadian federal government's approach to human-AI decision-making relationships is largely encapsulated by the Treasury Board Secretariat's Directive on Automated Decision-Making ("the Directive"), which was briefly mentioned in section 3(d).This directive creates scaling standards based on the impact of the decision that the automated decision system contributes to (Government of Canada, 2019, sec. Appendix C). These requirements include many components and concerns that have been found in the case studies, such as notice, explanations, testing and monitoring outcomes, data quality standards, peer review, employee training, contingency systems, and human intervention (Government of Canada, 2019, secs. 6.2, 6.3). While these requirements represent a promising start, they do not fully address the concerns surrounding human-AI decision-making relationships discussed in this thesis.

Much like the case studies, the Canadian federal government has also expressed a reliance on human discretion. For the most consequential decisions (levels III, and IV), "[d]ecisions cannot be made without having specific human intervention points during the decision-making process; *and the final decision must be made by a human*" (emphasis added), while "[d]ecisions may be rendered without direct human involvement" for lower impact decisions (levels I and II) (Government of Canada, 2019, sec. Appendix C). Similar to the cases, the Directive indicates that human discretion will ultimately be responsible for monitoring and counteracting harmful decision-making practices behind AI decision-making. However, while the case studies were generally unclear about the training that human decision-makers must go through, the Directive imposes high training requirements on human actors where initial training is required for level III systems and re-occurring training is required for level IV systems (Government of Canada, 2019, sec. Appendix C).

Training can do a lot of work to address the concerns surrounding human-AI decision-making relationships, but the Treasury Board Secretariat's approach does not fully account for the many ways in which the structure of the relationship can influence outcomes. While training can help decision-makers recognize their own biases and engage with AI decisions, there is still the risk that human discretion can be influenced by the various structures that define how human decision-makers interact with AI programs. For the many unaddressed factors that were revealed through the case study, it is important that any trust in human actors to control for harmful decision-making practices is accompanied by an understanding of how that role is structured to promote a positive human-AI dynamic.

7.3 Structuring Effective Decision-Making Relationships

Having analyzed existing human-AI decision-making relationships and finding the current reliance on human discretion to control for harmful outcomes, this last subsection will discuss some decision-making structures that present different trade-offs between controlling for harmful outcomes and realizing the benefits of AI technology.

Parallel Determinations:

Similar to the structure found in the iBorderCtrl program, AI programs can be used to make the same determination as a human decision-maker. In this structure, efficiency can be gained as both the human and AI make the decision with a reduced process and, if both come to the same conclusion, then the case is resolved. The initial reduction in process is compensated for by an automatic appeal to a full and robust decision-making process that is automatically triggered if the human and AI come to different conclusions.

Training:

A common theme throughout the cases is that the human decision-maker is ill-equipped to fully engage with the AI program's output, whether because the AI's reasoning is not made available or because the human is unaware of how the AI program can activate biases and heuristics. An immediate response would be to provide training to human decision-makers so that they are better equipped to critically analyze AI decisions as well as recognize and compensate for any biases and heuristics that can be activated.

Staggered Introduction:

Order effects were a consistent concern across the cases as AI outputs were often introduced at the beginning of the human decision-making process. In cases where AI programs support human discretion, order effects can be mitigated by introducing AI decisions in the middle of human decision-making processes. This structure still allows AI programs to support human decision-making while mitigating the potential for the program to activate certain biases and heuristics.

Dilution:

Similar to order effects, the manner in which AI decisions are presented can also impact how much weight they are given by human decision-makers. As discussed in the COMPAS case study, the impact of AI determinations can be "diluted" by presenting it alongside other information. In decision-making contexts, the trust or suspicion which human decision-makers hold towards AI can be mitigated by blending AI determinations with other sources of information.

Co-Determination

Decisions can consist of a variety of sub-determinations that are more suited for human cognition or AI processing. Co-determination seeks to find the parts of a decision that are best suited for AI determination to alleviate the workload placed on human actors. Here, efficiency and consistency would be gained because the AI program would be deciding part of the case with legal weight behind its determinations so that further human discretion is not required. Once the human actor and AI program have made their determinations, a further system would be required that would govern how they would be combined to form the final decision.

Escalation:

In this context, “escalation” refers to placing AI programs at the beginning of a decision-process that is followed by easy access to more robust processes. In this context, the AI program would be responsible for issuing a full determination on the case that they are presented with which would be legally valid. However, the significant amount of legal weight ascribed to the AI’s decision would be balanced against a robust appeal process that can be easily accessed. At the very least, this approach would filter out all of the cases that would be decided in the decision subject’s favour (as they would be unlikely to appeal a positive outcome) and only leave negative outcomes for human deliberation. This structure requires that the decision-subject is fully informed of their ability to easily appeal the decision. Further, a substantial effort should be made to reduce all barriers (e.g. cost, time, effort) associated with appealing the case. However, this structure still holds the risk that human decision-makers would hold the presumption that any matter they are hearing has already been rejected once which could influence their decision against the decision subject.

7.4 Future Steps

This thesis has worked to put research on cognitive biases and heuristics in conversation with the requirements of Canadian administrative law and the concerns surrounding bias and transparency in AI. At this broad level, this thesis could only identify areas where biases and heuristics could be expressed and affect the role of human decision-makers in controlling for harmful decision-making practices in AI. Consequently, some immediate next steps would be to conduct research on the degree to which biases and heuristics affect human decision-making when interacting with AI programs. More specific research into biases and heuristics in human-AI decision-making relationships can inform and support further work on structuring human-AI decision-making relationships to promote positive processes and outcomes.

8. Conclusion

This thesis has examined early implementations of human-AI decision-making relationships through a qualitative case study. Beginning with a review of administrative bodies, artificial intelligence, and how the qualities of AI can support administrative decision-making, this thesis then went on to survey the duties and obligations imposed on administrative decision-makers and how they are in tension with concerns about bias and transparency in AI programs. As the Canadian federal government indicated that human discretion can be used to control for harmful decision-making practices, the following section looked at how biases and heuristics can undermine this approach. The case study revealed that COMPAS, iBorderCtrl, and PredPol all subordinated AI output to human discretion, and trust the human actor to correctly weigh and evaluate the AI's determinations. Not only did these decision-making relationships not account for the ways in which AI can shape the exercise of discretion, but also limited the benefits that could be obtained from AI capabilities. In response to these unaddressed factors, this thesis concluded with a discussion of how further research could build upon this thesis, as well as how human-AI decision-making relationships can be differently structured to more effectively mitigate risk and realize the benefits of AI in administrative contexts.

References

Alpaydin, E. (2016). *Machine Learning: The New AI*. Retrieved from http://ebookcentral.proquest.com/lib/utoronto/detail.action?docID=4714219

Angwin, J. (2016a, February 1). *Sample COMPAS Core Risk Assessment*. Retrieved from https://assets.documentcloud.org/documents/2702103/Sample-Risk-Assessment-COMPAS-CORE.pdf

Angwin, J. (2016b, May 18). *Sample COMPAS Results*. Retrieved from http://www.documentcloud.org/documents/2839240-Sample-Risk-Assessment-COMPAS-Results.html

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine Bias. *ProPublica*. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*(1), 124–140. https://doi.org/10.1016/0749-5978(85)90049-4

Baker v. Canada (Minister of Citizenship and Immigration), 2 Supreme Court Reports 817 (Supreme Court of Canada 1999).

Baldus, D. C., Woodworth, G., & Grosso, C. M. (2007). Race and Proportionality Since McCleskey v. Kemp (1987): Different Actors with Mixed Strategies of Denial and Avoidance. *Columbia Human Rights Law Review*, *39*, 36.

Barocas, S., Bradley, E., Honavar, V., & Provost, F. (2017). Big Data, Data Science, and Civil Rights. *ArXiv:1706.03102 [Cs]*. Retrieved from http://arxiv.org/abs/1706.03102

Barocas, S., & Selbst, A. D. (2016). *Big Data's Disparate Impact* (SSRN Scholarly Paper No. ID 2477899). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2477899

BC Human Rights Tribunal. (n.d.). Leading Cases—Protected Characteristics. Retrieved July 21, 2018, from BC Human Rights Tribunal website: http://www.bchrt.bc.ca/law-library/leading-cases/protected-characteristics.htm

Beatson, J. (2018). AI-Supported Adjudicators: Should Artificial Intelligence Have a Role in Tribunal Adjudication? *Canadian Journal of Administrative Law & Practice; Scarborough*, *31*(3), 307–337.

