The Automated Tipster: How Implicit Bias Turns Suspicion Algorithms into BBQ Beckys

Christine Kumar*

TABLE OF CONTENTS

* J.D., May 2020, The George Washington University Law School, May 2020. Notes Editor, Federal Communications Law Journal, Vol. 72. B.A. Political Science, Johns Hopkins University, 2016. This Note is dedicated to my parents, Hannah and Suresh Kumar, and my sister, Cynthia Kumar, for their continued support and love and for inspiring me everyday. I would like to thank Dean Renée McDonald Hutchins for her guidance and scholarship on this subject as well as the Federal Communications Journal Vols. 71-72 staff for their work on this Note. I would also like to thank Emily S. Haselton for her work and support while writing this Note. And finally I would like to thank Daniel Wolman, Caitlyn Kretzschmar and Jarrod Carman for their support on this Note and throughout my time at GW.

I. INTRODUCTION

#BBQBecky, one of several satirical hashtags that went viral this summer, refers to Jennifer Schulte, a white woman who called the police on a group of black people barbecuing at an Oakland park after the group refused to stop grilling.[1] The group of black people were using a charcoal grill in a non-charcoal designated area, a rarely enforced rule considered "not a police matter."[2] The video, which captured only twenty minutes of a three-hour phone call Ms. Schulte had with the police department, shows Ms. Schulte on the phone reiterating that this specific area is not designated for charcoal grilling.[3] Despite the hesitations from dispatch and even a recommendation that Ms. Schulte be evaluated for a temporary psychiatric hold, the police eventually arrived and questioned the group of grilling black residents for over an hour.[4] While no arrests were made or citations issued, this successful dispatch of police for implicitly racially motivated reasons, involving no emergency, is just one example of how police are operating on, or are led to operate on, racist biases.[5] In fact, several hashtags, including #BBQBecky, #PermitPatty, and #CornerstoreCaroline, have become part of a national discourse as a response to these videos capturing white people alerting the police to, or threatening to call the police on, people of color doing similarly minor or non-criminal activities, including sleeping in their own dormitory or sitting at a Starbucks.[6] While these hashtags do bring some levity to these overt demonstrations of racism, the videos they are inspired by expose the weaponization of law enforcement against people of color in order to police social norms instead of actual criminal conduct.[7] This kind of policing—unnecessary interactions with lawful black Americans based on racial biases—concerningly results in increased instances of police violence leading to countless murders of innocent, non-threatening black Americans by police

1. Tom Cleary, *Jennifer Schulte, 'BBQ Becky': 5 Fast Facts You Need to Know*, HEAVY.COM (Jun. 23, 2018, 5:54 PM), https://heavy.com/news/2018/05/jennifer-schulte-bbq-becky/.

2. *Id.*

3. *Id.*; *see also* Hilary Hanson, *Listen to Full 911 Audio of 'BBQ Becky' Calling Cops on Black Men Grilling*, HUFFINGTON POST (Sept. 2, 2018, 1:56 PM), https://www.huffingtonpost.com/entry/bbq-becky-911-calls-grill_us_5b8c0f07e4b0162f4724a74c.

4. Hanson, *supra* note 3.

5. *See id.*; Jeffery Robinson, *Let's Address the Ridiculous 911 Calls that People of Color Endure*, THE HILL (June 11, 2018), https://thehill.com/opinion/civil-rights/391639-lets-address-the-ridiculous-911-calls-that-people-of-color-endure.

6. Jessica Guynn, *BBQ Becky, Permit Patty and Why the Internet is Shaming White People Who Police People 'Simply for Being Black*,*'* USA TODAY (July 23, 2018, last updated 12:17 PM), https://www.usatoday.com/story/tech/2018/07/18/bbq-becky-permit-patty-and-why-internet-shaming-white-people-who-police-black-people/793574002/; Gina Martinez, *Woman Dubbed 'Cornerstore Caroline' Faces Backlack After Falsely Accusing a 9-Year-Old Boy of Sexual Assault*, TIME (Oct. 16, 2018), https://time.com/5426067/cornerstore-caroline-backlash-sexual-assault-boy/.

7. Robinson, *supra* note 5.

officers.[8] In fact, this kind of police brutality and subsequent lack of justice, such as the acquittal of Trayvon Martin's killer, has sparked the Black Lives Matter movement.[9]

Under the Fourth Amendment, an officer must be able to demonstrate reasonable articulable suspicion (RAS) in order to legally conduct a stop under *Terry v. Ohio*.[10] Generally, given the risk of consequences for a false report, tips from a reliable source are sufficient to meet the RAS standard in order to conduct a stop, whereas anonymous tips require further corroboration in order to meet the standard.[11] The recent trend in inflammatory 911 calls made by white people against black people doing non-criminal activities demonstrates the need for higher scrutiny from police responding to 911 calls. Most states criminalize those who make false police reports with punishments including misdemeanor or felony charges, and several jurisdictions are even starting to introduce legislation that would criminalize 911 calls against people of color when there is no evidence of wrongdoing.[12]

While greater attention is being paid to these discriminatory calls to the police, both publicly and in terms of legal action, other prejudiced alerts to the police are not always as flagrant as #BBQBecky. With the rise of Big Data and data mining,[13] law enforcement agencies have begun to use computers and software in order to aid in crime prevention.[14] More specifically, "predictive policing," an evolving trend in law enforcement, uses data analyses and criminology theories in order to create models that can anticipate when or where a crime will occur.[15] In theory, these models, which are based on numerical statistics and scientific data, should offer neutral and accurate findings and thus lead to a more efficient and better-informed criminal justice

---

8. I. Bennett Capers, *Race, Policing, and Technology*, 95 N.C. L. REV. 1241, 1254 (2017); Ryan W. Miller, *Black Lives Matter: A Primer on What it Stands for*, USA TODAY (July 11, 2016, 9:53 PM), https://www.usatoday.com/story/news/nation/2016/07/11/black-lives-matter-what-what-stands/86963292/.

9. *Id.*

10. Terry v. Ohio, 392 U.S. 1 (1968).

11. Florida v. J.L., 629 U.S. 266, 270 (2000).

12. *See generally* S.C. CODE ANN. § 16-17-722 (2012); CAL. GOV'T CODE § 53153.5 (2012); MICH. CODE § 750.411a (2012); N.J. CODE § 2C:28-4 (2012); *see also* Morgan Gstalter, *NY State Senator Wants to Criminalize Calling 911 on Law-Abiding Black People*, THE HILL (Aug. 16, 2018, 4:00 PM), https://thehill.com/homenews/state-watch/402211-ny-state-senator-wants-to-criminalize-calling-911-on-law-abiding-black.

13. Steve Lohr, *How Big Data Became So Big*, N.Y. TIMES (Aug. 11, 2012), https://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html?auth=login-smartlock ("Big Data is a shorthand label that typically means applying the tools of artificial intelligence, like machine learning, to vast new troves of data beyond that captured in standard databases. The new data sources include Web-browsing data trails, social network communications, sensor data and surveillance data."); Daniel J. Steinbock, *Data Matching, Data Mining, and Due Process*, 40 GA. L. REV. 1, 3 (2005) (Data mining is the "computer application of statistical formulas to large bodies of data to identify relationships or patterns," specifically "identifying people who fit a designated computer-generated profile.").

14. Michael L. Rich, *Machines as Crime Fighters*, 30 CRIM. JUST. 10 (2016).

15. Lindsey Barrett, *Reasonably Suspicious Algorithms: Predictive Policing at the United States Border*, 41 N.Y.U. REV. L. & SOC. CHANGE 327, 334-35 (2017).

system.[16] However, the software and tools used by the police, including something known as Automated Suspicion Algorithms (ASAs), have demonstrated that these technologies inherit and further perpetuate implicit biases that could lead to discriminatory police alerts similar to those made by a #BBQBecky or a #PermitPatty.[17] However because these predictive policing technologies are trusted to theoretically eliminate any bias involved with human judgment, they are able to function without any of the scrutiny or risk of punishment that bigoted tipsters are now facing when they wrongfully alert the police.[18]

This Note will address the issue of how the police, in employing Big Data tools like data mining and machine learning, perpetuate discriminatory and harmful policing biases similar to stereotypical 911 callers, but without any of the scrutiny or legal recourse available to remedy these prejudices. It will argue that in order to prevent discrimination in predictive policing, the police should abide by the reasonable articulable suspicion standard as applied to anonymous tips before stopping an individual pegged by a predictive policing tool.[19] More specifically, this Note calls for any tip generated from one of these technologies to require support by further corroboration before being acted upon. Finally, this Note will argue that there should be a civil remedy available, namely under the product liability framework, against these technology developers, so as to deter discriminatory models and offer victims targeted by racial bias the opportunity to address their grievances.

