

NEW YORK UNIVERSITY

BIG DATA: CRITICAL PERSPECTIVES

PROJECT PROPOSAL

---

# Predictive Policing - Opaqueness, Impact and Legal Standards

---

*Author:*

Kevin GALLAGHER

*Professor:*

Helen NISSENBAUM

May 11, 2016

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Predictive Policing: Definition and Taxonomy . . . . .	2
1.2	Due Process . . . . .	3
1.3	Cognitive Psychology . . . . .	4
1.4	Machine Learning . . . . .	5
<b>2</b>	<b>The Novelty of Predictive Policing</b>	<b>5</b>
2.1	Similarities . . . . .	5
2.2	Differences . . . . .	5
<b>3</b>	<b>Verdict and Conclusion</b>	<b>5</b>

## **Abstract**

The abstract will go here.

# 1 Introduction

One Friday morning in July of 2011, two women were arrested in Santa Cruz, California after being caught peering into cars parked in a local parking garage. In addition to the suspicious nature of their activities, one of the women arrested had outstanding warrants, and the other was in possession of drugs. Though arrests of this nature are rarely reported on the news, this particular arrest was an exception due to its peculiar circumstances. These arrests were made after officers were dispatched because a computer algorithm predicted the area as likely to be effected by crime. [14] This technology, called Predictive Policing, is defined as the application of analytical techniques, particularly quantitative techniques, to identify promising targets for police intervention and prevent or solve crime. [20] As such, predictive policing can be seen as a branch of “Big Data” or “Data Science,” the collection and analysis of large quantities of data to form predictive tools.

The purpose of this paper is to analyze the effect of predictive policing on due process. To achieve this, the effects of predictive policing will be viewed through a multidisciplinary lens. This paper discusses the definition and role of due process under the law and its role in liberal democracies, cognitive psychology and models of human thought, and machine learning is discussed to understand the inner working of big data classification and regression. In order to tie all of these different disciplines together, this paper discusses all of them through the lens of probability and likelihood.

The remainder of this paper will be organized as follows, Section 1.1 will define predictive policing. Section 1.2 will define due process and discuss its role in liberal democracies, as well as offer a novel definition of the reasonable suspicion and probable cause standards through the lens of probability. Section 1.3 will introduce elements of cognitive psychology. Section 1.4 will discuss machine learning and probability and their roles in predictive policing. Section 2 will discuss the changes to due process introduced by predictive policing. Finally, Section 3 concludes.

## 1.1 Predictive Policing: Definition and Taxonomy

As mentioned in Section 1, predictive policing is by Perry et al. to be the application of analytical techniques, particularly quantitative techniques, to identify promising targets for police intervention and prevent or solve crime. In the book *Predictive Policing: The role of crime forecasting in law enforcement operations*, Perry et al. present a taxonomy of predictive policing methods. This taxonomy breaks predictive policing into four categories: [20]

1. Methods for predicting future crimes
2. Methods for predicting future offenders
3. Methods for predicting perpetrator’s identities
4. Methods for predicting victims of crimes

Category one, methods for predicting future crimes, focuses on predicting the times and places in which crimes are anticipated to occur. Category two, methods for predicting future offenders, focuses on predicting individuals and groups that are “at risk of offending in the future.” Category three, methods for predicting perpetrator’s identities, focuses on

identifying individuals who have committed past crimes. Lastly category four, methods for predicting victims of crimes, focuses on predicting individuals and groups that are “likely to become victims of crime.”

## 1.2 Due Process

**FiXme Fatal: Use Cornell dictionary, not OED.** Due Process is defined by the Oxford English Dictionary (US Edition) to be “[f]air treatment through the normal judicial system, especially as a citizens entitlement.” In modern liberal democracies (such as the United States), due process serves an important role; Due process protects citizens from unfair treatment under the law.

Liberal democracies have at the core of their philosophies the idea that every individual has equal protection of human rights, civil rights, civil liberties, and other political freedoms. As such, due process of law is necessary to protect the core tenants of this philosophy, and to protect the rights of citizens from governments that seek to violate the rights of its citizens. It ensures that the weak, poor and powerless have the same legal protections as the strong, wealthy and powerful, creating equal treatment under the law.