Benbouzid, B. (2018). *Values and Consequences in Predictive Machine Evaluation. A Sociology of Predictive Policing*. *31*. https://doi.org/10.13140/RG.2.2.25512.67846

Bender, P. (1983). The Canadian Charter of Rights and Freedoms and the United. *McGill Law Journal*, *28*(4), 56.

Blumenthal-Barby, J. S., & Krieger, H. (2015). Cognitive Biases and Heuristics in Medical Decision Making: A Critical Review Using a Systematic Search Strategy. *Medical Decision Making*, *35*(4), 539–557. https://doi.org/10.1177/0272989X14547740

Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. *ArXiv:1607.06520 [Cs, Stat]*. Retrieved from http://arxiv.org/abs/1607.06520

Brantingham, P. J., Valasik, M., & Mohler, G. O. (2018). Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial. *Statistics and Public Policy*, *5*(1), 1–6. https://doi.org/10.1080/2330443X.2018.1438940

Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. Retrieved from http://go.utlib.ca/cat/12222450

Calhoun, C. C. (2002). Adversarial system. In *Dictionary of the Social Sciences*. Retrieved from http://www.oxfordreference.com/view/10.1093/acref/9780195123715.001.0001/acref-9780195123715-e-28

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

CBC News. (2018). *LAPD tries to predict crime through data analysis*. Retrieved from https://youtu.be/F56v5xHt8g0

Choo, C. W. (2005). *The Knowing Organization: How Organizations Use Information to Construct Meaning, Create Knowledge, and Make Decisions*. Oxford University Press.

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic decision making and the cost of fairness. *ArXiv:1701.08230 [Cs, Stat]*. https://doi.org/10.1145/3097983.309809

Cummings, L. (2014). The “Trust” Heuristic: Arguments from Authority in Public Health. *Health Communication*, *29*(10), 1043–1056. https://doi.org/10.1080/10410236.2013.831685

d’Alessandro, B., O’Neil, C., & LaGatta, T. (2017). Conscientious Classification: A Data Scientist’s Guide to Discrimination-Aware Classification. *Big Data*, *5*(2), 120–134. https://doi.org/10.1089/big.2016.0048

de Laat, P. B. (2017). Big Data and Algorithmic Decision-making: Can Transparency Restore Accountability? *SIGCAS Comput. Soc.*, *47*(3), 39–53. https://doi.org/10.1145/3144592.3144597

Degeling, M., & Berendt, B. (2018). What is wrong about Robocops as consultants? A technology-centric critique of predictive policing. *AI & SOCIETY*, *33*(3), 347–356. https://doi.org/10.1007/s00146-017-0730-7

Dieterich, W., Mendoza, C., & Brennan, T. (2016, July 8). *COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity*. Retrieved from http://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf

Domtar Inc. V. Quebec (Commission d’appel en matière de lésions professionnelles), 2 SCR 756 (Supreme Court of Canada 1993).

Dunsmuir v. New Brunswick, 1 Supreme Court Reports 190 (Supreme Court of Canada 2008).

European Commission. (2018a, April 10). EU Member States sign up to cooperate on Artificial Intelligence [Text]. Retrieved April 30, 2019, from Digital Single Market—European Commission website: https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence

European Commission. (2018b, October 24). Smart lie-detection system to tighten EU's busy borders. Retrieved June 1, 2019, from https://ec.europa.eu/research/infocentre/article_en.cfm?artid=49726

Executive Office of the President. (2016). *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*. Retrieved from The White House website: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

Flood, C. M., & Sossin, L. M. (Eds.). (2013). *Administrative law in context* (2. ed). Toronto: Emond Montgomery Publications.

Frankish, K., & Ramsey, W. M. (Eds.). (2014). *The Cambridge handbook of artificial intelligence*. Cambridge, UK: Cambridge University Press.

Freeman, K. (2016). Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis. *North Carolina Journal of Law & Technology*, *18*(5), 75.

Government of Canada. (2019, February 5). *Directive on Automated Decision-Making*. Retrieved from https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592

iBorderCtrl. (2016a). Frequently Asked Questions | iBorderCtrl. Retrieved May 30, 2019, from https://www.iborderctrl.eu/Frequently-Asked-Questions

iBorderCtrl. (2016b). Technical Framework | iBorderCtrl. Retrieved November 9, 2018, from https://www.iborderctrl.eu/Technical-Framework

iBorderCtrl. (2016c). The project | iBorderCtrl. Retrieved November 9, 2018, from https://www.iborderctrl.eu/The-project

iBorderCtrl. (n.d.). IBorderCtrl Traveller Information Guide. Retrieved June 2, 2019, from https://www.iborderctrl.eu/tua/#/main-container/user-instruction-page

IEEE. (2016). *Ethically Aligned Design: A Vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems* (No. 1). Retrieved from http://standards.ieee.org/develop/indconn/ec/ead_v1.pdf

Jones, C. E. (2013). The troubling new science of legal persuasion: Heuristics and biases in judicial decision-making. *Advoc. Q.*, *41*, 49.

Kagan, R. A. (2009). *Adversarial Legalism: The American Way of Law*. Retrieved from http://go.utlib.ca/cat/11345548

Kahneman, D., Ritov, I., & Schkade, D. (1999). Economic preferences or attitude expressions?: An analysis of dollar responses to public issues. *Journal of Risk and Uncertainty; New York*, *19*(1), 203–235.

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Retrieved from http://go.utlib.ca/cat/4235399

Kahneman, D., Tversky, A., & Slovic, P. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Retrieved from http://go.utlib.ca/cat/2611354

Karppi, T. (2018). “The Computer Said So”: On the Ethics, Effectiveness, and Cultural Techniques of Predictive Policing. *Social Media + Society*, *4*(2), 2056305118768296. https://doi.org/10.1177/2056305118768296

Kilbertus, N., Rojas-Carulla, M., Parascandolo, G., Hardt, M., Janzing, D., & Schölkopf, B. (2017). Avoiding Discrimination through Causal Reasoning. *ArXiv:1706.02744 [Cs, Stat]*. Retrieved from http://arxiv.org/abs/1706.02744

Knight v. Indian Head School Division No. 19, 1 Supreme Court Reports 653 (Supreme Court of Canada 1990).

Kroll, J., Huey, J., Barocas, S., Felten, E., Reidenberg, J., Robinson, D., & Yu, H. (2017). Accountable Algorithms. *University of Pennsylvania Law Review*, *165*(3), 633.

Law, J. L. (2018). Enabling statute. In J. Law (Ed.), *A Dictionary of Law*. Retrieved from http://www.oxfordreference.com/view/10.1093/acref/9780198802525.001.0001/acref-9780198802525-e-1321

Law Reform Commission of Canada. (1980). *Administrative Law, Independent Administrative Agencies, Working Paper 25* [Working Paper]. Retrieved from http://www.lareau-legal.ca/LRCWP25.pdf

Lazer, D., & Radford, J. (2017). Data ex Machina: Introduction to Big Data. *Annual Review of Sociology*, *43*, 19–39. https://doi.org/10.1146/annurev-soc-060116-053457

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, 31.

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, *13*(5), 14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Macnish, K. (2012). Unblinking eyes: The ethics of automating surveillance. *Ethics and Information Technology*, *14*(2), 151–167. https://doi.org/10.1007/s10676-012-9291-0

Mann, T. M., & Blunden, A. B. (2010). Privative clause. In *Australian Law Dictionary*. Retrieved from http://www.oxfordreference.com/view/10.1093/acref/9780195557558.001.0001/acref-9780195557558-e-2602

Mergel, I., Rethemeyer, R. K., & Isett, K. (n.d.). Big Data in Public Affairs. *Public Administration Review*, *76*(6), 928–937. https://doi.org/10.1111/puar.12625

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative Research: A Guide to Design and Implementation*. Retrieved from http://ebookcentral.proquest.com/lib/utoronto/detail.action?docID=4040452

Miller, E. J. (2015). *Challenging Police Discretion* (SSRN Scholarly Paper No. ID 2649696). Retrieved from Social Science Research Network website: https://ssrn.com/abstract=2649696

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716679679

Northpointe. (2015, March 19). *Practitioner's Guide to COMPAS Core*. Retrieved from http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core-_031915.pdf

O'Leary, D. E. (2013). Artificial Intelligence and Big Data. *IEEE Intelligent Systems*, *28*(2), 96–99. https://doi.org/10.1109/MIS.2013.39

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *30*(3), 286–297.