Section II will begin looking at the foundations of implicit bias and how it has infiltrated machine learning technologies such that once it has been inputted into the algorithms, it is nearly impossible to extract, and thus perpetuates racist thinking. Section III will explore two possible solutions to this issue: first, this Note will argue that the government should be held accountable for the employment of these algorithms. Given the nature of machine learning and the constant consumption of information that continually adjusts the algorithm, machine learning technologies should be held to the standard courts use for anonymous tips when police are mobilized and should by default be seen as unreliable unless supported by corroboration.[20] Furthermore, to better ensure transparency in what data is being used to train and guide these algorithms, this Note will argue for data protection legislation similar to specific provisions regarding information and access to personal data passed within the European Union's recent data protection legislation in order to give transparency to individuals whose

16. *See* Lindsey Patterson, *How the Evolution of Big Data Is Influencing Law Enforcement*, TECHNOLOGY.ORG (July 28, 2017), https://www.technology.org/2017/07/28/how-the-evolution-of-big-data-is-influencing-law-enforcement/; Andrew Ferguson, *Is "Big Data" Racist? Why Policing by Data Isn't Necessarily Objective*, ARS TECHNICA (Dec. 29, 2017, 7:23 AM), https://arstechnica.com/tech-policy/2017/12/is-big-data-racist-why-policing-by-data-isnt-necessarily-objective/ [https://perma.cc/XW2A-ZARD].

17. *See* Rich, *supra* note 14, at 13.

18. *See* Barrett, *supra* note 15, at 341-42.

19. *See* Terry v. Ohio, 392 U.S. 1 (1968).

20. *See* Florida v. J.L., 629 U.S. 266 (2000).

personal data is being used by third party software companies in these algorithms. Second, this Note will argue that software developers should be held accountable if their technologies are found to be based on discriminatory and racially biased data or models. This accountability should be imposed through a products liability scheme that focuses on strict liability, as opposed to negligence, in order to impose legal obligations on these software developers and compel them to take affirmative, preventative action when creating these algorithms.

Ultimately, the growing reliance on machine learning technologies by law enforcement agencies requires greater knowledge and review of how these technologies function and better assurances to check if they are being employed in a manner that is neutral and accurate, both mechanically and legally. While the issue of #BBQBecky and others who wrongfully mobilize the police continues to be a problem, technologies that use similarly biased thinking and formulations often go unquestioned or unchallenged, and thus require heavier scrutiny as we enter a more data-driven society. Implicit bias, and the racist behavior it leads to, should be curbed, not further reinforced by technological advancements, in order to prevent unnecessary and racially motivated interactions with the police that generate violence against black Americans.

II. THE WRONGFUL MOBILIZATION OF THE POLICE: HOW IMPLICIT BIAS IN HUMANS AND TECHNOLOGIES CAN INFLUENCE POLICING

Subsection A will first look at implicit bias and how it infiltrates daily interactions between civilians and between civilians and the police. Then, Subsection A will examine issues of "profiling by proxy" as demonstrated by #BBQBecky and others who wrongfully mobilize the police.[21] Subsection B will then discuss the new forms of technology based on Big Data and machine learning that law enforcement is increasingly relying upon for crime detection and prevention, and the similar permeation of the same kind of implicit bias in the data and models used to create those technologies.

Despite inherently involving the bias that leads to the outrage with #BBQBecky and other bigoted tipsters, the results of these technologies go unquestioned without any legal pushback or standard, therefore making their effects potentially even more dangerous in terms of racial discrimination in law enforcement.

21. *See* Lisa Thurau & Bob Stewart, *Avoiding 'Profiling by Proxy*,*'* VERA INSTITUTE OF JUSTICE: THINK JUSTICE BLOG (Mar. 13, 2015), https://www.vera.org/blog/police-perspectives/avoiding-profiling-by-proxy.

A. Implicit Bias in Human and Police Interactions

Racial disparities within the criminal justice system are well known—black and Latino people are disproportionately stopped, prosecuted, and incarcerated as compared to white people who commit the same crimes, despite black and Latino populations making up a smaller percentage of the population.[22] Despite the fact that racial bias in the criminal justice system is both pervasive and well-studied, the racism that permeates interactions between minorities and law enforcement is not necessarily obvious or even cognizable by the person or group perpetrating it. Implicit bias is the internalization of stereotypes and perspectives concerning other races or people that can often unintentionally lead to discriminatory behavior.[23] Even in spite of "avowed or endorsed beliefs or principles," a person can unconsciously "activate" a network of negative stereotypes when confronted with pictures, symbolic representations, or members of a stereotyped group, which in turn can lead to discriminatory behavior.[24] More specifically, studies have shown that when exposed to African-Americans or their perceived "culture," such as seeing a member of that group or listening to a certain type of music or language often attributed to that group, participants have displayed "negative emotional arousal" and "evidence of self-regulatory or executive control activity" in response.[25]

Implicit bias is remarkably prevalent in police interactions with the public. Different from overt racism, which is the conscious activation of harmful stereotypes that manifest in outwardly racist behavior, implicit bias is much more difficult to detect and therefore harder to correct.[26] For example, overt racist police behavior would be a police officer deciding "I'm stopping all black people," while implicit bias would be a police officer deciding "I'm stopping all dangerous people," which in effect would be the stopping of only black people due to the officer's unconscious conflation of black with dangerous.[27] Implicit bias in the police is particularly perilous when police officers tend to use more force against black Americans because of latent stereotypical thinking that a black person needs more force to be subdued as

---

22. *See* Radley Balko, *There's Overwhelming Evidence that the Criminal-Justice System Is Racist. Here's the Proof*, WASH. POST (Sept. 18, 2018), https://www.washingtonpost.com/news/opinions/wp/2018/09/18/theres-overwhelming-evidence-that-the-criminal-justice-system-is-racist-heres-the-proof/?noredirect=on&utm_term=.bb3ceddb9a16#section9.

23. *See* Anthony G. Greenwald & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations*, 94 CALIF. L. REV. 945, 951 (2006).

24. *Id.*; *See* Robert J. Smith & Justin D. Levinson, *The Impact of Implicit Racial Bias on the Exercise of Prosecutorial Discretion*, 35 SEATTLE U. L. REV. 795, 798-801 (2012).

25. *Id.*

26. Katherine Lee Goyette, *Implicit Bias & Police Encounters*, 87 J. KAN. B. ASS'N 9, 19 (2018); Megan Quattlebaum, *Let's Get Real: Behavioral Realism, Implicit Bias, and the Reasonable Police Officer*, 14 STAN. J. C.R. & C.L. 1, 7 (2018) ("[R]acial profiling will be defined as 'the use of race or ethnicity, or proxies thereof, by law enforcers as the basis of judgments of criminal suspicion,' except with trustworthy information, relevant to the locality and timeframe, that links a person of a particular race or ethnicity to an identified criminal incident or scheme.").

27. *Id.*

compared to their white counterpart.[28] Accordingly, implicit bias has led to black Americans being one-third more likely to be stopped by the police, three times more likely to be searched by the police, and three times more likely to have been the subject of force as compared to their white counterparts.[29] Implicit bias has also led to a phenomenon of "stereotype threat," where someone concerned about being perceived as part of a negative stereotype subconsciously acts in accordance with that negative association, further affirming and perpetuating negative stereotypes in the mind of a police officer and unintentionally escalating already dangerous situations.[30] For example, a black individual may become more self-regulatory in their actions when interacting with an officer and may subsequently exhibit behavior that a police officer would view as "deceptive" or that may affect the individual's ability to resist pressure in interrogative situations.[31]

Not only does implicit bias guide police interactions, from both the perspectives of the officer and the suspected individual, implicit bias is also part of the mobilization of the police. "Profiling by Proxy" is a phenomenon in which individuals alert the police to false claims of misconduct by people or groups of people against whom these callers are biased or dislike.[32] Some of the more notable cases of profiling by proxy involved direct racial profiling such as those of "Permit Patty," where a women called the police on 8-year-old for not having a permit to sell water on a corner,[33] "Cornerstore Caroline," where a woman called the police to report being "sexually assaulted" by a child because his backpack brushed against her back while he walked behind her,[34] or the white women who called 911 on a black Yale graduate student taking a nap in her dormitory's common room.[35] However, incidents like

28. Dakshana Bascaramurty, *Implicit Bias Linked to Lethal Police Force, Research Suggests*, THE GLOBE AND MAIL (July 2017).