In the United States, due process of law is separated into two groups, procedural due process and substantive due process. Procedural due process focuses on rights that apply to government proceedings that may result in the denial of an individual’s right to life, liberty or property [4]. As part of procedural due process, it is well accepted that the following rights apply: [12]

1. An unbiased tribunal.
2. Notice of the proposed action and the grounds asserted for it.
3. Opportunity to present reasons why the proposed action should not be taken.
4. The right to present evidence, including the right to call witnesses.
5. The right to know opposing evidence.
6. The right to cross-examine adverse witnesses.
7. A decision based exclusively on the evidence presented.
8. Opportunity to be represented by counsel.
9. Requirement that the tribunal prepare a record of the evidence presented.
10. Requirement that the tribunal prepare written findings of fact and reasons for its decision.

Substantive due process is “[a] doctrine holding that the 5th and 14th Amendments require all governmental intrusions into fundamental rights and liberties be fair and reasonable and in furtherance of a legitimate governmental interest.” [6] One part of substantive due process is the application of the Bill of Rights, which contains as part of it the Fourth Amendment, or the protection against unreasonable searches and seizures by the government.

In modern policing, two legal standards keep the actions of law enforcement agencies in check, reasonable suspicion and probable cause. Reasonable suspicion is the standard

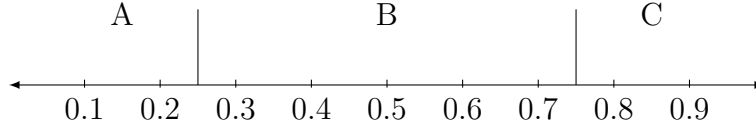


Figure 1: The number line separated into three sections by  $p_1$  and  $p_2$ .

that must be met in order for law enforcement officials to justify brief stops and detentions, but not full searches. [5] Reasonable suspicion must be based on “specific and articulable facts... taken together with rational inferences from those facts,” which must be associated with a particular individual. [7] In order to receive a warrant for a search or to make an arrest, a law enforcement agent must meet the standard of probable cause. Probable cause requires a “reasonable basis for believing that a crime may have been committed (for an arrest) or when evidence of the crime is present in the place to be searched (for a search).” [3] These standards keep the police from arbitrarily targeting individuals for investigation of crimes, instead requiring reasons for the infringement on their rights.

To frame this into the discussion of probabilities, we can think of redefining reasonable suspicion and probable cause as probability threshold values  $p_1$  and  $p_2$  respectively, where  $0 < p_1 < p_2 < 1$ , where the values of  $p_1$  and  $p_2$  are defined by societal norms, legislation and case law. (Picking these values is not trivial, though the abstraction is still useful; we will not discuss the values of  $p_1$  or  $p_2$  in this paper.) These values divide a number line on the interval  $[0,1]$  into three parts, as shown (with example values of  $p_1$  and  $p_2$ ) in Figure 1. If the perceived probability  $P$  (we will define this term contextually in both Sections 1.3 and 1.4) is greater than  $p_1$ , but less than  $p_2$ , (i.e.  $p_2 \geq P > p_1$ ), only the standard of reasonable suspicion is met. This can be seen as section B of the numberline. If the perceived probability is above  $p_2$  (i.e.  $P > p_2$ ), the standards of probable cause and reasonable suspicion are met. This can be seen as section C of the number line. This reflects the fact that reasonable suspicion is a lesser standard than probable cause. If the perceived probability is less than  $p_1$  (i.e.  $P \leq p_1$ ), neither probable cause nor reasonable suspicion are met. This corresponds to section A of the number line. Though this numerical definition of these two standards may seem overly complex, they are necessary to tie together the disciplines of law, psychology and computer science.