PredPol. (2018). Overview. Retrieved May 29, 2019, from PredPol website: https://www.predpol.com/about/

PredPol. (2018). Predictive Policing Technology. Retrieved May 30, 2019, from PredPol website: https://www.predpol.com/technology/

Puente, M. (2019, May 5). LAPD to scrap some crime data programs after criticism. *LA Times*. Retrieved from https://www.latimes.com/local/lanow/la-me-lapd-predictive-policing-big-data-20190405-story.html

Rehaag, S. (2017). "I Simply do not Believe...": A Case Study of Credibility Determinations in Canadian Refugee Adjudication. *Windsor Review of Legal and Social Issues*, *38*, 34.

Romei, A., & Ruggieri, S. (2014). A multidisciplinary survey on discrimination analysis. *The Knowledge Engineering Review*, *29*(5), 582–638. https://doi.org/10.1017/S0269888913000039

Shipley, D. E. (2008). Due Process Rights Before EU Agencies: The Rights of Defense. *GA. J. INT'L & CoMP. L.*, *37*(1), 52.

Smith, M. (2018, October 30). *Can we predict when and where a crime will take place?* Retrieved from https://www.bbc.com/news/business-46017239

Smith, M. (2019). *Review of Selected Los Angeles Police Department Data-Driven Policing Strategies* (p. 52). Retrieved from Office of the Inspector General website: http://www.lapdpolicecom.lacity.org/031219/BPC_19-0072.pdf

State v. Loomis, 881 N.W.2d 749 (Supreme Court of Wisconsin 2016).

Thompson, D. (2019, June 20). Should We Be Afraid of AI in the Criminal-Justice System? *The Atlantic*. Retrieved from https://www.theatlantic.com/ideas/archive/2019/06/should-we-be-afraid-of-ai-in-the-criminal-justice-system/592084/

Thompson, S. (2018, September 24). "Predictive policing" just a new spin on old law enforcement biases, critics say | CBC News. *CBC*. Retrieved from https://www.cbc.ca/news/world/crime-los-angeles-predictive-policing-algorithms-1.4826030

Treasury Board Secretariat of Canada. (2018). *RFI for AI Services, solutions and Products* (White Paper No. 000). Retrieved from Government of Canada website: https://buyandsell.gc.ca/cds/public/2018/06/28/05baa93c08b6f2d3000855170f831066/ABES.PROD.PW__EE.B017.E33657.EBSU000.PDF

Trump, D. (2019). Executive Order 13859: Maintaining American Leadership in Artificial Intelligence. *Federal Register*, *84*(31), 3967–3972.

United States Probation and Pretrial Services, W. D. of W. (n.d.). Presentence Investigation | Western District of Wisconsin. Retrieved May 30, 2019, from https://www.wiwp.uscourts.gov/presentence-investigation

Washington, A. L. (2018). How to Argue with an Algorithm: Lessons from the COMPAS-ProPublica Debate. *Colo. Tech. LJ*, *17*, 131.

WIRED. (2018). *How Cops Are Using Algorithms to Predict Crimes*. Retrieved from https://youtu.be/7lpCWxlRFAw

Wisconsin State Public Defenders. (2014). *Using the COMPAS Electronic Pre-Sentence Investigation Report* (p. 20). Retrieved from https://www.wispd.org/attachments/article/272/COMPAS%20PSI%20Presentation%20by%20DOC.pdf

Wisser, L. (2019). Pandora's Algorithmic Black Box: The Challenges of Using Algorithmic Risk Assessments in Sentencing. *Am. Crim. L. Rev.*, *56*, 1811.

Appendix A: COMPAS Risk Assessment Survey

Risk Assessment

PERSON

Name:	Offender #:	DOB:
[redacted]	[redacted]	[redacted]

[redacted]	Gender:	Marital Status:	Agency:
	Male	Single	DAI

ASSESSMENT INFORMATION

Case Identifier:	Scale Set:	Screener:	Screening Date:
[redacted]	Wisconsin Core - Community Language	[redacted]	[redacted]

Current Charges

☐ Homicide ☑ Weapons ☑ Assault ☐ Arson
☐ Robbery ☐ Burglary ☐ Property/Larceny ☐ Fraud
☐ Drug Trafficking/Sales ☐ Drug Possession/Use ☐ DUI/OUIL ☑ Other
☐ Sex Offense with Force ☐ Sex Offense w/o Force

1. Do any current offenses involve family violence?
   ☑ No ☐ Yes
2. Which offense category represents the most serious current offense?
   ☐ Misdemeanor ☐ Non-violent Felony ☑ Violent Felony
3. Was this person on probation or parole at the time of the current offense?
   ☑ Probation ☐ Parole ☐ Both ☐ Neither
4. Based on the screener's observations, is this person a suspected or admitted gang member?
   ☐ No ☑ Yes
5. Number of pending charges or holds?
   ☑ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4+
6. Is the current top charge felony property or fraud?
   ☑ No ☐ Yes

Criminal History

Exclude the current case for these questions.

7. How many times has this person been arrested before as an adult or juvenile (criminal arrests only)?
   5
8. How many prior juvenile felony offense arrests?
   ☐ 0 ☐ 1 ☐ 2 ☐ 3 ☑ 4 ☐ 5+
9. How many prior juvenile violent felony offense arrests?
   ☐ 0 ☐ 1 ☑ 2+
10. How many prior commitments to a juvenile institution?
    ☐ 0 ☑ 1 ☐ 2+

Note to Screener: The following Criminal History Summary questions require you to add up the total number of specific types of offenses in the person's criminal history. Count an offense type if it was among the charges or counts within an arrest event. Exclude the current case for the following questions.

11. How many times has this person been arrested for a felony property offense that included an element of violence?
☐ 0 ☑ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
12. How many prior murder/voluntary manslaughter offense arrests as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3+
13. How many prior felony assault offense arrests (not murder, sex, or domestic violence) as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3+
14. How many prior misdemeanor assault offense arrests (not sex or domestic violence) as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3+
15. How many prior family violence offense arrests as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3+
16. How many prior sex offense arrests (with force) as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3+
17. How many prior weapons offense arrests as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3+
18. How many prior drug trafficking/sales offense arrests as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3+
19. How many prior drug possession/use offense arrests as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3+
20. How many times has this person been sentenced to jail for 30 days or more?
☐ 0 ☑ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
21. How many times has this person been sentenced (new commitment) to state or federal prison?
☐ 0 ☑ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
22. How many times has this person been sentenced to probation as an adult?
☑ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+

Include the current case for the following question(s).

23. Has this person, while incarcerated in jail or prison, ever received serious or administrative disciplinary infractions for fighting/threatening other inmates or staff?
☑ No ☐ Yes
24. What was the age of this person when he or she was first arrested as an adult or juvenile (criminal arrests only)?
14

Non-Compliance

Include the current case for these questions.

25. How many times has this person violated his or her parole?
☑ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
26. How many times has this person been returned to custody while on parole?
☑ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
27. How many times has this person had a new charge/arrest while on probation?
☐ 0 ☐ 1 ☐ 2 ☐ 3 ☑ 4 ☐ 5+
28. How many times has this person's probation been violated or revoked?
☐ 0 ☐ 1 ☑ 2 ☐ 3 ☐ 4 ☐ 5+

29. How many times has this person failed to appear for a scheduled criminal court hearing?
☑ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+

30. How many times has the person been arrested/charged w/new crime while on pretrial release (includes current)?
☐ 0 ☐ 1 ☐ 2 ☑ 3+

Family Criminality

The next few questions are about the family or caretakers that mainly raised you when growing up.