29. Tom James, *Can Cops Unlearn Their Unconscious Biases?*, THE ATLANTIC (Dec. 23, 2017), https://www.theatlantic.com/politics/archive/2017/12/implicit-bias-training-salt-lake/548996/.

30. Cynthia J. Najdowski et al., *Stereotype Threat and Racial Differences in Citizens' Experiences of Police Encounters*, 39 LAW & HUM. BEHAV. 463, 464 (2015).

31. *Id.*

32. Thurau & Stewart, *supra* note 21.

33. Kalhan Rosenblatt, *White Woman Dubbed "Permit Patty' for Calling Police on Black Girl Denies It was Racial*, NBC NEWS (June 25, 2018, 9:30 AM), https://www.nbcnews.com/news/us-news/white-woman-dubbed-permit-patty-calling-police-black-girl-denies-n886226.

34. Ryan Grenoble, *White Woman Apologizes for Falsely Reporting That a Black Boy Groped Her*, HUFFINGTON POST (Oct. 17, 2018, 5:56 PM), https://www.huffingtonpost.com/entry/cornerstone-caroline-black-boy-false-grope_us_5bc785d5e4b055bc947d04ac.

35. Alan Pyke, *A Black Yale Grad Student Took a Nap in her Dorm's Common Room, and a White Woman Called the Cops*, THINK PROGRESS (May 9, 2018, 3:54 PM), https://thinkprogress.org/white-woman-calls-cops-on-black-yale-grad-student-for-crime-of-napping-in-a-common-room-10826f736ce3/.

"BBQ Becky,"[36] and the Starbucks employees in Philadelphia who asked two black men to leave for sitting in the store without purchasing anything,[37] display how implicit bias—or, in these cases, internalized racist and stereotypical views of black people that inspired unwarranted feelings of fear and suspicion—can trigger unnecessary and potentially even more discriminatory or harmful police interactions.[38] Given how dire the current situation is between officers and minorities in the United States, addressing implicit bias at the outset of police interactions—specifically how they are getting information that leads to racially targeting individuals—is crucial in order to begin correcting overall bias in policing.

B. Big Data, Machine Learning and the Police

Implicit and explicit biases, specifically concerning an individual's race, are embedded in all aspects of police interactions—from what mobilizes the police to who is arrested. Increasingly, law enforcement agencies have been looking to technological tools in order to aid in crime detection and prevention, with the hope that relying on technology, as opposed to human judgement and bias, can lead to more efficient, and more accurate, results.[39] The rise of Big Data[40] has led to novel methods of discovering latent information within this mass accumulation of data in order to better inform and guide how businesses, service providers, and the government can function more effectively without constant human supervision or, in some cases, human input.[41] One such method is data mining, which is the computerized application of statistical information to these large aggregations of data in order to find relationships or patterns, or to identify people who meet a certain

36. Christina Zhao, *'BBQ Becky,' White Woman Who Called Cops on Black BBQ, 911 Audio Released: 'I'm Really Scared! Come Quick!'*, NEWSWEEK (Sept. 4, 2018, 5:42 AM), https://www.newsweek.com/bbq-becky-white-woman-who-called-cops-black-bbq-911-audio-released-im-really-1103057.

37. Scott Neuman, *Men Arrested in Philadelphia Starbucks Reach Settlements*, NPR (May 3, 2018, 1:06 AM), https://www.npr.org/sections/thetwo-way/2018/05/03/607973546/men-arrested-in-philadelphia-starbucks-reach-settlements.

38. Robert J. Smith, *Reducing Racially Disparate Policing Outcomes: Is Implicit Bias Training the Answer?*, 37 U. HAW. L. REV. 295, 298 (2015) ("Biases could shape whether an officer decides to stop an individual for questioning in the first place, elects to interrogate briefly or at length, decides to frisk the individual, and concludes the encounter with an arrest versus a warning.").

39. Christopher Rigano, *Using Artificial Intelligence to Address Criminal Justice Needs*, NATIONAL INSTITUTE OF JUSTICE: NIJ JOURNAL (Jan. 2019), https://www.ncjrs.gov/pdffiles1/nij/252038.pdf.

40. Neil M. Richards & Jonathan H. King, *Three Paradoxes of Big Data*, 66 STAN. L. REV. 41, 42 (2013) ("Big data analytics depend on small data inputs, including information about people, places, and things collected by sensors, cell phones, click patterns, and the like. These small data inputs are aggregated to produce large datasets which analytic techniques mine for insight.").

41. Savan Patel, *Chapter 0: What is Machine Learning?*, MEDIUM, MACHINE LEARNING 101 (Apr. 29, 2017), https://medium.com/machine-learning-101/chapter-0-what-is-machine-learning-ad136361c618.

profile.[42] These discovered relationships, patterns, or profiles are then collected into "models," which create automatic processes in order to classify and organize a particular interest in a practice called "machine learning."[43] Machine learning, in other words, is artificial intelligence that uses relationships, patterns, or profiles found within historical data sets and applies that information in new and unpredictable data sets without much or any human supervision.[44] For example, law firms have begun using machine learning software for document review in the discovery phase of litigation to evaluate large numbers of documents, flag those that are relevant, and send those that are "questionable" to the attorney, all without needing a paralegal or other human oversight.[45] Remarkably, these algorithms continue to learn based on analyzing new data sets and thus "improve their performance on a task with experience."[46] Therefore, those "questionable" documents sent to the attorney who then determines whether the document is relevant or not will be absorbed by the algorithm to better understand what "relevant" means, or could mean, eventually leading to a lower number of "questionable" documents and even less need for attorney input.

Machine learning technology is beginning to transform law enforcement through "predictive policing," which offers information based on criminological theories applied to police records, camera footage, and other information already in use or available to the police to use for both crime detection and crime prevention.[47] For example, officers in Santa Cruz, California are using an algorithm that gives beat cops "crime forecasts" for that particular day in a particular area, so that those officers can then include that area in their patrol and help suppress or uncover any potential crime.[48] An example of this crime forecast could be something like "there is a 10.36% likelihood of a car theft in a particular downtown garage on a particular day. The times when those car thefts are most likely to occur are also listed." "CrimeScan" is another such piece of predictive policing software that offers geographical location information based on past police data in order to preempt high violence crimes.[49] With the assumption that "violent crime is

---

42. Daniel J. Steinbock, *Data Matching, Data Mining, and Due Process*, 40 GA. L. REV. 1, 3 (2005).

43. Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 677-78 (2016).

44. Laura A. Odell et al., *A State Cyber Hub Operations Framework*, INSTITUTE FOR DEFENSE ANALYSES (June 2016), https://apps.dtic.mil/dtic/tr/fulltext/u2/1013836.pdf.

45. Bernard Marr, *How AI and Machine Learning are Transforming Law Firms and the Legal Sector*, FORBES (May 23, 2018, 12:29 AM), https://www.forbes.com/sites/bernardmarr/2018/05/23/how-ai-and-machine-learning-are-transforming-law-firms-and-the-legal-sector/#7fea06e932c3.

