The Struggle over Predictive Analytics in U.S. Urban Policing

A Sociotechnical Research Paper
presented to the faculty of the
School of Engineering and Applied Science
University of Virginia

by

Grady Roberts

April 6, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Grady Roberts

Sociotechnical advisor: Peter Norton, Department of Engineering and Society

The Struggle over Predictive Analytics in U.S. Urban Policing

In a 2015 panel on big data in cities, NYPD Commissioner Bill Bratton stated, "The 'Minority Report' of 2002 is the reality of today" (New York Times Events, 2015, 21:10). Police forces in many major U.S. cities deploy predictive analytics—tools for predicting future crime based on historical data. PredPol[1], the largest vendor of predictive policing algorithms in the U.S., claims that police departments use its product to police "one out of every 33 [Americans]" (PredPol, n.d.). The technology's proliferation has led its critics—civil liberties advocates, think tanks, constitutional rights lawyers, among others—to call for reforms and even abolition. According to a legal scholar, "The technology has far outpaced any legal or political accountability" (Ferguson, 2017, p. 1109). In the absence of regulation, predictive policing's proponents and opponents compete to influence its perception. This struggle is defined by the scale at which each group characterizes the technology. Proponents of predictive analytics cite its benefits for the community at large, while opponents warn of threats to individuals. Proponents argue that predictive analytics improve community safety by reducing bias and increasing policing accuracy and efficiency. Opponents argue that predictive analytics infringe on civil liberties and reify human bias with biased data.

Review of Research

In a study of Illinois governor George Ryan's 2003 decision to commute all death sentences, McCann (2007) found that pro- and anti-death penalty groups use the word "victim" differently. McCann classifies the speech into two categories: therapeutic and material

[1] In March of 2021 PredPol rebranded as Geolitica, "a mashup of 'geographical analytics'" to "better represent the direction" of the company (Geolitica, 2021).

victimhood. Therapeutic victimhood refers to an "affective strategy that emphasizes personal suffering and healing in the context of criminality" and material victimhood, in contrast, refers to a strategy "grounded in political, historical, and economic contextualization" (McCann, 2007, p. 383). These categories characterize the death penalty at different scales: the personal and the systemic. McCann does not ascribe one category to pro- and the other to anti-death penalty speech, claiming that speech from both groups can fit either category.

Van Kessel and Quinn (2020) surveyed U.S. adults for descriptions of how the COVID-19 pandemic has negatively affected them. The authors characterize the difference between Republican and Democratic sentiments towards mask-wearing. Like predictive policing, mask mandates have little legal precedent (Storace, 2020), leaving participant groups to struggle over their public perception. The authors found a strong political divide in mask rhetoric, with Republicans arguing that masks violate their personal rights and Democrats arguing that anti-maskers endanger their communities. The authors do not draw the conclusion that pro- and anti-mask groups use scale to frame their arguments, but the data support this contention.

Cardona (2020) analyzes the struggle between neoliberal ideologies and the public health response to COVID-19 in Australia. Cardona argues that a history of "personalization rhetoric" centering individual responsibility and freedom of choice in healthcare are the cause for anti-mask movements. Cardona claims that this rhetoric "shift[s] the blame for poor health outcomes from government to individuals" (Cardona, 2020, p. 1). While this work discusses the scale used to characterize healthcare, it only analyzes public policy rather than the views and rhetoric of groups trying to influence policy.

Kaufmann et al. (2019) contend that patterns are the "epistemological foundation" (p. 675) of predictive policing because they inform the way that algorithms process and 'understand' their

input data to make predictions about future crime. The authors argue that the personal values and ideas of what crime is vary between groups implementing predictive policing, leading to different types of patterns. While the authors find differences in the scale of data used by each pattern, they do not consider scale used in rhetoric around predictive policing. The authors do not discuss what conceptions of crime are common among opponents of the technology.

Bias

Proponents of predictive policing products argue that the technology is not susceptible to bias their models do not include demographic attributes. Jeffrey Brantingham is an anthropologist and a founder of PredPol. In a 2016 *American Scientist* interview, Brantingham argued that his company avoids bias and privacy concerns, stating "We work with event data. Because we're only looking at the what, where, and when of the crimes themselves, it's not actually about targeting people" (Saunders, 2016, p. 265). Brantingham adds "It's very difficult to figure out how do we actually use information about poverty or socioeconomic status as a component of models to forecast crime. All of that additional information is already built into the events themselves" (p. 264). These statements conflict—PredPol's algorithm is not biased because it is trained on demographic-free event data, but event data encodes demographic characteristics. This suggests that Brantingham believes bias occurs at the scale of the individual while his algorithm works at the scale of the community. Santa Cruz Police Department Deputy Chief Steve Clark argues "There's nothing in there about demographics" (Eidam, 2016).