31. Which of the following best describes who principally raised you?
☐ Both Natural Parents
☐ Natural Mother Only
☐ Natural Father Only
☐ Relative(s)
☐ Adoptive Parent(s)
☐ Foster Parent(s)
☑ Other arrangement

32. If you lived with both parents and they later separated, how old were you at the time?
☑ Less than 5 ☐ 5 to 10 ☐ 11 to 14 ☐ 15 or older ☐ Does Not Apply

33. Was your father (or father figure who principally raised you) ever arrested, that you know of?
☑ No ☐ Yes

34. Was your mother (or mother figure who principally raised you) ever arrested, that you know of?
☑ No ☐ Yes

35. Were your brothers or sisters ever arrested, that you know of?
☐ No ☑ Yes

36. Was your wife/husband/partner ever arrested, that you know of?
☑ No ☐ Yes

37. Did a parent or parent figure who raised you ever have a drug or alcohol problem?
☑ No ☐ Yes

38. Was one of your parents (or parent figure who raised you) ever sent to jail or prison?
☑ No ☐ Yes

Peers

Please think of your friends and the people you hung out with in the past few (3-6) months.

39. How many of your friends/acquaintances have ever been arrested?
☐ None ☐ Few ☑ Half ☐ Most

40. How many of your friends/acquaintances served time in jail or prison?
☐ None ☐ Few ☑ Half ☐ Most

41. How many of your friends/acquaintances are gang members?
☐ None ☑ Few ☐ Half ☐ Most

42. How many of your friends/acquaintances are taking illegal drugs regularly (more than a couple times a month)?
☑ None ☐ Few ☐ Half ☐ Most

43. Have you ever been a gang member?
☐ No ☑ Yes

44. Are you now a gang member?
☐ No ☑ Yes

Substance Abuse

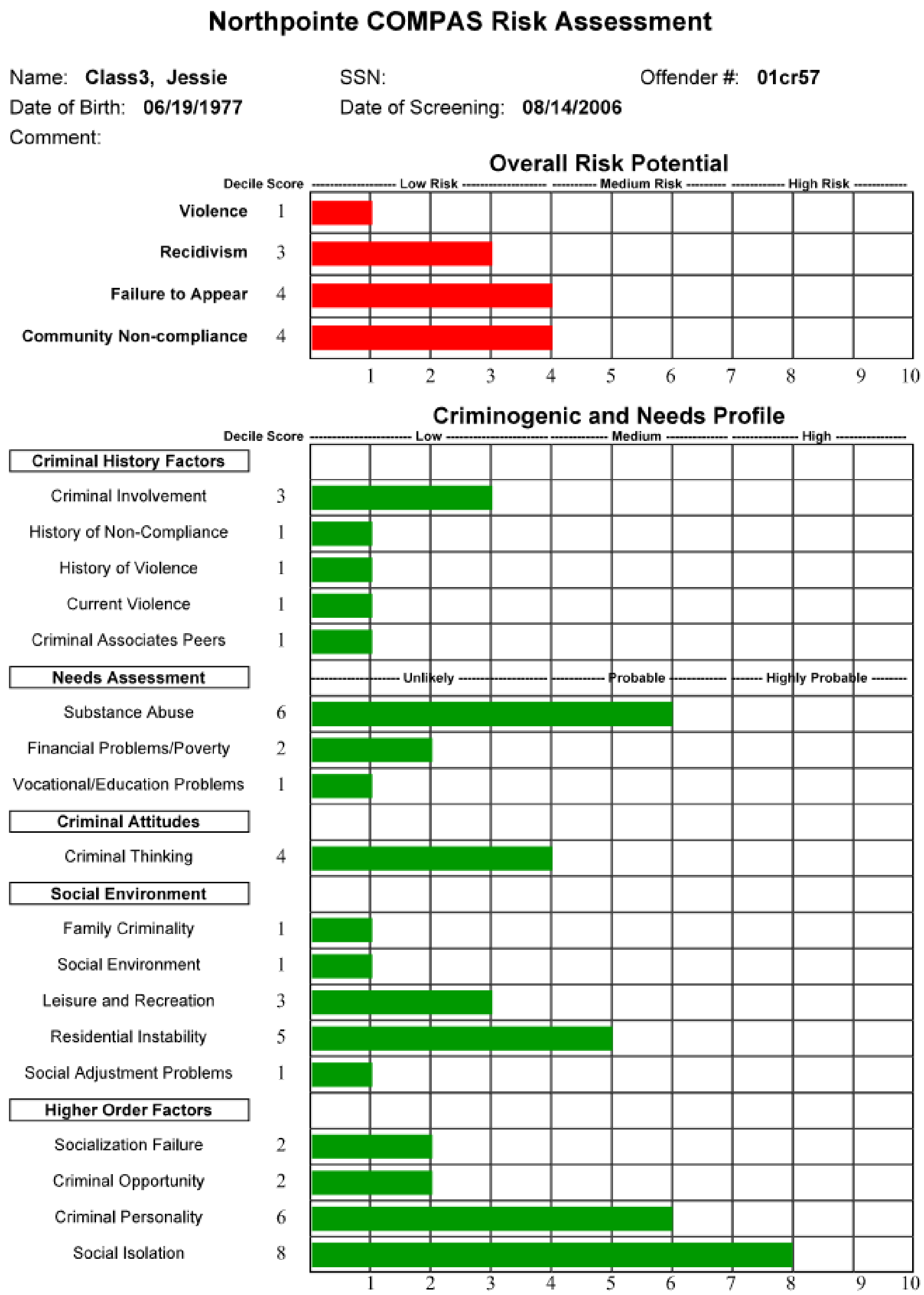
What are your usual habits in using alcohol and drugs?

[redacted]

[redacted] Page 3 / 8

(Angwin, 2016a).

Appendix B: Sample COMPAS Risk Assessment



(Angwin, 2016b).

Appendix C: Sample Presentence Investigation Report

PRESENT OFFENSE

Description of Offense:

According to the Criminal Complaint on 7-17-06 at 1270 W. Glendale Ave in the City of Milwaukee Robert Noname, having been previously convicted of a felony offense was illegally in possession of a firearm.

Milwaukee Police were alerted to a vehicle driven by this defendant reportedly carrying a gun in his lap. This was witnessed by an anonymous citizen in a higher vehicle who had provided vehicle information to the police.

Upon stopping this vehicle, the gun was found in plain view of the floor board at which time this defendant admitted the weapon was his and he had previously been convicted of a felony offense in 2000.

Defendant's Version:

Mr. Noname stated on 7-17-06 he was indeed found with a weapon on his lap in his fiancé's car. He admitted to police the weapon was his. However, it was not his weapon but he did not want his fiancé to be punished. He stated prior to that date he had a BBQ at his house, however no drugs or alcohol were used. He invited a "friend" who inadvertently brought a weapon with him. He warned the individual that he was not to be carrying a weapon at his house. The friend, Brandon, asked him if Mr. Noname could hold the weapon for him until he was leaving to which he agreed. Mr. Noname then explained that Brandon was younger and he had known him from about the age of 13 when he started cutting his hair. He admittedly had known Brandon to be in trouble from time to time and in need of guidance. He further stated this reminded him much of himself at that age and he wanted to assist him in taking a better path for himself.

When the party was over, Brandon's ride was no longer there. Mr. Noname agreed to drive him home, but refused to ride with the gun in his car telling him he would need to arrange to have someone pick it up. Mr. Noname stated he was uncomfortable having a weapon in his house as he has children who visit him on the weekends and he did not want to keep it there. Mr. Noname admitted he attempted to call Brandon and even stopped by his house inquiring when the weapon was to be taken. He was alerted Brandon had been arrested for selling drugs. It was then Mr. Noname decided Brandon did not need this gun and furthermore he was not keeping it any longer at his house. Mr. Noname stated he and his fiancé set out to sell the weapon. He put the gun in his lap and got into his car with his fiancé as the passenger. Someone saw him with the gun in his lap as he was driving and reported this to police. He stated when stopped by police, he indicated the weapon was his because the car belonged to his girlfriend and he did not want her to receive any punishment.

In regard to punishment, Mr. Noname stated he knows his behavior was wrong, but he isn't certain what would be appropriate for this law violation. He wanted to leave the punishment to the judge and God. He did request the judge be informed he takes full and complete responsibility for his actions and will face whatever is bestowed upon him "like a man." He admitted he wasn't thinking about potential consequences and believed he was doing well, causing him to forget that having a felony on his record made the possession of a weapon illegal. He confessed he would not make this mistake again. He stated he has much to lose by incarceration. He believes he has the power to positively influence people and people look up to him. He is ashamed of his behavior, but has learned from his mistake.