46. Rich, *supra* note 14, at 10.

47. Barrett, *supra* note 16, at 334-35.

48. Andrew Guthrie Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L.J. 259, 268 (2012).

49. Randy Rieland, *Artificial Intelligence is Now Used to Predict Crime. But Is It Biased?*, SMITHSONIAN MAGAZINE (Mar. 5, 2018), https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/.

like a communicable disease," CrimeScan takes in police records ranging from prior high violence crimes, past 911 calls involving gun possession or shots fired, and temporal information regarding high violence crimes in order to build a type of map where high violence is likely to occur, but stops before attempting to predict who might commit these crimes.[50]

Predictive policing, however, has in fact gone beyond just offering geographical or temporal predictions of when crime could occur and now has begun offering the likelihood that an individual is in the progress of committing a crime.[51] "Automated Suspicious Algorithms" (ASAs) determine the likelihood that an individual is engaging in criminal activity and generate probabilistic guesses about an individual's level of suspicion.[52] Based on the historical data already available to the police—for example, description of a suspect and facts of an earlier arrest for cocaine possession—and based on continual learning through ongoing experiences in executing the algorithm—for example, whether individuals tagged by the ASA for similar crimes were arrested or not—a computer can identify patterns that typically signal when an individual is engaging in hand-to-hand cocaine sales.[53] The computer then alerts the police when this level of suspicion reaches a programmed level of confidence, such as a numerical predictive percentage, which can lead to officers being dispatched to the individual that is allegedly dealing cocaine.[54] The computer, therefore, is determining an individual's level of suspicion, as opposed to a bigoted tipster that inheres latent motivations biased against a certain class of people. This kind of suspicion software is already in use in several police departments. In Chicago, the "Strategic Subject List" (SSL) is one working example of an ASA currently in use whose algorithm generates risk scores and identifies individuals who are at the highest risk of danger, either as the victim or the perpetrator of a crime.[55] Factors to determine this score include the individual's prior arrests, including for violent offenses and narcotics, and the number of times an individual was the victim of a shooting or aggravated battery or assault.[56] However, because the workings of the algorithm have not been disclosed, it is unclear how heavily each factor weighs in how the score is actually generated.[57] Individuals who have a score of 250 or above are then tagged by

---

50. *Id.*; Adam Mann, *How Science Is Helping Stop Crime Before It Occurs*, NBC NEWS (Oct. 6, 2017), https://www.nbcnews.com/mach/science/how-science-helping-stop-crime-it-occurs-ncna805176.

51. Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 876 (2016).

52. *Id.*

53. Rich, *supra* note 14, at 10.

54. *Id.* (". . . there is a 62 percent chance the highlighted individual is currently engaged in hand-to-hand cocaine transactions.").

55. Jeff Asher & Rob Arthur, *Inside the Algorithm that Tries to Predict Gun Violence in Chicago*, N.Y. TIMES (June 13, 2017), https://www.nytimes.com/2017/06/13/upshot/what-an-algorithm-reveals-about-life-on-chicagos-high-risk-list.html.

56. *Id.*

57. Brianna Posadas, *How Strategic is Chicago's "Strategic Subjects List"? Upturn Investigates.*, MEDIUM (June 22, 2017), https://medium.com/equal-future/how-strategic-is-chicagos-strategic-subjects-list-upturn-investigates-9e5b4b235a7c.

the list and put "on [the Chicago Police Department's] radar," although it is not clear how the police actually use this information.[58] Ironically, when the *New York Times* conducted a study based on the information the Chicago Police Department released to the public, they discovered that violence in Chicago is not necessarily concentrated among those with the highest SSL risk scores and that despite the use of the SSL for years, Chicago is still contributing to a large share of increasing urban murders country-wide.[59] The algorithm has, therefore, had minimal, if any, positive effect on crime prevention and reduction.

In theory, the use of a machine to determine an individual's level of suspicion should offer a more neutral, informed, and reliable outcome than a human who is prone to implicit biases and has the potential for bad motivations. However, despite the fact that these algorithms are based on numerical data from past police records and studies and the lack of human supervision in the machine's perpetual learning, these algorithms are just as likely to inhere implicit biases as their human counterparts but, concerningly, without the scrutiny or legal mechanisms to counter these biases.[60] Machine learning AI is typically viewed as a "black box:" although it can learn and "make predictions and decisions as humans do," it does so "without being able to communicate its reasons for doing so" and in methods that humans may not be able to comprehend.[61] In addition to this general lack of knowledge about what actually motivates an algorithm's decision making, implicit bias can easily permeate machine learning processes (and be further perpetuated by machine learning), specifically when it comes to assessing an individual's level of suspicion.[62]

More specifically, data mining technologies, despite their self-learning capabilities, still involve considerable human input when building these algorithms.[63] Data mining involves two types of processes: interpretable processes, which use a limited number of variables in discovering relationships, patterns, or profiles, and thus require human scrutiny; and non-interpretable processes, which involve so many variables (up to thousands) that the way the result came about is nearly impossible to determine.[64] Furthermore, the human analyst "predefine[s] the parameters of the search" that the machine conducts when discovering relationships, patterns, or profiles, meaning that the way the algorithm fundamentally functions is at the analyst's discretion.[65] In either process, therefore, human determination is intrinsically involved in the creation of these models and therefore any human

58. *Id.*
59. Asher & Arthur, *supra* note 55.
60. *See* Barocas, *supra* note 43, at 680-81, 684, 686; Barrett, *supra* note 13, at 339-41.
61. Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 890, 893 (2018).
62. Barocas, *supra* note 43, at 680-81.
63. Tal Z. Zarsky, *Governmental Data Mining and Its Alternatives*, 116 PENN ST. L. REV. 285, 293 (2011).
64. *Id.*
65. *Id.* at 292-93.

bias can also still seep into results of the algorithms, although the extent of their involvement is unknown.

Moreover, even without the active input of human discretion, the results of the machines are dependent on the quality of data used in how these machines are developed.[66] More specifically, the "models," or the collection set of discovered relationships, are learned based on "training data," the data that is used to teach the machine how to behave.[67] The data that is used to train the model is critical in how the machine interacts with data. More specifically, if the training data is biased, then the machine will be biased and perpetuate that bias. Biased data can happen if the data that is used was decided by or involved prejudice, and as such the machine would just be reproducing the same prejudice in their suspicion predictions.[68] For example, if the data used to train the model is based on old police cases for gang-related violence in Los Angeles, the model would then adopt and perpetuate the same kind of discriminatory bias against black and Latino people used by the LAPD that led to excessive incarceration of those populations.[69] Additionally, some of the data actually used in these kinds of predictive policing software "rely on commercial data brokers and data gleaned from social media" which can lead to "acontextual and inaccurate results," as inappropriate motivations, such as profit, or the implicit biases of those brokers and social media accounts are uncertain.[70]

Bias can also infiltrate ASAs if the data used allows the machine to make discriminatory inferences, also known as "collection bias."[71] If the data is based on a particular sample of the population, like if the model is based solely on data regarding cocaine sales from a predominantly black neighborhood, this will lead to an overrepresentation of black people in this data and so the machine would infer a connection between black people and cocaine sales, leading to higher suspicion rates for black people just because they are black.[72] The population sample used in the training data therefore should in theory be representative of the entire population; however, given that these are only samples, this is unlikely.[73]

Ultimately, the implicit bias that has driven racially disparate policing, in terms of both bigoted tipsters and actual police conduct, is just being recycled into these algorithms as the data being used to "teach" these

66. Barocas, *supra* note 43, at 687; Barrett, *supra* note 16, at 340.
67. Barocas, *supra* note 43, at 680-81.
68. *Id.*
69. Donna Murch, *Crack in Los Angeles: Crisis, Militarization, and Black Response to the Late Twentieth-Century War on Drugs*, J. AM. HIST. 162, 164 (2015) ("Punitive campaigns against drugs and gangs in Los Angeles rationalized a new martial infrastructure . . . [A]s in counter-insurgency strategy, the geographic application of force meant that the particular populations were at high risk not only because of their age and race but also because of their location. Indeed, by 1992 city sheriffs listed nearly half of the African American men under age twenty-five in Los Angeles County as gang members.")
70. Barrett, *supra* note 16, at 339.
71. Barocas, *supra* note 43, at 684.
72. *See id.* at 680-81.
73. *Id.* at 686.

algorithms is not reflective of actual crime information.[74] The failure of this technology used by the police to provide accurate results—as well as entirely rid itself of human bias—is not just an indictment of shoddy artificial intelligence but also has serious consequences in terms of an already problematic record of police abuse in the United States.[75] Media and U.S. Department of Justice reports from Baltimore, Cleveland, Ferguson, Chicago, Los Angeles, New Orleans, Albuquerque, and Portland have revealed that police departments in those cities had used excessive force and abuse, even going so far as to suggest that some police officers treated people, specifically minorities, "as animals or subhuman[s]."[76] Despite these serious deficiencies, state police departments and federal law enforcement agencies continue to move forward, as more trials and demonstrations of this technology continue to be put into effect.[77]

Moreover, the uncertainty surrounding the significance of the impact of implicit bias is further exacerbated given the lack of transparency about these algorithms from the developers.[78] Northpointe, a for profit company that created the most-used algorithm for risk assessment recidivism scores has not disclosed any specific calculations that go into the algorithms.[79] Pro Publica, however, used their methodology and found different results in the scores, as well as significant racial disparities.[80] For example, the Florida Supreme Court is deciding a case in which defendant Willie Lynch was arrested for selling approximately fifty dollars' worth of crack cocaine based solely on low quality photos that were run across an algorithmic facial recognition system.[81] "Face Analysis Comparison Examination System" (FACES), the facial recognition system implicated in this case, awarded Lynch "one star"—with the stars correlating to the likelihood of a match—but the analyst who testified to the results admitted that she was neither aware of the maximum number of stars possible nor how the algorithm actually worked, and the other

74. Rieland, *supra* note 49.

75. German Lopez, *Cities Across the Country Have Been Riddled with Accusations of Police Abuse*, VOX (Nov. 14, 2018, 4:12 PM), https://www.vox.com/identities/2016/8/13/17938200/police-shootings-abuse-brutality-justice-department.

