VOLUME 3 November 2022 ISSUE 2

JOURNAL ON EMERGING TECHNOLOGIES

ARTICLES

ETHICAL AI IN AMERICAN POLICING

Elizabeth E. Joh

ETHICAL AI IN AMERICAN POLICING

*Elizabeth E. Joh**

INTRODUCTION

We know there are problems in the use of artificial intelligence in policing, but we don't quite know what to do about them.[1] Artificial intelligence (AI) systems[2] are becoming conventional and widespread in routine policing. License plate reader systems routinely scan thousands of plates per minute.[3] At least 117 million Americans are included in databases where facial recognition searched are conducted.[4] Predictive algorithms try to forecast future places or persons warranting law enforcement attention.[5] Autonomous drones can follow a suspect or record activity with the push of a button.[6] Increasingly the issue is not whether, but under what circumstances, these tools will be used.

---

*Professor of Law, U.C. Davis School of Law. Thanks to the editorial staff of the Notre Dame Journal on Emerging Technologies for their editorial work, and to the inter-journal collaboration at Notre Dame Law School for organizing the Race & the Law: Interdisciplinary Perspectives symposium.

[1] This isn't just an American problem. The director of the UK Police Foundation stated in January 2022 that "national guidance on ethical considerations [for emerging technologies] would be especially welcome." *See* GLORIA GONZÁLEZ FUSTER, EUROPEAN PARLIAMENT POL'Y DEP'T FOR CITIZEN'S RTS. & CONST. AFFS., ARTIFICIAL INTELLIGENCE AND LAW ENFORCEMENT: IMPACT ON FUNDAMENTAL RIGHTS (2020) [hereinafter IMPACT ON FUNDAMENTAL RIGHTS] ("The magnitude and seriousness of challenges triggered by AI in the field of law enforcement and criminal justice . . . do not appear to be conveniently addressed by ongoing reflections."); Claudia Glover, *Policing Minister Rejects Need for Ethical Guidance on Emerging Tech*, TECH MONITOR (Jan. 13, 2022), https://techmonitor.ai/policy/regulating-use-of-technology-in-uk-police.

[2] By using the terns "AI applications" or "AI systems," I refer to the application of algorithms and substantial amounts of computing power to enormous amounts of digitized data.

[3] *See, e.g.*, Ángel Díaz & Rachel Levinson-Waldman, *Automatic License Plate Readers: Legal Status and Policy Recommendations for Law Enforcement Use*, BRENNAN CTR. (Sept. 10, 2020), https://www.brennancenter.org/our-work/research-reports/automatic-license-plate-readers-legal-status-and-policy-recommendations (noting "93 percent of police departments in cities with populations of 1 million or more use their own ALPR systems, some of which can scan nearly 2,000 license plates per minute").

[4] Clare Garvey et al., *The Perpetual Line-Up: Unregulated Police Face Recognition in America*, GEORGETOWN L. CTR. ON PRIV. AND TECH. (Oct. 18, 2016) [hereinafter *Perpetual Line-Up*], https://www.perpetuallineup.org/ (noting that at least 26 states allow the police to run face recognition searches against driver's license and ID photos).

[5] *See infra* Part I.

[6] *See infra* Part I.

With artificial intelligence, the police can perform their traditional functions not just on a faster and larger scale, but in novel ways that have prompted strong criticism. Some of these issues are familiar to a legal audience. If the police can track everywhere you've been in public, what does that mean for the usual lack of constitutional protections in public spaces? If the police can easily identify every face in a public protest, how does that dampen free speech rights? Other voices in this backlash have arisen out of what has been called the algorithmic accountability movement: scholars and activists who have focused on the harms posed by the particulars of the technologies themselves.[7] For instance, the now quite well-documented issue of racial and gender bias in many facial recognition technology programs means that the costs of mistaken matches are borne disproportionately by people of color and women.[8] At the same time, law enforcement officials have embraced these technologies as promising innovations. Automation both in and around policing is growing, with few signs of slowing down.

One can also find many reports and white papers today offering principles for the responsible use of AI systems by governments, civil society organizations, and the private sector. Increasingly common too are calls for the fair use of artificial intelligence across fields like housing, employment, consumer credit, and criminal justice. This comes at a time when automated decision-making might determine whether you'll be hired,[9] whether you'll be fired,[10] whether you'll receive one medical

---

[7] We can also include here the development of the field of Fairness, Accountability, and Transparency in Machine Learning. See, e.g., *Fairness, Accountability, and Transparency in Machine Learning*, FATML, http://fatml.org (last visited Feb. 26, 2022).