### 1.3 Cognitive Psychology

In order to understand judgements made by police before the creation of predictive policing, one must understand how the mind deals with decisions of a statistical nature. These ideas are detailed greatly by the work of Kahneman and Tversky, two psychologists who created and studied a useful model of human statistical thinking. This section briefly outlines the work done by Kahneman and Tversky in this field and is based entirely on Kahneman’s book *Thinking Fast and Slow*. [18]

In *Thinking Fast and Slow*, Kahneman describes his work with Tversky in different areas of psychology. The central thesis of the book, and of Kahneman and Tversky’s work itself, is the idea of a human’s thought process being separated into two modes of thought, called “system one and system two” both in the academic literature and in the book itself. In the book, Kahneman explains that system one is fast and effortless. It acts

as a statistician, counting occurrences and making quick predictions and reactions based on input stimuli. As such, it is trained over many years of input data and outcomes, calculating correlations and storing them for later use. System one is also responsible for expert reactions to domain specific problems. System two is the antithesis of system one. It is slow, effortful and logical, and it analyzes new problems and stimuli that system one was incapable of handling, or was proven to have handled incorrectly. Due to the effortful nature of system two, most decisions, predictions, judgements and reactions are made by system one, the quick and intuitive mode of thinking. In order to avoid activating the slow and effortful system two, there exist a list of heuristics that allow individuals to reach decisions using only system one, including framing, availability, loss aversion, and others.

Since system one is responsible for quick expert decisions, it is reasonable to believe that when a law enforcement agent attempts to assess whether a standard of probable cause or reasonable suspicion are met, they use the quick and intuitive “statistician” that is system one. To place this in context of our frame of probability, this can be viewed as a law enforcement agent estimates the perceived probability  $P = f(x_1, \dots, x_n)$ , that is the probability that an observed individual has committed or is going to commit a crime. This probability is estimated based on expert training (police training) and previous similar experiences. Then, the law enforcement agent estimates the values of  $p_1$  and  $p_2$  based on his or her understanding of the legislation and case law. Lastly, he or she compares the perceived probability  $P$  with the thresholds  $p_1$  and  $p_2$  as described in Section 1.2.

## **1.4 Machine Learning**

# **2 The Novelty of Predictive Policing**

## **2.1 Similarities**

## **2.2 Differences**

## **3 Verdict and Conclusion**

## References

- [1] Definition of due process in english:.
- [2] Due process.
- [3] Probable cause.
- [4] Procedural due process.
- [5] Reasonable suspicion.
- [6] Substantive due process.
- [7] Terry v. ohio, 1968.
- [8] L. Amore. Data derivatives on the emergence of a security risk calculus for our times. *Theory, Culture & Society*, 28(6):24–43, 2011.
- [9] D. K. Citron and F. A. Pasquale. The scored society: due process for automated predictions. *Washington Law Review*, 89, 2014.
- [10] K. Crawford and J. Schultz. Big data and due process: Toward a framework to redress predictive privacy harms. *BCL Rev.*, 55:93, 2014.
- [11] A. G. Ferguson. Predictive policing and reasonable suspicion. 2012.
- [12] H. J. Friendly. Some kind of hearing. *University of Pennsylvania Law Review*, 123(6):1267–1317, 1975.
- [13] L. Gitelman. *Raw data is an oxymoron*. MIT Press, 2013.
- [14] E. Goode. Sending the police before theres a crime, Aug 2011.
- [15] M. Hardt. How big data is unfair: Understanding sources of unfairness in data driven decision making, 2014.
- [16] M. Hildebrandt and K. De Vries. *Privacy, due process and the computational turn: the philosophy of law meets the philosophy of technology*. Routledge, 2013.
- [17] E. E. Joh. Policing by numbers: Big data and the fourth amendment. 2014.
- [18] D. Kahneman. *Thinking, fast and slow*. Macmillan, 2011.
- [19] W. J. Perry and C. M. Vest. Protecting individual privacy in the struggle against terrorists a framework for program assessment: Executive summary. *Journal of Privacy and Confidentiality*, 2(1):5, 2010.
- [20] W. L. Perry, B. McInnis, C. C. Price, S. C. Smith, and J. S. Hollywood. *Predictive policing: The role of crime forecasting in law enforcement operations*. Rand Corporation, 2013.
- [21] D. Robinson, H. Yu, and A. Rieke. Civil rights, big data, and our algorithmic future. *Robinson & Yu*, 2014.



- [22] J. Scahill and G. Greenwald. The nsa's secret role in the u.s. assassination program, Feb 2014.
- [23] D. o. R. S. South Dekota Department of Human Services. Lsi-r: Level of service inventory-revised introductory training course, Apr 2012.