76. *Id.*

77. Rigano, *supra* note 39.

78. Julia Angwin et al., *Machine Bias*, PRO PUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/B6T7-7GQJ]

79. *Id.*

80. *Id.*

81. Somil Trivedi & Nathan Freed Wessler, *Florida is Using Facial Recognition to Convict People Without Giving Them a Chance to Challenge the Tech*, AMERICAN CIVIL LIBERTIES UNION (Mar. 12, 2019, 5:15 PM), https://www.aclu.org/blog/privacy-technology/surveillance-technologies/florida-using-facial-recognition-convict-people?fbclid=IwAR37ossykYzI5zgUc1BWyKX-cG-xDLhb1LuGndyyVIyxmL6DJLZEWWxrv64 [https://perma.cc/RYC4-MVQ4].

photos FACES had tagged as possible matches were not disclosed.[82] The lack of knowledge involved in how FACES's system operates, plus the influence it had in Lynch's conviction, namely that the government's case was wholly based on the officer's testimony that they recognized Lynch as the man selling the crack cocaine, demonstrates the need for more scrutiny in how these technologies are used and how they are built.[83]

III. LEGAL MECHANISMS THAT CAN PROTECT AGAINST IMPLICIT BIAS IN POLICE-USED MACHINE LEARNING TECHNOLOGIES

In Section III, this Note will look at two possible approaches to confronting implicit bias in ASAs and other machine learning technologies used by law enforcement. Subsection A will propose that the police should treat ASA results as anonymous tips, which must be deemed sufficiently reliable or be further corroborated in order to be used to lawfully justify a police stop under the reasonable articulable suspicion standard. Furthermore, the algorithms themselves should be held to certain international data protection standards, specifically Articles 13 through 15 of the European Union's data protection legislation, in order to ensure adequate transparency and algorithm accountability such that, at the very least, these software companies are named and shamed. Subsection B will offer that, because ASAs are technically products that are purchased by police departments, software developers should be held accountable for any defects, namely disproportionate levels of racial disparities, under tort product liability standards.

A. Holding the Government Accountable: Fourth Amendment Checks on Implicit Bias

Unlike bigoted tipsters, like #BBQBecky and #PermitPatty, who face social pushback and are expected to face the appropriate level of scrutiny under the Fourth Amendment in order to be considered reliable, ASAs and other machine learning algorithms used by the police encounter similar issues of implicit bias but do not receive much pushback, either legally or socially. Therefore, ASAs and other machine learning technologies used by the police should, at the very least, be held to the same standard as calls from BBQ Beckys—anonymous tips that require the evaluation of the tip's reliability or some other corroborating information.

Jurisprudence for stops based on the Fourth Amendment requires that police officers must have a reasonable, articulable suspicion (RAS) that the individual is involved in an imminent or pending criminal activity or that the individual has just committed a felony in order to lawfully stop them without

82. *Id.*; Aaron Mak, *Facing Facts*, SLATE (Jan. 25, 2019, 12:49 PM), https://slate.com/technology/2019/01/facial-recognition-arrest-transparency-willie-allen-lynch.html [https://perma.cc/XCJ4-6V7W].

83. Trivedi & Wessler, *supra* note 81; Mak, *supra* note 82.

their consent.[84] While the Supreme Court has yet to specifically define what RAS means, the standard required to briefly detain a person in order to question or resolve the officer's suspicion is lower than what is required to establish probable cause.[85] In terms of implicit bias, the Court has deemed any pretextual motivation, including racial animus, irrelevant under the Fourth Amendment as long as the officer is able to provide either probable cause or RAS.[86] Even under the Fourteenth Amendment's Equal Protection Clause, plaintiffs must provide evidence that the officer intentionally discriminated, as it is "not sufficient to show disproportionate impact."[87] An officer's intentions, therefore, are largely immaterial when doing a stop under RAS.

When police work on tips from 911 callers, such as #BBQBecky or #PermitPatty, they similarly have to ensure that these tips are sufficiently reliable in order to establish either probable cause or RAS.[88] In evaluating the reliability of tips, courts evaluate the quantity and quality of the information along with the degree of reliability of the tip.[89] More specifically, the Supreme Court found in *Illinois v. Gates* that courts should evaluate the "veracity, reliability and basis of knowledge" of the tipster under a totality of the circumstances analysis, specifically determining whether the tipster is reliable and how the tipster came upon the information.[90] Under a totality of the circumstances analysis, should one element of this reliability determination be considered weak—i.e. the basis of knowledge is based on hearsay and not personal knowledge—the required level of suspicion is not automatically defeated as long as the veracity of the tipster is established.[91] In establishing the veracity of the tipster, the "absence of an apparent motive to falsify an independent police corroboration of the details provided by the informant" can be considered sufficient.[92]

Tips from anonymous tipsters, as opposed to known tipsters, therefore, generally lack an "indicia of reliability" at the outset to justify a stop under *Terry v. Ohio*, but can become sufficient if the police have other compelling

---

84. *See generally* Terry v. Ohio, 392 U.S. 1 (1968).

85. *Id.* at 27 (finding that there must be something more than an "inchoate and unparticularized suspicion or 'hunch'" and "some minimal level of objective justification.")

86. Whren v. United States, 517 U.S. 806, 814 (1996) (holding that pretextual motivations are irrelevant to a Fourth Amendment search or seizure as long as it meets probable cause standards); Ashcroft v. Al-Kidd, 563 U.S. 731, 740 (2011) (citing Saucier v. Katz, 533 U.S. 194, 201-02, (2001)) (holding that pretextual motivations are irrelevant for lower suspicion stops as long as RAS is met).

87. Wayne C. Beyer, *Police Misconduct: Claims and Defenses under the Fourteenth Amendment Due Process and Equal Protection Clauses*, 30 URBAN LAWYER 65, 113-14 (1998).

88. Alabama v. White, 496 U.S. 325, 330 (1990).

89. *Id.* (noting that the same totality of the circumstances test applied in *Illinois v. Gates* to determine probable cause should be used for establishing RAS, albeit it at a lower standard than probable cause).

90. Illinois v. Gates, 462 U.S. 213, 230 (1983) (holding that the "rigid" two-prong test under *Aguilar* and *Spinelli* in assessing the reliability of tips is no longer the law); United States v. Angulo-Lopez, 791 F.2d 1394, 1396 (9th Cir. 1986).

91. *Angulo-Lopez*, 791 F.2d at 1396.

92. *Id.* at 1397.

corroborative information and can find no underlying motivation.[93] In addition, anonymous tips only offering "innocent" details, such as the particulars about the individual's clothing or appearance, are only helpful in identifying the individual, but do not speak to either the tipster's "knowledge of concealed criminal activity," or to whether there is an illegality that can give rise to RAS.[94] Tipsters like #BBQBecky and #PermitPatty, therefore, are only so helpful in identification of individuals since they only describe the suspicious person as opposed to any legitimate unlawful activity and, as such, may reveal their implicit bias and improper, racially based motivation for calling the police. Thus, in order to establish RAS and lawfully stop the individuals in question, the police must have some other corroborating information about an imminent or impending illegality.

Similarly, ASAs are used primarily to identify individuals who meet certain factors of "suspicion" and should not be considered reliable unless there is some other corroborating information.[95] Therefore, each time an ASA alerts the police once an individual reaches the pre-set level of confidence, the police officer should assess the "veracity, reliability, and basis of knowledge" of the ASA itself.[96] The reliability of the information provided by the ASA, however, involves just as much, if not more, implicit bias than #BBQBecky and #PermitPatty and thus its reliability should be considered weak by default. Furthermore, the ASAs are created by third party, for-profit companies that use unverified employees who may have implicit or explicit biases—financial motivations, a lack of concern for ensuring accuracy, or even simply having different hiring standards than officers—that are then learned by the machine they've created.[97]

As a result, an ASA's reliability is fundamentally in question and can often lead to inaccurate, discriminatory results that could amount to a Fourth Amendment violation. Even so, remedies for violations of the Fourth Amendment are limited—an individual who is charged with a crime is entitled to invoke the "exclusionary rule" in order to suppress any illegally obtained evidence.[98] However, several exceptions to the rule, such as the Independent Source Doctrine,[99] Good Faith (or "*Leon*") Doctrine,[100] and the

93. Terry v. Ohio, 392 U.S. 1 (1968); Adams v. Williams, 407 U.S. 143, 147 (1972) (holding that a tip can carry sufficient "indicia of reliability" to establish RAS but not sufficient to establish probable cause like when information comes from a known informant).