2 NONAME, ROBERT

PRIOR RECORD

Adult Record:

Date of Offense	Date of Conviction	Date of Sentencing	Location	Offense
06/29/1999	08/13/1999	09/05/1999	Milwaukee	Disorderly Conduct, 99CM000551
Disposition: 2 years probation. Revoked				
07/28/2001	08/30/2001	09/15/2001	Milwaukee	Fleeing, Hit and Run, 01CF00256
Disposition: 3 years Wisconsin Prison System. 18 months initial confinement, 18 months extended supervision				

Juvenile Record:

Date of Offense	Date of Adjudication	Date of Disposition	Location	Offense
unknown	unknown	unknown		
No juvenile record				

Other Relevant Offenses:

Not Applicable

Correctional Experience:

In November of 1998, the defendant had his first contact with law enforcement. He was arrested for loitering and received a ticket which he never paid.

On 1/27/99, the defendant was issued a municipal citation for disorderly conduct.

On 7/1/99 he was again arrested for disorderly conduct. He was found guilty and on 9/07/99 he was sentenced 60 days jail. The sentence was stayed and the defendant was placed on probation for two years. He followed his rules but failed to enroll back in school.

On 7/21/99, he was issued a citation for SAR and on 8/20/99 he received a citation for a pedestrian violation.

On 6/30/00, the defendant was arrested for Take and Drive a Vehicle Without Owner's consent when he was stopped by police while driving a stolen car. The charges were dropped because of confusion over whether the defendant knew the car was stolen. However, he was cited for excessive sound and SAR.

On 07/30/01, he was arrested for Fleeing and was also charged with two counts of Hit and Run, which were dismissed and read in. On 09/17/01 he was sentenced to prison. When released on ES supervision, he had several violations including a dirty UA (THC) and continuing to drive without a valid license.

On 3-25-04, he was questioned by police and gave the false name of Donte Rice and at the same time had in his possession a picture ID of himself as Donte Rice. This occurred at 42nd Street and Holton. Revocation proceeding were started, however, the defendant was offered an ATR by being placed on Electronic Monitoring and allowed to work as a barber.

He discharged from ES supervision on 6-23-04, successfully without incident on EMP.

Pending Charges:

Presently, this defendant does not have any additional cases pending in the Milwaukee County Circuit Court.

Defendant's Explanation of Record:

The defendant's prior offense involved him engaging in a feud with other teenagers and confronting a group of juveniles as they exited a building. He removed a chrome semi-automatic handgun from his waistband and held it in the air to intimidate the other subjects. The defendant fled the scene before the police arrived.

When asked to explain his prior record, Mr. Noname relayed his involvement in the Criminal Justice System was the result of growing up and making poor choices/mistakes. He stated he failed to consider consequences of his actions and was hanging around with the wrong people and was doing things out of character. Mr. Noname considered himself a trusting person who believes certain people are his friends. This consistently ends in trouble. Mr. Noname indicated he was successful in his incarceration term as well as on extended supervision.

FAMILY

IDENTIFYING INFORMATION

Mother

Name	Street Address	City	State	Zip
Geena Noname	4391 Sumac Pl	Milwaukee	WI	55555

Geena Noname received her Bachelor's degree in Education in 1975. She was employed in the Milwaukee School District for 25 years and is currently retired. According to CCAP, she has no criminal record.

Father

Name	Street Address	City	State	Zip
Robert Noname	2601 Evansville Rd	Milwaukee	WI	55555

Robert Noname is a high school graduate and owns his own construction business titled, Noname and Associates Construction. He also has owned and managed various properties. According to CCAP, he was previously convicted of Disorderly Conduct and received probation.

Spouse/Significant Other

Name	Street Address	City	State	Zip
Latoya Harris	1405 Langdon St Apt. A	Milwaukee	WI	55555

The Defendant's fiancé, Latoya Harris currently resides with the defendant in Milwaukee. Latoya completed high school and cosmetology school and is currently employed as a hairstylist. According to CCAP, she was arrested for retail theft and served 30 days in jail.

Dependents

Name	Street Address	City	State	Zip
Aquarius Campbell	unknown	Milwaukee	WI	55555
Nyiara Noname	unknown	Milwaukee	WI	55555
Savion Noname	8934 Peachtree St	Atlanta	GA	34126

Mr. Noname has three children with three different women. Aquarius Campbell (10 yrs) is the oldest child of the defendant and Tequilla Campbell. He has another daughter, Nyiara Noname (8 yrs) who was born to he and Marla Strauder. Savion Noname (6 yrs) is his youngest son born to he and Shaneka Sloan.

Siblings

Name	Street Address	City	State	Zip
Jared Noname	3099 Locust St	Milwaukee	WI	55555

Mr. Noname has one brother, Jared Noname. He stated that he has maintained a good relationship with him over the years. Jared (30 yrs) is a college graduate and is currently employed as a nurse at Children's Hospital in Milwaukee. He has no criminal record.

Other

Name	Street Address	City	State	Zip

Not Applicable

FAMILY STABILITY, ATTITUDES, & VALUES

Family Criminality

Family Criminality Scale Score: Probable

Mr. ROBERT NONAME's family members (parents and/or siblings) were reported as having some involvement in criminal activity, drugs, and/or alcohol abuse. Mr. ROBERT NONAME may need to minimize or structure the contact with certain members of his/her family to minimize adverse influence or exposure to inappropriate substance use and to avoid modeling of violent or criminal behaviors.

Socialization Failure

Socialization Failure Scale Score: Unlikely

The Socialization Failure Scale score suggests that Mr. ROBERT NONAME's is unlikely to have an impaired socialization.

Agent Comments:

Mr. Noname was born in Milwaukee to parents, Geena Noname and Robert Noname. He stated his parents were never married and he was the only child of their relationship. Until the age of 7, Mr. Noname was raised by his mother and he reported not having a close relationship with her. At that time, his father requested full custody of his son and, due to financial reasons, was granted his request. The defendant stated he has not seen his mother in years. Geena Noname does not have a criminal record according to CCAP, whereas Robert does have a Disorderly Conduct. Both of the defendant's parents have had stable employment through his childhood into adulthood.

Mr. Noname explained his early childhood with his mother was difficult due to her distant nature, and he was financially deprived of many things. After moving in with his father, both money and relational issues were better. His father was very educationally involved and demanded that Mr. Noname receive good grades. He was frequently ordered to study for many hours every day and had his social activities limited as a result. He felt his father was very strict in this regard. When he was allowed freedoms as he got older, he went "wild." He acknowledged making poor choices. He called his dad a "military father." He was a strict disciplinarian, utilizing grounding and privilege withdrawal as his punishments. He did not report any physical abuse. When describing his relationship with his father, Mr. Noname stated he and his dad are very close. He feels he can talk to him about anything. He stated his father is very disappointed in this present offense, but has told his son that now he is a man and must accept responsibility like a man.

This writer did have the opportunity to speak with the defendant's father, Robert Noname. He stated he raised his son from the age of 5 until adulthood. He felt his son was an intelligent man. Mr. Noname had no complaints concerning his son when he was a child. He stated his son hung around with the wrong crowd. To date, he feels his son is with a woman with whom he can make a future. He has four children he needs to support and has the ability to do that.

During the interview with Robert Noname senior he stated his son has recently been having some difficulties with his personal safety. Mr. Noname contended his son has been "jumped, jacked and robbed" several times by the same individuals. He further stated most likely his son purchased the weapon for protection. Mr. Noname stated he is always there for his son, but now that he is an adult he is responsible for his own decision making skills. He has taught him to consider consequences and weigh outcomes and hopes now he has internalized the importance of this. Mr. Noname further reported his son is aware that he will need to somehow change his social life to live more conflict free. Mr. Noname hoped his son would be given a second chance and stated sometimes that is necessary. However, he also felt he would rather see his son in jail than visit his grave. He was well aware that fights are no longer customary and people do not hesitate to kill in settling disputes. Furthermore, he did not wish to make his son a statistic of the madness. In closing Mr. Noname stated life choices make or break a person and for every action there is a consequence that needs consideration.