Civil liberties advocates assert that predictive analytics reinforce human bias with biased datasets. Dr. Bernice King, CEO of The King Center for Nonviolent Social Change, stated in a February 2021 interview that "Technology reflects who we are. … We're recreating the whole

world that we currently have with all of the inequities" (OHUB, 2021, 11:30). Jamelle Watson-Daniels, Director of Research at Data for Black Lives, argues "We should not treat data as if it has been collected in a vacuum separate from society nor should we assume that algorithms are inherently fair or equitable simply because they are not human" (Watson-Daniels, qtd. in Finley, 2021). In a June 2020 notice to the American Mathematical Society, a group of over 1400 academics and professional mathematicians petitioned their colleagues to stop collaborating with police on predictive analytics projects, arguing "It is simply too easy to create a 'scientific' veneer for racism" and that "We are really hoping that this is a moment in which more of our colleagues are ready to accept ethical responsibility for the outcomes of algorithms or other applications of mathematics to policing" (Aougab et al., 2020). The petitioners use the scale of personal ethical responsibility to convince their colleagues to agree with their position.

Accuracy

Opponents of predictive policing center the potential consequences of inaccurate predictions, arguing that statistics on crime obscure individual harm. Regarding Chicago's Strategic Subject List (a list of predicted violent criminals), the Illinois ACLU stated "The list is not harmless. People on the list are subject to enhanced prosecutions and have armed police officers show up at their door. Enhanced prosecutions can lead to more time in prison for someone—a serious life-altering event" (Sheley, 2016). The ACLU centered the individual impacts of Chicago's SSL to argue for greater transparency in how the CPD decided who to include on the list. The ACLU uses the personal scale to encourage readers to empathize with those on the list and turn public opinion against it. In January 2020, Detroit police arrested Robert Williams and falsely accused him of theft based on an erroneous prediction by a facial

recognition algorithm (Hill, 2020). In a complaint to the Detroit Public Safety Headquarters, the Michigan ACLU argued that the police should not use this technology for investigations, stating:

> Robert Williams was arrested ... in his front lawn, as his wife Melissa looked on and as his daughters wept from the trauma of their father being taken from them. ... Mr. Williams' false arrest disrupted his family life, resulted in his unjustified jailing, and violated all norms of reasonable policing and investigation. (Mayor, 2020)

By describing Robert Williams' traumatic arrest, the ACLU encourages readers to imagine themselves in the same situation to turn them against the technology.

Opponents of predictive analytics believe that humans are too complex to for data-driven models to depict accurately. In a 2019 statement on the LAPD's use of PredPol, activists from the Stop LAPD Spying Coalition argued "[PredPol] wrongly assumes that human behavior can be operationalized and simplified into a universal mathematical model/algorithm; on the contrary, our lives and stories will not and cannot be datafied." The Stop LAPD Spying Coalition is appealing to the reader's individuality to argue that predictive analytics oversimplify human behavior, and that this will lead to inaccurate predictions. Dorothy Roberts, University of Pennsylvania professor and founding director of the Penn Program on Race, Science and Society, said in a July 2020 panel on predictive analytics:

> Risk assessment is no longer about determining whether an individual is going to do something, it's about whether the individual belongs to a population that the state wants to manage. ... law enforcement databases and algorithms are predicting that *toddlers* are going to be gang members. (Roberts, 2020, 27:21)

Roberts argues that predictive technology can not accurately predict individual criminality because it was not designed to do so (even if it claims to). To Roberts, predictive analytics is algorithmic justification for the police to target certain groups over others.

Proponents of predictive analytics argue that the technology can predict crime more accurately than human analysts. In the Los Angeles Police Department's five-year strategic plan, PredPol is praised for helping analysts be "100% more effective than they are with traditional hotspot mapping" (LAPD, 2015). The National Institute of Justice is an agency of the United States Department of Justice that invests in predictive policing research and promotes its use. According to a computer scientist at NIJ, analysts employing predictive analytics transitioned from basic crime mapping to crime forecasting and prediction (Hunt, 2019). Andrew Ferguson, author of "The Rise of Big Data Policing," said in an interview with *The Economist* that "New surveillance technologies let police map physical movements, digital communications and suspicious associations in ways that can reveal previously hidden patterns of criminal activity in otherwise overwhelming amounts of data" (Ferguson, 2018). Ferguson argues that modern police departments generate too much data for human analysts to discover crime patterns.