94. Florida v. J.L., 629 U.S. 266, 271-72 (2000).

95. *Adams*, *supra* note 84, at 147.

96. *Illinois v. Gates*, 462 U.S. at 230.

97. Angwin, *supra* note 78.

98. *See, e.g.*, Weeks v. United States, 232 U.S. 383, 397-98 (1914) (holding exclusionary rule applies in federal cases); Mapp v. Ohio, 367 U.S. 643, 655-57 (1961) (extending the "exclusionary rule" to states, implying constitutional underpinnings to the rule).

99. *See, e.g.*, Segura v. United States, 468 U.S. 796, 805 (1984); Murray v. United States, 487 U.S. 533, 537 (1988).

100. *See, e.g.*, United States v. Leon, 468 U.S. 897, 923-24 (1984); Davis v. United States, 564 U.S. 229, 246-47 (2011) (holding that officers' illegal conduct had to be deliberate, reckless, or grossly negligent).

Herring doctrine,[101] make the exclusionary rule nearly ineffective. In addition, those who are not defendants in a criminal action and cases where illegally seized evidence is being introduced are forced to seek alternate remedies that are similarly unsatisfactory.[102] Civil suits under 42 U.S.C. § 1983 and *Bivens* actions both require funding that may exclude indigent defendants and are also bound by officers' qualified immunity, which protects discretionary actions by officers as long as "conduct does not violate clearly established statutory or constitutional rights of which a reasonable person would have known."[103] Even the Supreme Court has conceded that "exclusion presents the only effective remedy to deter wrongful police conduct," suggesting that some forms of police misconduct within the ambit of the Fourth Amendment cannot be remedied.[104]

Despite the limitations on rectifying Fourth Amendment violations, police officers should still be vigilant in the face of growing reliance on technology. More specifically, officers should be wary of ASAs and their built in biases, as they would be with a #BBQBecky or #PermitPatty and their more blatant racial prejudices, taking care to scrutinize any results produced by ASAs before automatically dispatching an officer. Corroborative information should be acquired before the police interact with a targeted individual in order to ensure that actual illegal conduct is in question and that the individual has not been targeted for racist reasons.

1. Holding the Government Accountable: Data Protection Legislation

Another way of ensuring that ASAs are reliable tools that avert, rather than perpetuate, implicit bias would be to ensure that the data used within these algorithms and used in training models adheres to certain protections. More specifically, there should be certain safeguards and limits on using personal data, such as factors associated with identification (including addresses and physical features like those used on driver's licenses), when used by third party businesses and by the government. Advocacy organizations like the Electronic Privacy Information Center (EPIC) have been a part of litigation against the government and technology companies like Facebook seeking to obtain records of the type of data being used and how it's being used, and to ensure checks on algorithmic transparency are put

101. Herring v. United States, 555 U.S. 135, 144 (2009) (holding that application of the exclusionary rule should be determined on a case by case basis in order to promote deterrence sufficient to be "worth the price paid by the justice system" and as such should not be automatically applied).

102. *See generally*, Brent E. Newton, *The Supreme Court's Fourth Amendment Scorecard*, 13 STAN. J. C.R. & C.L. 1, 13-14 (2017).

103. Wilkerson v. Goodwin, 774 F.3d 845, 851 (5th Cir. 2014).

104. *See* Elkins v. United States, 364 U.S. 206, 220 (1916).

in place.[105] EPIC has also proposed legislation through its Public Voice coalition titled the "Universal Guidelines for Artificial Intelligence," which would impose obligations on businesses to have a final determination made by a human when using algorithms, and obligations on institutions using artificial intelligence generally to ensure that there is no unfair bias or impermissible discriminatory decision making taking place.[106]

Furthermore, in 2016, the European Union passed the General Data Protection Regulation (GDPR), which, in Chapter 3, gave data subjects—any EU citizen whose personal data is being used—the right to ask what of their personal information is being used and how it is being used.[107] Article 35 of the GDPR also requires a company or organization to employ a "data protection officer" whenever personal data, such as ethnicity, religious beliefs, or genetic data, is used in order to ensure that the company or organization using this data is using it in compliance with GDPR standards.[108] Finally, Article 79 sets a penalty for non-compliance, which can rise up to 4% of a company's global annual revenue based on the violation.[109] Although this Note does not suggest adopting all of the GDPR, the move towards regulation ensuring algorithmic transparency, specifically in allowing citizens to at least inquire about their data, should motivate US legislators. This may already be in motion, as California recently passed their own data privacy law.[110]

The United States on both federal and state levels should take California's example and push further, passing legislation that specifically investigates the algorithms employed in the criminal justice system with the goal of uncovering the breadth and severity of implicit bias within these algorithms and its discriminatory effects in the real world. The United States should impose policies similar to the data access protection provisions within the GDPR specifically allowing data subjects, or in this case those stopped by police officers for seemingly no legitimately lawful reason, access to inspect what personal information is being used by the software developers generating ASAs and how their algorithms then use that data.[111] For the US

---

105. Brief for EPIC as Amici Curiae Supporting Appellee, *Patel v. Facebook, Inc.*, 932 F.3d 1264 (9th Cir. 2018) (No. 18-15982) (noting that Facebook's collection of biometric data, namely the facial recognition feature used for photos, is a prohibition of the Illinois Biometric Information Privacy Act which asserts "certain obligations on private entities that collect or possess biometric identifiers."); Brief for EPIC as Amici Curiae Supporting Appellee, *United States v. Miller*, [No. 16-47-DLB-CJS, 2017 WL 2705963 (E.D. KY. 2017) (No. 18-5578) (arguing that the search of Miller's emails, which were only searched upon Google becoming aware that there were flagged images of apparent child pornography being sent through the emails, was an invasion of privacy and a violation of the Fourth Amendment).

106. *Universal Guidelines for Artificial Intelligence*, THE PUBLIC VOICE (Oct. 23, 2018), https://thepublicvoice.org/ai-universal-guidelines/.

107. *See generally* 2016 O.J. (L 119) 39-47.

108. *Id.* at 53; Juliana De Groot, *What is the General Data Protection Regulation? Understanding and Complying with GDPR Requirements in 2019*, DIGITAL GUARDIAN (Jan. 3, 2019).

109. De Groot, *supra* note 108.

110. Dipayan Ghosh, *What You Need to Know About California's New Data Privacy Law*, HARV. BUS. REV. (July 11, 2018), https://hbr.org/2018/07/what-you-need-to-know-about-californias-new-data-privacy-law.

111. General Data Protection Regulation, 2016 O.J. (L 119) 1-88.

government to even acknowledge that these algorithms may be inherently biased, especially when used to detect suspicion levels or identify certain types of people, may be helpful in combatting these algorithms' discriminatory outputs. Moreover, tasking the government with monitoring and regulating how these algorithms operate (specifically in regards to the misuse of personal data that leads to discriminatory results) makes the government responsible for the prevention of automated implicit bias.

B. Holding Developers Accountable: Product Liability

Ultimately, ASAs and other machine learning technologies used by the police are products sold to them by specialized technology and software development companies.[112] Third-party companies that create ASAs and other machine learning technologies employed by the police should be held accountable by implementing a financial penalty in cases of wrongful use of data, such as biased data collection or use of biased historical records, creating a better incentive to be more vigilant in ensuring that algorithms are not incorporating biases in their initial stages and perpetuating inaccurate racial disparities, or at least in attempting to correct these biases within the models. Therefore, one way of ensuring accountability and decreasing the risk of incidents of implicit bias in action would be to hold the developers legally accountable for recklessly or intentionally created ASAs and the like that effectuate disproportionately discriminatory results. An immediate issue with this solution, however, is the lack of consensus on whether artificial intelligence generally should be regarded as a product or as a service, which affects the theory of tort liability under which AI should be evaluated.