With regard to other family involvement, Mr. Noname has three children with three different women. His oldest daughter is Aquarius Campbell (10 yrs). Mr. Noname stated he and her mother, Tequilla Campbell, grew up together and had been best friends for a very long time before developing a relationship resulting in the birth of their daughter. He stated the romantic interest lasted about a year, but they continue to remain good friends at present. Mr. Noname's second daughter is Nyiara Noname (8 yrs). Her mother, Marla Strauder was introduced to him by a friend at the barber shop. He stated they dated for two years, but mutually decided to separate after deciding they had grown apart. His youngest child is Savion Noname (6 yrs). Mr. Noname indicated the mother, Shaneka Sloan, had moved with their son to Atlanta, Georgia causing their separation. She has since moved back allowing more consistent contact with his son. Mr. Noname contends he maintains routine contact with all of his children, but stated he is behind in court ordered child support by about $13,000. He stated he is making serious efforts at payments as able.

The defendant and Latoya Harris were engaged 3 years ago and reportedly have a "great" relationship. During an interview with Ms. Harris, she stated they communicate well, have similar interests and are always there for each other. She also discussed that the defendant continues to maintain an active role with his children and she has also developed a relationship with them. With regard to the offense, she believes he realizes his actions were wrong and is working on making more positive choices in many areas of his life. She added she will continue to support him.

PERSONAL HISTORY

VOCATION/EDUCATION/EMPLOYMENT

Education History

Name	City	State	Zip Code	Phone Number	Length
Rufus King High School	Milwaukee	WI	53209	(414) 267-0700	N/A
Kilmer South	Milwaukee	WI	53216		N/A
Milwaukee Area Technical College	Milwaukee	WI	53233		N/A

Employment History

Employer	Occupation	Supervisor	Phone	Length
Julia Myrick	Barber	Julia Myrick		09/2013 -
Nex-Level	Barber	Unknown		09/2000 - 10/2004

Vocation/Education

Vocational/Education Scale Score: Unlikely

The Vocational and Education Scale score indicates Mr. ROBERT NONAME does not need vocational or educational treatment intervention.

Agent Comments:

The SLOSSON ORAL READING TEST was administered to Mr. Noname to which he scored the reading ability of a 12.5 grade student.

Mr. Noname indicated he did receive his GED while incarcerated at Prairie Du Chien in 2002. This information was verified through an official copy of his transcript and GED results located in his terminated DOC file.

Mr. Noname stated his first two years of High School were spent at Rufus King. He recalled he was behind in credits and therefore sent to Kilmer South to catch up. While in school he stated his grades were rather low due to poor peer influence. He felt high school was an entirely different environment for him and he was more involved in partying than his education at that time. His attendance and overall behavior were good. He did acknowledge he was the class clown, not extremely disruptive, but enjoyed making jokes. While in school he was involved in Basketball, Forensics and the Debate Team.

Mr. Noname did provide this writer with paper copies of his certifications for his GED as well as high school report cards indicating his attendance as stated. It should also be noted he did attend course work at WCTC for Milwaukee Career Quest in 1995.

Mr. Noname is presently enrolled at MATC for Barbery/Cosmetology. He indicated he has an additional 1.5 years to complete. He is presently working on getting his cutting hours completed and does cut hair at a Barber Shop presently. Following his education, he plans to continue in school for his instructor's license.

Mr. Noname is presently employed by Julia Myrick and works full time cutting hair. He brought in pictures to demonstrate his work and stated that he is very involved in creative cutting including shaved artwork. He stated many of his clients are neighboring high school students who specifically come to him for his superior hair cutting ability. He stated many of these individuals look up to him and are very disappointed to hear of his recent involvement in this offense. He stated he has utilized this as a tool to tell them to stay out of trouble and protect their freedom.

Mr. Noname indicated he has also worked at Nex-Level doing hair cutting as well. This position lasted from 2000-2004. He stated while he enjoyed his work there he did not feel the atmosphere was professional and wanted the recognition of a more professional salon. This information is congruent with the assessment results.

MENTAL ABILITY

Agent Comments:

Mr. Noname stated that he has never been diagnosed with a developmental disability. This writer also observed that the defendant was able to actively engage in the interview process and reported understanding all questions asked.

FINANCIAL

Financial

Financial Scale Score: Highly Probable

Mr. ROBERT NONAME's Financial Scale score indicates that he is likely to have financial problems. He is likely to worry about financial stability, has trouble paying bills, and may have conflicts with family or friends over money. If this is the case, he would likely benefit from a class on financial management, job skills or vocational/employment training. He may also require assistance understanding and negotiating social assistance such as welfare, food stamps, and unemployment compensation. A case plan may also prioritize stabilizing his income and developing budgeting skills.

Agent Comments:

Presently, Mr. Noname is earning approximately $400 per week. He stated he has interests in designing his own hair book as well as sponsoring a hair show in the future. While he reports little to no problems paying current bills, he does comment on his concern of long term financial stability. This in part stems from significant child support debts, which are currently approximately $13,000. As is reflected in his assessment he does worry about paying back his child support and reports he is working on creating a plan to do just that. In addition to child support he also acknowledges a need to address his outstanding court costs and fines from his latest court case, which total approximately $1800.

COMPANIONS

Criminal Associates/Peers

Criminal Associates/Peers Scale Score: Probable

The Criminal Associates and Peers Scale score indicates Mr. ROBERT NONAME likely has a modest involvement with antisocial friends. Restricting his contact with any current antisocial associates may help to reduce criminal opportunity. Encourage Mr. ROBERT NONAME to build more affiliations with pro-social peers in various pro-social activities.

Social Isolation

Social Isolation Scale Score: Probable

Mr. ROBERT NONAME's Social Isolation Scale score suggests that he may lack a supportive social network that he relies on regularly and during difficult times. He may also exhibit feelings of isolation and loneliness. If this is the case, he may benefit from participating in activities with pro-social people such as becoming involved in a church group, recreational activities, volunteering, and taking fitness classes.

Agent Comments:

As is indicated in the Assessment results, Mr. Noname has both positive and negative influences in his life. Mr. Noname reports having some friends who he has known since high school who are criminally involved and/or are gang members.

Some of these friends do have criminal records. He did report he is trying to minimize the time he spends with his friends that are still gang involved, as he feels they are not a good influence on him. He also states some of his friends are positive and enjoy playing basketball, cards and watching sporting events. They support him in prosocial activities and these are the friends he is trying to spend more time around. Mr. Noname states he does not associate with any known drug users or dealers. Likewise, he reports never being involved in a gang.

In discussing positive supports in his life, Mr. Noname references his good friend and coworker Ray who has acted as a mentor and support in the past. He indicates a desire to remain close with Ray because he respects how successful Ray has been in the haircutting business and would like to learn from him.

EMOTIONAL AND PHYSICAL HEALTH

Agent Comments:

Mr. Noname stated he has never been diagnosed with a learning disability, mental health or substance abuse problem. He is not presently taking any medications and has never been prescribed medications to aid with anxiety, depression, mood swings, thinking problems or controlling his behavior. Mr. Noname denied any suicidal thoughts or attempts at this endeavor. Mr. Noname reported having had a hernia surgery in 2001 while in prison. Aside from that he is in fine physical health. He also does not report any sexual victimization or traumatic events in his past.

SEXUAL BEHAVIOR

Agent Comments:

Not Applicable

ATTITUDES AND BELIEFS

Criminal Personality

Criminal Personality Scale Score: Highly Probable

Mr. ROBERT NONAME's scale score indicates a tendency toward an antisocial personality. This may include factors such as: impulsivity, risk-taking, boredom, no guilt, selfishness, anger, and so on. Referral for a more in-depth personality assessment may be warranted. This score, in some cases, may also indicate a resistance to treatment. Impulsive decision-making, if detected, may be amenable to some form of cognitive therapy. Mr. ROBERT NONAME may need high levels of control.

Criminal Thinking Self Report

Criminal Thinking Self Report Scale Score: Highly Probable

Mr. ROBERT NONAME's Self-Report Criminal Thinking Scale score indicates that he is likely to rationalize his criminal behavior. He is unlikely to accept responsibility for his actions and may minimize the seriousness and consequences of his criminal behavior. If this is the case, then a cognitive restructuring intervention is advisable. This program should focus on modifying his criminal attitudes and thinking patterns and implementing pro-social reframes. Responsivity to other forms of treatment may be low while the person continues to excuse and rationalize his behavior; therefore cognitive restructuring interventions might be sequenced prior to other types of programs, treatments or placements.