[8] Researcher Joy Buolamwini was among the first to identify the issue of bias. Steve Lohr, *Facial Recognition is Accurate, if You're a White Guy*, N.Y. TIMES (Feb. 9, 2018), https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html (citing Buolamwini's work finding up to 35% error rate for darker skinned women compared to 1 percent error rate for white men). The National Institute of Standards and Technology similarly found in 2019 that the facial recognition programs it studied mistakenly identified people of color far more often than white people. *See NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software*, NIST (Dec. 19, 2019), https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software (evaluating 189 software algorithms and finding that "for one-to-one matching, the team saw higher rates of false positives for Asian and African-American faces relative to images of Caucasians.").

[9] Rebecca Heilweil, *Artificial Intelligence Will Help Determine if You Get Your Next Job*, VOX (Dec. 12, 2019), https://www.vox.com/recode/2019/12/12/20993665/artificial-intelligence-ai-job-screen ("recruiters are increasingly using AI to make the first round of cuts and to determine whether a job posting is even advertised to you.").

[10] Spencer Soper, *Fired by Bot at Amazon: 'It's You Against the Machine,'* BLOOMBERG (June 28, 2021), https://www.bloomberg.com/news/features/2021-06-28/fired-by-

treatment over another, or whether you'll be granted bail. In 2021, Congress established a National AI Advisory Committee, tasked with providing recommendations about the use of AI and its impact on society.[11] The White House Office of Science and Technology Policy plans to publish an Algorithmic "Bill of Rights."[12] The European Union is preparing to adopt a comprehensive regulatory framework for the use of AI in 2022.[13]

Yet, largely missing from the current debate in the United States is a shared framework for thinking about the ethical and responsible use of AI that is specific to policing.[14] Leading an average-sized law enforcement agency in the United States in the 2020s means responding to very different pressures: to reduce crime, to address bias and

---

bot-amazon-turns-to-machine-managers-and-workers-are-losing-out ("Increasingly, the company is ceding its human-resources operation to machines as well, using software not only to manage workers in its warehouses but to oversee contract drivers, independent delivery companies and even the performance of its office workers.").

[11] The Committee is one of several governance bodies created by the National Artificial Intelligence Initiative Act of 2020. *See National Artificial Intelligence Advisory Committee (NAIAC)*, NIST (Oct. 27, 2021), https://www.nist.gov/artificial-intelligence/national-artificial-intelligence-advisory-committee-naiac.

[12] Eric Lander & Alondra Nelson, *Americans Need a Bill of Rights for an AI-Powered World*, WIRED (Oct. 8, 2021), https://www.wired.com/story/opinion-bill-of-rights-artificial-intelligence/ ("In the coming months, the White House Office of Science and Technology Policy (which we lead) will be developing such a bill of rights, working with partners and experts across the federal government, in academia, civil society, the private sector, and communities all over the country.").

[13] *2021 Artificial Intelligence and Automated Systems Annual Legal Review*, GIBSON DUNN (Jan. 20, 2022), https://www.gibsondunn.com/wp-content/uploads/2022/01/2021-artificial-intelligence-and-automated-systems-annual-legal-review.pdf ("With the new Artificial Intelligence Act, which is expected to be finalized in 2022, it is likely that high-risk AI systems will be explicitly and comprehensively
regulated in the EU."). The proposed EU regulations focus on "harmonised rules for the development, placement on the market and use of AI system in the Union following a proportionate risk-based approach." *See Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, at 3, COM (2021) 206 final (Apr. 21, 2021).

[14] The same is true elsewhere, as observed by the U.K. government's Centre for Data Ethics & Innovation. *See* CTR. FOR DATA ETHICS & INNOVATION, REVIEW INTO BIAS IN ALGORITHMIC DECISION-MAKING 7 (2020) [hereinafter CDEI] ("Though there is strong momentum in data ethics in policing at a national level, the picture is fragmented with multiple governance and regulatory actors, and no single body fully empowered or resourced to take ownership."). The CDEI is a "government expert body enabling the trustworthy use of data and AI." *See About Us*, CDEI, https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation/about#:~:text=Overview-,The%20CDEI%20is%20a%20government%20expert%20body%20enabling,use%20of%20data%20and%20AI.&text=The%20CDEI%20is%20committed%20to,core%20component%20of%20its%20work (last visited Feb. 27, 2022).

discrimination, to cut costs, and to innovate. In this context, AI systems offer tools that promise faster and more efficient methods of investigation and police administration. But their adoption into police decision-making and tactics also introduces complications. Any police department interested in guidelines for ethical use of AI systems would "find a field with few existing examples and no established guidelines or best practices."[15]