1. AI as a Product: Strict Liability

Contemplating artificial intelligence as a product in tort litigation could prompt strict liability if the product, or computer-based technology, is found to be defective.[113] Strict liability notably does not consider fault, but rather is based purely on causation. Therefore, any safeguards taken by the maker or developer in order to prevent the defect would be irrelevant.[114]

Strict liability is found if a plaintiff can prove that the product was "defective and unreasonably dangerous, that the defect existed when it left the

112. Karen Hao, *Police Across the US are Training Crime-Predicting AIs on Falsified Data*, MIT TECH. REV. (Feb. 13, 2019), https://www.technologyreview.com/s/612957/predictive-policing-algorithms-ai-crime-dirty-data/ (mentioning Palantir and PredPol, third party companies developing predictive policing software).

113. Ryan Abbott, *The Reasonable Computer: Disrupting the Paradigm of Tort Liability*, 86 GEO. WASH. L. REV. 1, 15 (2018).

114. Rylands v. Fletcher, L.R. 3 H.L. 300 (1868) (holding that the owner who brought a reservoir onto his land that was considered "abnormally dangerous" is *prima facie* answerable to any damage arising from it); Greenman v. Yuba Power Products, 377 P.2d 897 (Cal. 1963) (en banc) (holding that manufacturers of defective products are strictly liable).

hands of [the] defendant and the defect caused the harm."[115] A defect, as defined by the Restatement (Second) of Torts, requires a showing that the product was unexpectedly dangerous to the ordinary customer, also known as the "consumer expectation test."[116] In *Camacho v. Honda Motor Co., Ltd.*, the Supreme Court of Colorado evaluated whether the product is unreasonably dangerous under a totality of the circumstances test to determine whether the product should be held to a strict liability standard.[117] Factors for a court to evaluate under this test include: the utility of the product to user and public; the safety of the product (which would inquire into the likelihood that the product will cause injury or probable seriousness of harm); the availability of a substitute product meeting the same need and that is not as unsafe; the user's ability to avoid harm with due care, and other factors.[118] Under the Restatement (Third) of Torts, a strong case for applying strict liability for an abnormally dangerous activity is essentially based on whether there was knowledge of significant risk and whether that risk was still disregarded.[119] The factors to evaluate under this theory include whether: "(1) the activity creates a foreseeable risk of physical harm; (2) the risk is a 'highly significant' risk; (3) the risk remains 'even when reasonable care is exercised by all actors;' and (4) 'the activity is a matter of not common usage.'"[120]

Applying strict liability to ASAs that result in disproportionately racist identifications would be a direct way to combat implicit bias in these algorithms. ASAs and other machine learning technologies that are used to identify individuals inherently carry significant risk and thus require constant vigilance from the outset, instead of requiring some error to be seen as dangerous.[121] If a developer is aware that they are strictly liable for any defect in their algorithm's execution, perhaps ASA developers would be more incentivized to have more oversight and controls in place that ensure implicit bias is detected and corrected before it gets integrated into models and gets lost.[122] However, traditional tort product litigation under strict liability is most accessible to users of a product who suffer personal, physical injury, as opposed to a bystander or third party suffering non-physical harms such as

115. Alvin S. Weinstein et al., *Product Liability: An Interaction of Law and Technology*, 12 DUQ. L. REV. 425, 428-29 (1974); DAVID G. OWEN, PRODUCTS LIABILITY LAW 1 (3d ed. 2014); RESTATEMENT (SECOND) OF TORTS § 402A (1965).

116. Roger Traynor, *The Ways and Meanings of Defective Products and Strict Liability*, 32 TENN. L. REV. 363, 366 (1965); RESTATEMENT (SECOND) OF TORTS § 402A (1965).

117. Camacho v. Honda Motor Co., Ltd., 741 P.2d 1240, 1244 (Colo. 1987).

118. *Id.* at 1247.

119. Elizabeth Fuzaylova, *War Torts, Autonomous Weapon Systems, and Liability: Why a Limited Strict Liability Regime Should be Implemented*, 40 CARDOZO L. REV. 1327, 1360 (2019).

120. *Id.*; RESTATEMENT (THIRD) OF TORTS § 20 (AM. LAW. INST. & UNIF. LAW COMM'N 2010).

121. *See generally* Rich, *supra* note 51.

122. Abbott, *supra* note 113, at 22 ("[S]trict liability creates a stronger incentive for manufacturers to make safer products, and manufacturers may be better positioned than consumers to insure against loss.")

economic or emotional pain.[123] The "harms" that result from racially prejudiced algorithms do not necessarily result in physical harms although the effect—unlawful police stops—is still damaging. Furthermore, the "harms" of the defective product are not inflicted on the user, the police officer, but on third parties who are identified by the defective product and stopped by the police. Therefore, meeting even the basic requirements of strict liability, particularly that the defect caused this harm, may be difficult under traditional tort liability theory.[124]

However, should courts move in a direction that recognizes non-physical harms brought on by reliance on machine learning, ASAs and other machine learning technologies used by the police that adopt and employ implicit bias may meet the necessary factors for determining a defect. First, the utility of machine learning in law enforcement is fairly high—all computer-based technologies that involve automation, such as an ASA, should in theory be safer in executing a design than a human since it eliminates the risk of human error.[125] Former General Motors ("GM") vice chairman, Bob Lutz, was quoted predicting that GM's first autonomous car will have a significantly lower accident rate than cars driven by humans.[126] Specifically, within the law enforcement context, the fairly low ratio of officers to citizens, as well as the enormous prison population, suggests that additional, reliable assistance, such as suspicion detecting cameras and facial recognition, would be beneficial.[127] Despite the considerable potential of these kinds of machine learning technologies, the risk of harm inherent in their use, specifically in regards to implicit bias, outweighs the possible benefits of ASAs and thus strengthens the argument that ASAs that incorporate racial bias are defective. More specifically, the safety of an ASA is questionable, as there is clearly a high likelihood of error in terms of racial bias that can lead to unlawful stops or other more intensive police interactions.[128] Although the concern with unquestioned reliance on ASAs is that an unlawful police stop could lead to violence and even death, this would still require courts to broaden their view of "harm" to include non-physical injuries to ensure that all instances of police misconduct initiated by ASAs are punishable, not just those that lead to violence. Second, the availability of

123. Cathy Bellehumeur, *Recovery for Economic Loss Under a Products Liability Theory: From the Beginning Through the Current Trend*, 70 MARQ. L. REV. 320, 321-22 (1987) ("Most courts do not allow tort recovery for purely economic loss in the absence of any personal injury or property damage. However, the method for categorizing a claim is so varied that a claim not recoverable in tort because it constitutes an economic loss in Idaho may be recoverable in tort in Illinois where it is classified as a claim for property damage.").

124. *See* Alvin S. Weinstein et al., *supra* note 102, at 428-29.

125. Abbott, *supra* note 113, at 18-19.

126. *Id.*

127. *Police Employee Data*, Federal Bureau of Investigation: Criminal Justice Information Services Division (2011), https://ucr.fbi.gov/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/policeemployees_main_final.pdf (in 2011, there were 3.4 full-time law enforcement officers for every 1,000 inhabitants).

128. *See* Renata M. O'Donnell*, Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause*, 94 N.Y.U. L. REV. 544, 553 (2019).

a substitute product without implicit bias may not currently be possible as the algorithms typically rely on police records, which inherently involve bias.[129] However, police officers are beginning to undergo implicit bias trainings to defeat this bias and so could technically be considered an appropriate substitute to an ASA.[130] Finally, even with due care, an officer may not be able to avoid the harms, or rather the perpetuation of harms, that come with relying solely on suspicion algorithms since the implicit bias may be so embedded that it could be difficult to detect where the bias came from and how it influenced the algorithm's results.[131]

Despite courts thus far failing to apply strict liability to AI on technical grounds,[132] the inherent danger of AI used in the context of police identifications should be sufficient to allow courts to view AI software as unreasonably dangerous even though the harm may not necessarily be obvious or physical. Strict liability was born out of a need to find liability when product failures became increasingly difficult to prove under traditional theories of negligence[133] and as such should be the starting standard for how courts evaluate machine learning technology. On a policy level, imposing strict liability would not only on its face imply that there is significant risk in using this technology—which would in application force police officers to use ASAs only with caution—but would also force developers to reduce the risk from the beginning in order to avoid being held liable. Furthermore, companies, as opposed to users, are much better equipped to find and correct defects in products, as seen with companies, such as Volvo and Google, recognizing their ability to correct and prevent harm, announcing that they will accept full responsibility if their self-driving products cause a collision despite no legal compulsion to do so.[134] Courts should therefore create a similar legal obligation on all software developers selling AI products, especially those who provide AI to government agencies, in order to ensure that these algorithms are being generated in good faith and to acknowledge their particular predisposition to biased input. Applying strict liability would also ensure that those who were unlawfully stopped by the police or were the victim of other constitutional violations have the opportunity for redress and thus the traditional strict liability paradigm should expand to include non-physical, third party harms.