Anger

Anger Scale Score: Probable

Mr. ROBERT NONAME's Anger Management Scale score suggests he is likely to have difficulty managing and controlling his anger. He may be quick to anger and often loses his temper. If this is the case, then he may benefit from anger management classes or assertiveness training. These classes will help him understand what triggers his anger and alternative ways of expressing his feelings.

Cognitive Behavior

Cognitive Behavioral Scale Score: Probable

Mr. ROBERT NONAME's Cognitive Behavioral/Psychological score suggests the presence of some anti-social attitudes. In some cases these may include moral justification for his criminal behavior, refusing to accept responsibility, blaming the victim, rationalizations (excuses) that minimize the seriousness and consequences of his criminal activity, etc. He may have some elements of a high risk lifestyle such as: idleness, boredom and impulsive decision-making. If such issues are detected a cognitive therapy program coupled with more positive role models, more socially productive activities, and the development of positive social bonds may be warranted. In some cases a more in-depth mental health assessment may be in order.

Social Adjustment Problems

Social Adjustment Problems Scale Score: Highly Probable

A highly probable score on this scale suggests that Mr. ROBERT NONAME is likely to have problematic relationships in multiple social contexts such as family, school and work. He would likely benefit from classes that can improve his social skills that may help build his social supports, particularly prosocial supports. A cognitive program aimed at improving his success in social contexts may be appropriate.

Agent Comments:

Mr. Noname told this writer that he is very sorry for what happened and he knows now that his behavior was wrong. He states at the time that he forgot he could not carry a weapon. Due to these statements he seems able to take partial, but not full responsibility. He does appear to be rationalizing and minimizing his behavior to a degree by commenting "...but it really wasn't that big of a deal, no one got hurt." Mr. Nomane's assessment indicates low impulse control. This is exhibited in his behavior of possessing the gun despite knowing he should not have one in his possession.

Mr. Noname stated he is trying to make positive changes in his life and believes he can be successful in staying out of trouble in the future. He believes everyone has the ability to succeed if they just put their mind to it. He maintains barbering has given him a second chance on life. His passion and motivation in this line of work will pull him through these hard times.

In the past Mr. Noname reported he has not been as focused on the success of his children and fiancé, but realizes he needs to put them first more than he does. He comments "life is more than just looking out for number one."

Mr. Noname believes he should be held accountable for what he did. Despite saying he forgot at the time, he now acknowledges he should not have possessed the firearm and whatever the judge decides will be what he deserves.

SUBSTANCE USAGE HISTORY

Substance	Amount	Frequency	Last Used
THC (Marijuana, Hashish, etc)	1 joint	infrequent	07/2013
Alcohol	1 shot	Infrequent	11/2013

Substance Abuse

Substance Abuse Scale Score: Unlikely

Mr. ROBERT NONAME's substance abuse scale score suggests that he is unlikely to have a serious substance abuse problem. However, in some cases a substance abuse education program may still be appropriate.

Agent Comments:

Mr. Noname stated he does not consider himself to be an alcohol or drug abuser. He tried alcohol for the first time at the legal age of 21 and generally utilizes it in celebration having a "shot" after a hair show. He stated he tried Marijuana at the age of 17 and last used in early July of this year. He too contends this use is purely for special occasions and not a daily habit. Mr. Noname denied the use of any additional illegal substances. He has never used needles to inject drugs. Again, he reiterated he does not consider himself to have a drug habit. He stated his initial introduction to alcohol was at the "club scene." He claimed that use has no effect on him.

In regard to treatment, Mr. Noname denies any need in this area. The UNCOPE AODA screening tool was administered to Mr. Noname which confirmed his claim that treatment in this area is not necessary.

MILITARY

Agent Comments:

Not Applicable

LEISURE TIME ACTIVITIES

Leisure and Recreation

Leisure and Recreation Scale Score: Probable

Mr. ROBERT NONAME's Leisure and Recreation Scale score suggest he may not engage in pro-social activities. He may exhibit feelings of boredom, restlessness, or an inability to maintain interest in a single activity for any length of time. If this is the case, then increasing prosocial activities that are of interest to him may be important. Brainstorming and identifying the types of activities available in his community may be helpful.

Criminal Opportunity

Criminal Opportunity Scale Score: Probable

Mr. ROBERT NONAME's criminal opportunity scale score suggests a tendency towards a fairly high-risk lifestyle. This may involve limited prosocial ties, and a rising level of high-risk high-crime opportunity situations, perhaps in the company of high-risk persons. There may also be an inadequate level of pro-social or constructive activities (e.g., working, spending time with family) and an absence of social ties, high boredom and high restlessness. It may be important to promote increased involvement in more positive and socially constructive persons and activities. Case management strategies may focus on structuring daily activities and minimizing idle time.

Agent Comments:

Mr. Noname states he enjoys spending time with his children. He takes them to parks and sporting events when he is not working. In addition to his passion for his barbering, he also states he likes to cook.

The defendant also told this writer he plays a weekly basketball game with his friends. He also reports he does not usually have any downtime until after his kids go to bed.

RESIDENTIAL HISTORY

Address	City	State	Zip	Date Resided
1405 Langdon Street	Milwaukee	Wisconsin	55555	08/01/2013

Residential Instability

Residential Instability Scale Score: Unlikely

Mr. ROBERT NONAME's relatively low scale score suggests he likely has a fairly stable lifestyle, residence and adequate social ties to the community.

Social Environment

Social Environment Scale Score: Highly Probable

Mr. ROBERT NONAME's Social Environment Scale score suggests that he is likely to live in an environment characterized by high rates of crime, drugs and gangs. If this is the case, then he may benefit from relocating to a lower risk neighborhood, if possible. Explore potential alternative living arrangements with prosocial family and friends. If relocating is not possible, seeking safety or implementing strategies (e.g. staying inside after dark) to keep him safe and out of trouble in his neighborhood may be important. Treatment may also focus on improving residential arrangements and employability.

Agent Comments:

The defendant lived with his mother until he was 7. They lived in the same house for that whole time. At the age of 7 the defendant's father got full custody and he lived with his father in the house that his father still resides in.

Mr. Noname has had stable residence for the past year in the apartment he rents on Langdon St. He did report that the neighborhood has some less than desirable areas, but he is trying to stay away from those and feels they are lessening. Prior to this he resided in the same neighborhood for the past 7 years with the mothers of his children.

DEFENDANT STRENGTHS

Motivated in his current line of employment
Current stable living situation
High school graduate or GED
Age 30 or Greater
Currently Employed
Current Skill or Trade


Overall Risk Potential
Risk
Violent Recidivism Risk Medium
General Recidivism Risk Low
Pretrial Release Risk Low
Decile 1 2 3 4 5 6 7 8 9 10
Criminogenic Need Scales
Criminal Involvement
Criminal Involvement Low
History of Non-Compliance Medium
History of Violence Medium
Current Violence High
Relationships/Lifestyle
Criminal Associates/Peers Probable
Criminal Opportunity Probable
Leisure and Recreation Probable
Social Isolation Probable
Substance Abuse Unlikely
Personality/Attitudes
Criminal Personality Highly Probable
Criminal Thinking Self Report Highly Probable
Anger Probable
Cognitive Behavioral Probable
Family
Family Criminality Probable
Socialization Failure Unlikely
Social Exclusion
Financial Highly Probable
Vocational/Education Unlikely
Residential Instability Unlikely
Social Adjustment Problems Highly Probable
Social Environment Highly Probable

SUMMARY AND CONCLUSIONS

The COMPAS is an actuarial assessment tool which has been validated on a national norming population. This means that it predicts the general likelihood that those with a similar history of offending are either less likely or more likely to commit another crime generally within the two year period following release from custody. The COMPAS assessment does not, however, attempt to predict specifically the likelihood that an individual offender will commit a certain type of specific offense within the same two year period. For that prediction, an alternative screening which is normed specifically for that offender population (i.e. use of screeners such as the STATIC-99R, VASOR, etc. for sex offenders) may be used.

In addition to identifying general levels of risk to re-offend, COMPAS also identifies criminogenic needs specific to that offender which are most likely to impact future criminal behavior. For purposes of Evidence Based Sentencing, actuarial assessment tools are especially relevant to: 1. Identify offenders who should be targeted for interventions. 2. Identify dynamic risk factors to target with conditions of supervision. It is important to remember that while risk scores may assist in informing sentencing decisions based on the risk principle to target medium and high risk individuals for intervention, they should never be the sole and deciding factor in determining the severity of the sentence or whether an offender is incarcerated.