Efficiency

For its proponents, predictive analytics is an opportunity to more efficiently allocate policing resources. If police can predict where crimes will occur, they argue they can allocate more officers to this area to prevent the crime. In a 2015 statement to POLICE magazine, then-CEO of PredPol Larry Samuels claimed:

> By placing your officers in the right place at the right time, you will reduce crime in your community. You will also save considerable sums of money, as you will be able to return

> significant amounts of officer time that would be spent writing crime reports back into the community and that, too, will perpetuate crime decreases. (Griffith, 2015)

Samuels' goal is to improve the perception of PredPol among police officials to sell more software. When talking about PredPol's benefits, Samuels uses the scale of community safety by focusing on overall crime statistics. Palantir, another vendor of predictive policing technology, boasts that their product can help police "respond to crime as it happens" (Palantir, n.d.). Baltimore Police Commissioner Darryl De Sousa stated in a 2018 interview that "The whole concept behind predictive policing is to get in front of the crime before it even occurs" (Ritter, 2018). Speaking in support of predictive technology at a 2018 police conference, Deputy Attorney General Rod Rosenstein said:

> I visited Camden, New Jersey, to learn from Police Chief Scott Thomson how his department identifies potential hotspots and directs resources to cool them down. ... Chief Thomson rebuilt the department to focus on crime prevention. Murders in Camden declined by 70 percent, and violent crime fell by 39 percent. (Rosenstein, 2018)

Rosenstein uses high-level statistics about crime to convince his audience, police officials from around the country, that predictive analytics can help their departments.

In addition to reducing overall crime statistics and improving community safety, proponents argue that predictive analytics are necessary to alleviate overstretched departments. A crime analyst from the Santa Cruz police department, an early adopter of predictive analytics, said in 2011:

> We're facing a situation where we have 30 percent more calls for service but 20 percent less staff than in the year 2000 ... We have to deploy our resources in a more effective way, and we thought this model would help. (Goode, 2011)

San Francisco Police Department's chief of police, George Gascón, stated in 2010 that his department weathered recession budget cuts with predictive analytics: "With predictive policing, we have the tools to put cops at the right place at the right time or bring other services to impact crime, and we can do so with less" (Pearsall, 2010). A 2014 product feature in *The Police Chief* magazine boasts an "863 percent return on investment" in Memphis and that the city of Lancaster saved an "estimated $134 million" over four years after implementing predictive policing technologies (Harris, 2014). By framing money spent on predictive technology as an investment with significant return, the product feature appeals to police officials facing budget cuts.

Surveillance

Police officials avoid damaging public opinion of predictive analytics by fighting to keep the extent of its use for surveillance a secret. In San Diego, budget documents detailing how police spent $2 million on Palantir software over the last five years were labeled as "exempt from public disclosure" and only revealed after a records request (Marx, 2021). Further records requests about the nature of San Diego's use of Palantir have been denied, the city stating "The public interest in disclosure is outweighed by other factors. Disclosure of the information would chill the City's ability to have open and frank discussions about pre-decisional, deliberative matters" (Laurita, 2020). According to an investigation by *The Verge*, the New Orleans Police Department has secretly partnered with Palantir since 2012 without the knowledge or consent of the New Orleans city council (Winston, 2018). In New York, the NYPD fought the passage of the Public Oversight of Surveillance Technology (POST) Act, a law requiring 90 days of notification and a public comment period before the NYPD uses new surveillance technologies

(Winston, 2017). During a city council hearing on the POST Act, NYPD Deputy Commissioner of Intelligence and Counterterrorism John Miller argued:

> This proposal would require us to advertise sensitive technologies that criminals and terrorists do not fully understand. ... In effect, it would create a one-stop shopping guide to understanding these tools and how to thwart them for criminal elements and terrorists across the nation. (The New York City Council, 2017, 30:42)

Miller uses the threat of national terrorism to argue that the NYPD's use of these technologies is beyond the scale of New York and they should therefore keep them secret. In addition to fighting the POST Act, the NYPD fought a Freedom of Information Law request for documents on its contracts with predictive policing vendors for two years in court (Brennan Center for Justice, 2017).