---

129. Barocas, *supra* note 43, at 680-81, 684; *see also* William Isaac & Andi Dixon, *Why Big Data Analysis of Police Activity is Inherently Biased*, PBS (May 10, 2017), https://www.pbs.org/newshour/nation/column-big-data-analysis-police-activity-inherently-biased.

130. James, *supra* note 29.

131. Barocas, *supra* note 43, at 673-74.

132. Chatlos Systems Inc. v. National Cash Register Corp., 479 F. Supp. 738, 740-41 n. 1 (D.N.J. 1979), *rev'd on other grounds*, 635 F.2d 1081 (3d Cir. 1980), *aff'd after remand*, 670 F.2d 1304 (3d Cir. 1981) (holding that strict liability could not be applied to a service).

133. David C. Vladeck, *Machines without Principals: Liability Rules and Artificial Intelligence*, 89 WASH. L. REV. 117, 132 (2014).

134. Ben Taylor, *Who's liable for decisions AI and robotics make?*, BETANEWS (Mar. 21, 2017), https://betanews.com/2017/03/21/artificial-intelligence-robotics-liability/.

Ultimately, the use of the algorithms is more often than not going to result in racially disparate and discriminatory results,[135] so despite difficulties in applying strict liability under its traditional tort framework, on a larger policy level, developers should be held liable for the harms they can cause or for failure to properly identify and attempt to correct the implicit bias.

2. AI as a Service: Negligence

Under the Uniform Commercial Code (UCC), software that is specifically designed for a customer, as opposed to mass-produced and for the public, should be considered a service as opposed to a product.[136] As a service, harms caused by AI could then be evaluated under a traditional negligence standard.[137] The plaintiff would have to prove that the defendant, in this case the software developer, had a duty of care, that they breached that duty by failing to conform to a reasonable standard of behavior, and that that breach caused the injury to the plaintiff—essentially determining who is at fault.[138] Duty would, in theory, not be at issue here since there is privity between the software developer and the police officer. Further, many jurisdictions allow for foreseeable bystanders to recover, so, here, individuals targeted by ASAs could recover as well.[139] However, courts have refused to establish professional standards for technology developers as there is no licensing requirement, and as such "computer malpractice" has found limited acceptance in today's jurisprudence.[140] Furthermore, under a negligence paradigm, there is a broader range of harms that could be litigated, including a "knowledge base [that] is incomplete or inadequate; incorrect or inadequate warnings and documentation; [or] . . . the user . . . supplying faulty input or selecting the incorrect program for the task."[141] However, the burden of proof to demonstrate a breach of that nature rests on the plaintiff, a task which is already difficult for software developers given the intricacies of machine learning technology, and likely even more difficult for the injured individual

135. Barocas, *supra* note 43, at 673-74.

136. Rottner v. AVG Technologies USA, Inc., 934 F. Supp. 2d 222, 230 (D. Mass. 2013); Motorola Mobility, Inc. v. Myriad France SAS, 850 F. Supp. 2d 878 (N.D. Ill. 2012) (holding that computer code is considered a service and liability for defective software should be evaluated under a breach of warranty under the UCC as opposed to product liability).

137. J.K.C. Kingston, *Artificial Intelligence and Legal Liability*, *in Research and Development in Intelligent Systems XXXIII: Incorporating Applications and Innovations in Intelligent Systems XXIV* 269-279 (Max Bramer & Miltos Petridis eds., 2016).

138. *Id.*

139. MacPherson v. Buick Motor Co., 217 N.Y. 382, 390 (1916) (holding that privity of contract is no longer necessary to establish duty of care in negligence actions and thus, if it is reasonably foreseeable that a product is dangerous, duty of care goes beyond buyer and seller).

140. Marguerite E. Gerstner, *Liability Issues with Artificial Intelligence Software*, 33 SANTA CLARA L. REV. 239, 258-59 (1993).

141. *Id.* at 248.

who has no access to how the algorithm is developed and what information is being used.[142]

Given how complex and impenetrable the neural networks that drive ASAs and other machine learning technologies are, determining fault is nearly impossible to discern within the AI context.[143] Even prior to analyzing the breach inquiry of a negligence claim, the plaintiff would have to demonstrate that the reasonable standard of behavior of a computer is safer than a human and thus the implicit bias that leads to racialized results would be a failure to conform.[144] Furthermore, as is the case with strict liability, causation is hard to prove since ASAs only suggest certain actions, requiring another agent (the officer) to take the action that causes the harm of unlawful police interactions.[145] However, given that there have been cases where officers have solely relied on these algorithmic results to make their arrests suggests that these algorithms are directly causing potentially unlawful arrests or at least lending themselves to potential Fourth Amendment violations.[146]

At the very least, software developers can be compelled to satisfy a duty to warn about the risk that results of ASAs will contain inherent racial bias, so officers can be aware of this risk when relying on the ASA. However, a warning may not be sufficient to offset any police action based on the results of the ASA. Although negligence is what is traditionally used with services it does not do enough to compel software developers to be vigilant and active in identifying sources of bias in generating algorithms. Furthermore, given that these technologies continue to learn and make decisions without human input, and given that companies are already taking full responsibility for their technologies not functioning,[147] courts adopting a strict liability approach would offer a more straightforward, effective means of assigning liability to software developers. Ultimately, placing liability on the developer who has control over the creation of these algorithms sheds light on how implicit bias can permeate automated technologies and help prepare for a future where these machines operate without any human input at all.

IV. CONCLUSION

Implicit bias is a significant issue in the criminal justice system that requires constant vigilance and checks in order to be effectively combatted. The rise of ASAs and other machine learning technologies used by law enforcement should prevent implicit bias rather than perpetuate it, and so both officers and the software development companies that profit off these algorithms should be held liable for being complicit in the perpetuation of

142. Bathaee, *supra* note 62, at 892-93 (noting the near impossibility of discerning how a machine learning algorithm functions given its advanced learning abilities that even humans may not be able to comprehend).

143. Curtis E.A. Karnow, *Liability for Distributed Artificial Intelligences*, 11 BERKELEY TECH. L.J. 147, 192 (1996).

144. Abbott, *supra* note 125, at 22.

145. *Id.* at 22-26.

146. *See* Trivedi & Wessler, *supra* note 81.

147. Taylor, *supra* note 134.

implicit bias in new technology. It should be standard practice for police departments to treat the results from ASAs as anonymous tips in order to meet the requisite RAS to remain compliant with Fourth Amendment protections against search and seizure. Corroborative information ensuring that the ASA identified an illegality as opposed to a suspicious person based on race is necessary to combat the implicit biases that are entangled within these algorithms similar, to how BBQ Beckys are treated. In addition, software developers need to be transparent about the type of information going into these algorithms and therefore should be compelled to disclose to individuals how and what data is being used under a similar access scheme to the GDPR's Chapter 3.[148] Finally, software developers should be held legally responsible for their software through tort strict liability in order to incentivize developers, the only people who have the opportunity and skill to eliminate these biases, to keep vigilant for implicit biases as well as allow targeted individuals the opportunity to recover from the harms they suffer from being unlawfully targeted by the police. Furthermore, while this Note does not address the privacy issues of using this kind of personal data by third parties, as well as the lack of individualized suspicion involved in the employment of ASAs under a Fourth Amendment analysis, these issues add to the need for machine learning to have more transparency and accountability in their models and results.

As algorithms become increasingly universal but with no additional requirements for adequate transparency or accountability, the racial bias of #BBQBecky that has proven to be pervasive in and perpetuated by machine learning algorithms used by the police continues to become more cemented in the criminal justice system and in society. More jurisdictions should mirror California in being attentive to these technologies, but more specifically in how they are relied on by law enforcement, in order to promote equity in our criminal justice system and to help prevent further violence perpetrated by the police against black Americans.

148. *See generally* 2016 O.J. (L 119) 39-47.