AGENT IMPRESSIONS

Mr. Noname is scheduled to appear before the Court on a Class G Felony offense. This offense represents extremely poor judgment on the part of this offender. He was recently supervised on extended supervision, after spending time in prison for a felony. This writer has a difficult time believing that he "forgot" being in possession of a weapon constituted additional criminal behavior. The COMPAS assessment indicates high criminogenic need areas pertaining to criminal thinking patterns. However, over the course of the last month, Mr. Noname seems to be coming into a clearer realization of the seriousness of his behavior. Initially in both the assessment, as well as during the interview, the defendant seemed to rationalize his behavior and minimize the impact of offense because there was no immediate victim. However, in subsequent interactions, he did voice an increased level of responsibility and willingness to accept the consequences of this behavior.

Mr. Noname has a track record of being involved with the criminal justice system. However, while interviewing him, his current behavior did appear out of character for him, in relation to his current life goals. He appeared very motivated to achieve his dream of becoming a renowned Barber.

Mr. Noname is fortunate the public was not directly harmed during this offense. It is believed this can be used as a learning experience for this defendant to always keep in mind his prior convictions and the implications those convictions will have on him for the rest of his life.

Mr. Noname appears to have a very stable family with appropriate support. He also reports being very motivated in his career. This is an important area in which to foster additional growth and skills, as his financial status does appear to be a criminogenic area for him. Continued career development will allow Mr. Noname more opportunities to address his outstanding financial obligations. He is also older now and reports his maturity is a benefit in making better decisions regarding his career and life goals.

Considering Mr. Noname's motivation for change and current strengths, this writer believes he should receive an imposed and stayed jail term with probation supervision. This recommendation is based on the fact that he now appears to have stability in many key areas and he previously successfully completed his last term of supervision, albeit on Electronic Monitoring.

RESTITUTION

Due to the nature of this offense no restitution is to be determined for this defendant.

AGENT RECOMMENDATION

Mr. Noname stands before the court on a class G Felony for FIPOF. This offense carries a maximum imprisonment term of 10 years with an initial confinement of 5 years. Based on this defendant's recent positive advances in the community and yet taking into account the serious nature of this offense the following recommendation is being made.
It is respectfully recommended that Mr. Noname receive a one year sentence at the HOC straight time to be imposed and stayed placing him on a period of 2 years probation supervision with the conditions listed below.

RECOMMENDED CONDITIONS OF PROBATION OR EXTENDED SUPERVISION

(Consideration given for local programing such as specialty courts, restorative justice, or other diversion.)

Full time employment
No weapons
Payment of all court obligations and supervision fees
Complete all other programming as deemed appropriate by the COMPAS assessment and the Department of Corrections.

ANTICIPATED SUPERVISION PLAN

If the Court orders probation, the Department would supervise Mr. Noname in accordance with the risk assessment and Departmental policies and procedures. Given that he does not pose a significant risk to reoffend per the assessment, his reporting frequency and exposure to the probation office would be limited.

Together with the agent, a realistic supervision case plan would be developed to address the top criminogenic need areas highlighted above. This would likely include intensive homework, role play, and skill building sessions aimed at challenging his predisposition toward criminal thinking. These interventions would likely be individualized as opposed to a group process, considering his lower risk score.

Mr. Noname's general progress would be monitored through to successful completion and reviews of his adjustment under supervision would occur every 6 months. The agent would also monitor the payment of Court obligations and Department supervision fees. Violations of supervision would be swiftly addressed with appropriate sanctions in a manner consistent with established policies and procedures.

Respectfully submitted,

Abbey Fuszard
Probation and Parole Agent #33333

Patty Beyer-Robinson
Corrections Field Supervisor
Unit 333

______________________________ ______________________________

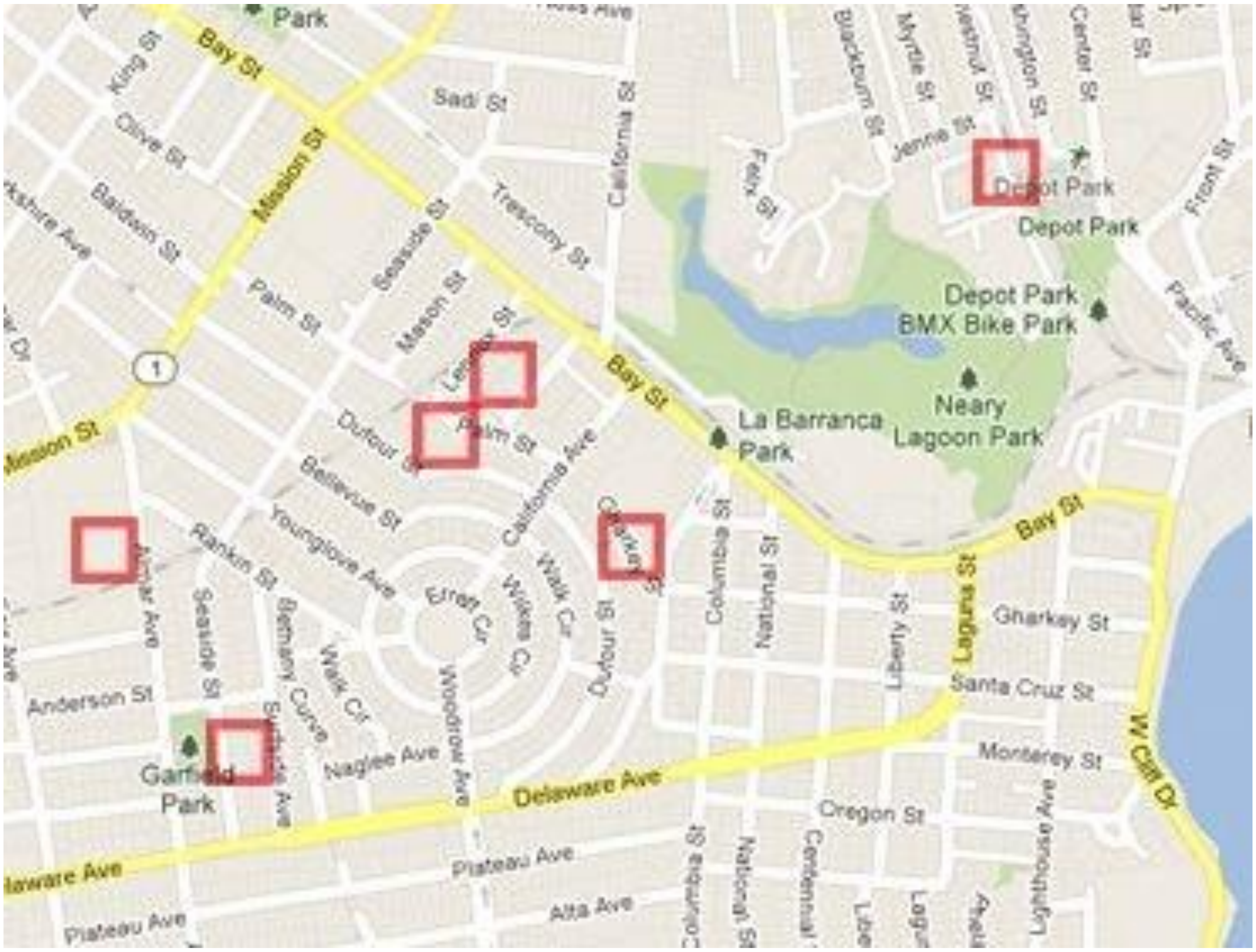
Sources of Information

1. Interview with the offender
2. Records of the Milwaukee Police Department.
3. Records of the Milwaukee District Attorney's office
4. Records of NCIC/CIB and FBI (Portal 100)
5. Records of the Wisconsin Circuit Court Automation Program
6. Milwaukee County Children's Court Records
7. Certificate verification
8. Prairie Du Chien GED scores
9. Phone interview with defendant's father, Robert Noname

20 NONAME, ROBERT

(Wisconsin State Public Defenders, 2014).

Appendix D: Example PredPol Hotspot Map



(Benbouzid, 2018, p. 3).