Vendors of predictive analytics argue that it is not their responsibility to decide what constitutes excessive surveillance. Alex Karp, CEO of Palantir, stated in a 2019 opinion article:

> Silicon Valley companies are taking the power to decide these issues [regulation of predictive analytics] away from elected officials and judges and giving it to themselves. ... This is not the way consequential policy decisions should be made. I don't believe I should have that authority. (Karp, 2019)

By deferring to the legal system, Karp deflects critics who say Palantir has an ethical obligation to uphold the spirit of current privacy laws that do not specifically regulate predictive technology (Posner, 2019). Clearview AI is a vendor of facial recognition technology powered by a database of over three billion images scraped from online social media accounts. In addition to identifying people, Clearview can retrieve search subjects' addresses and associates (Hill, 2019). One Clearview investor, David Scalzo, stated:

> I've come to the conclusion that because information constantly increases, there's never going to be privacy. Laws have to determine what's legal, but you can't ban technology. Sure, that might lead to a dystopian future or something, but you can't ban it. (Hill, 2019)

Scalzo argues that issues of privacy are bigger than Clearview and that privacy violations are inevitable with technology. An employee in the Los Angeles county Chief Information Office working on a project to aggregate government data on residents said in an interview that "consent is anachronistic" (Brayne, 2020, p. 54). Vendors of predictive technologies expand their clients' surveillance capabilities while arguing that the only thing that can stop them is a legal system that can't keep up.

Opponents of predictive analytics argue that predictive policing technologies infringe on their civil liberties. Albert Cahn, Executive Director for the New York-based Surveillance Technology Oversight Project (STOP), spoke before the New York city council in support of the POST Act, stating "Transparency is crucial in every area of government, it is nowhere more vital than in policing, where mistakes can quickly rob New Yorkers of their liberty, or even their life" (Cahn, 2019). Police deploy Automatic License Plate Readers (ALPR) to track the movement of vehicles and generate data for some predictive algorithms, including Palantir (Palantir, n.d.). Hayley Tsukayama, an activist at the Electronic Frontier Foundation, stated in favor of a California law regulating ALPRs:

> Law enforcement agencies, with little oversight or regard for people's civil liberties or privacy, have for years used automated license plate readers, which track and record the movements of millions of ordinary people—most of whom are not connected to a crime. (Weiner, 2021)

The Lucy Parsons Labs is a Chicago-based activist group petitioning for a ban on facial recognition technology in the city. Their website states:

> By pressing pause on the disturbing trends in biometric surveillance, we help create a safer world where civil liberties will be better protected, and police officers will not arrest innocent people for crimes they did not commit due to inaccurate algorithms. (Lucy Parsons Lab, n.d.)

Opponents of predictive technologies for police surveillance speak in terms of the impact on individuals, encouraging their audience to empathize with victims.

Conclusion

Debate over predictive analytics in U.S. urban policing continues. Understanding how proponents and opponents view the technology and attempt to influence its use is an important step towards facilitating productive discussion on policing. All groups want fair and effective policing but disagree on what this looks like due to differences in the scale applied in their analysis. For police responding to large populations, effective policing is represented by trackable crime statistics. For those opposing predictive technology, fair and effective policing must be reflected in the lived experiences of individuals in their communities. These findings reveal the danger of a disconnect between those implementing policy and those experiencing its effect, suggesting that a bottom-up community-driven approach may have more success for policies with wide-reaching impact. Further research is needed to determine if the findings hold in smaller jurisdictions where police may be better integrated with the community.

References

Aougab, T., Ardila, F., Athreya, J., Goins, E., Hoffman, C., Kent, A., Khadjavi, L., O'Neil, C., Patel, P., & Wehrheim, K. (2020, June 15). *Boycott collaboration with police*.

Brayne, S. (2020). *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press.

Brennan Center for Justice (2017, May 19). *Brennan Center for Justice v. New York Police Department.*

Cahn, A. F. (2019). *Testimony before NYC Council Committee on Technology*.

Cardona, B. (2020). The pitfalls of personalization rhetoric in time of health crisis: COVID-19 pandemic and cracks on neoliberal ideologies. *Health Promotion International*.

Eidam, E. (2016, September). *The Role of Data Analytics in Predictive Policing*. Government Technology.

Ferguson, A. G. (2018, June 4). How data-driven policing threatens human freedom. *The Economist*.

Ferguson, A. G. (2017). Policing Predictive Policing. *Washington University Law Review, 94*(5), 1109–1190.

Finley, K. (2021, February 18). *Black voices bring much needed context to our data-driven society.* The ReadME Project.

Geolitica: A new name, a new focus. (2021, March 2). Geolitica.

Goode, E. (2011, August 16). Sending the Police Before There's a Crime. *The New York Times.*

Griffith, D. (2015, June 30). *Predictive Policing: Seeing the Future*. POLICE Magazine.

Harris, S. (2014). Product Feature: Predictive Policing Helps Law Enforcement "See Around the Corners." *Police Chief Magazine, 81,* 44–45.

Hill, K. (2020, January 18). The Secretive Company That Might End Privacy as We Know It. *The New York Times*.

Hill, K. (2020, June 24). Wrongfully Accused by an Algorithm. *The New York Times*.

Hunt, J. (2019). From Crime Mapping to Crime Forecasting: The Evolution of Place-Based Policing. *National Institute of Justice Journal, 281,* 96–102.

Karp, A. (2019, September 5). I'm a tech CEO, and I don't think tech CEOs should be making policy. *Washington Post*.

Kaufmann, M., Egbert, S., & Leese, M. (2019). Predictive Policing and the Politics of Patterns. *The British Journal of Criminology*, *59*(3), 674–692.

Kessel, P. van, & Quinn, D. (2020, October 29). *Both Republicans and Democrats cite masks as a negative effect of COVID-19, but for very different reasons*. Pew Research Center.

LAPD in 2020: The Safest Big City In America. (2015). LAPD.

Laurita, A. (2020, November 1). *Public Record Request #20-4435*. City of San Diego.

Biometric Privacy. (n.d.). Lucy Parsons Labs.

Marx, J. (2021, February 2). *Local Law Enforcement Quiet on Relationships With 'Predictive Policing' Company*. Voice of San Diego.

Mayor, P. (2020, June 24). *ACLU of Michigan Complaint re Use of Facial Recognition*. American Civil Liberties Union.

McCann, B. J. (2007). Therapeutic and Material <Victim>hood: Ideology and the Struggle for Meaning in the Illinois Death Penalty Controversy. *Communication and Critical/Cultural Studies*, *4*(4), 382–401.

New York Times Events. (2015). *Cities for Tomorrow 2015—Data Mining in the Modern City*.

OHUB. (2021, February 8). *OHUB x Blacktocats: Black History Month Experience with Dr. Bernice King*.

Palantir. (n.d.). *Law Enforcement*. Palantir.

Pearsall, B. (2010). Predictive Policing: The Future of Law Enforcement? *National Institute of Justice Journal*, *266*, 16–19.

Posner, M. (2019, September 12). *How Palantir Falls Short of Responsible Corporate Conduct*. Forbes.

PredPol. (n.d.). *PredPol Mission | About Us | Aiming to reduce victimization keep communities safer.*

Ritter, R. (2018, February 28). *Baltimore Leaders Turn To Predictive Policing To Prevent Crime*.

Roberts, D. (2020, July 8). *Policing Without the Police: Race, Technology and the New Jim Code*. Haymarket Books.

Rosenstein, R. (2018, December 5). *Deputy Attorney General Rod J. Rosenstein Delivers Remarks at the Project Safe Neighborhoods National Conference*.

The New York City Council (2017, June 14). *Meeting of Committee on Public Safety on 6/14/2017 at 10:00 AM*.

Saunders, F. (2016). First Person: Andrea L. Bertozzi and P. Jeffrey Brantingham. *American Scientist, 104*(5), 264–265.

Sheley, K. (2016, June 7). *Statement on Predictive Policing in Chicago*. ACLU of Illinois.

Stop LAPD Spying Coalition. (2019, October 15). *Abolish Carceral Technologies: The Peoples Response*.

Storace, R. (2020, December 10). *What's the Law Surrounding Biden's Proposed Mask Mandate?* Connecticut Law Tribune.

Weiner, S. (2021, January 21). Senator Wiener Introduces License Plate Privacy Act. *East County Today*.

Winston, A. (2017, July 7). *NYPD Attempts to Block Surveillance Transparency Law With Misinformation*. The Intercept.

Winston, A. (2018, February 27). *Palantir has secretly been using New Orleans to test its predictive policing technology*. The Verge.