Commitments to ethical and responsible principles in the police use of AI have a role here. They aren't substitutes for regulation or judicial decision-making. However, legislators and judges have been slow. The United States lacks a national, comprehensive approach to the regulation of AI systems.[16] Instead, state and local governments have been left to decide whether and how to regulate AI systems either based on a particular industry or on specific use cases. Similarly, there have been a small number of cases challenging the use of AI systems in the courts, but not enough to conclude that a body of rules have been developed.[17] This means that policing in particular is guided by an uncertain set of rules and legal decisions for the adoption and use of AI-based systems. And while ethical and legal principles share common concerns, ethical principles broaden the set of possible questions police departments should consider.[18]

Many AI policy guidance documents exist now, but their value to the police is limited. Simply repeating broad principles about the responsible use of AI systems are less helpful than ones that 1) take into account the specific context of policing, and 2) consider the American experience of policing in particular. There is an emerging consensus

---

[15] *See Use of New Artificial Intelligence Technologies Policy – Public Consultation*, TORONTO POLICE SERV. BD. (2022) [hereinafter Toronto Police Services Board], https://tpsb.ca/ai.

[16] *See* Heather Sussman et al., *U.S. Artificial Intelligence Regulation Takes Shape*, ORRICK (Nov. 18, 2021), https://www.orrick.com/en/Insights/2021/11/US-Artificial-Intelligence-Regulation-Takes-Shape (contrasting developments in EU while noting "there is currently no federal regulation of AI in the U.S.").

[17] Jessica Field et al., *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*, (2020) [hereinafter Berkman Klein Report], https://cyber.harvard.edu/publication/2020/principled-ai ("Litigation over the harmful consequences of AI technology is still nascent, with just a handful of cases having been brought. Similarly, only a few jurisdictions have adopted regulations concerning AI . . .").

[18] *Cf.* Lexo Zardiashvili et al., *AI Ethics for Law Enforcement: A Study into Requirements for Responsible Use of AI at the Dutch Police*, 2 DELPHI 1, 2 (2019) (arguing that "for such spaces left open by the law, the police can, and we advise that they should incorporate 'ethics' through practical measures to ensure responsible use of AI and contribute toward enhancing (rather than limiting) legitimacy of and trust in the police.").

about what ethical and responsible values should be part of AI systems. This essay considers what kind of ethical considerations can guide the use of AI systems by American police.

I. AI SYSTEMS AND THEIR USE IN POLICING

Anyone taking a first look at the use of AI systems in policing would be justifiably confused. New tools are alternatively described as data-driven, based on artificial intelligence, powered by algorithms, or new surveillance technologies. Are these terms meaningfully different? We can begin by looking at what we mean by an AI system, and how police are using these tools.

First, there is no single widely accepted definition of artificial intelligence.[19] But many policy documents from around the world define AI in terms of software that can achieve a complex goal by acting upon collected information and then processing or interpreting that data. Sometimes an AI system will adapt its behavior by analyzing the environment changed by its previous actions.[20] This use of algorithms, combined with cheap and powerful computer processing, and massive amounts of data has also sometimes been referred to as the use of "big data."[21]

To add to these ambiguities, some discussions of AI systems in policing might also use the term "data-driven" policing: a term that captures both AI systems today and earlier efforts dating back to the 1990s that simply emphasize the increasing reliance of police decision-making on statistics.[22] Finally, discussions of AI systems in policing like

---

[19] The term "artificial intelligence" was first coined by John McCarthy in 1955, who defined it as "the science and engineering of making intelligent machines," but that definition is just one among many today. *See* PETER STONE ET AL., STAN. UNIV., ARTIFICIAL INTELLIGENCE AND LIFE IN 2030: REPORT OF THE 2015 STUDY PANEL 50 (2016) ("McCarthy is credited with the first use of the term "artificial intelligence" in the proposal he co-authored for the workshop with Marvin Minsky, Nathaniel Rochester, and Claude Shannon."); *see also* Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 404 (2017) ("There is no straightforward, consensus definition of artificial intelligence.").

[20] This particular definition is derived from the European Commission's High-Level Expert Group on Artificial Intelligence, but many similar ones exist. *See, e.g.*, Berkman Klein Report, *supra* note 17, at 4-5.

[21] *See, e.g.*, IMPACT ON FUNDAMENTAL RIGHTS, *supra* note 1, at 21 (defining AI as including "data, algorithms, and computer power" and acknowledging overlap with "big data").

[22] *See, e.g.*, Annie Gilbertson, *Data-Informed Predictive Policing was Heralded as Less Biased. Is It?*, THE MARKUP (Aug. 20, 2020), https://themarkup.org/ask-the-markup/2020/08/20/does-predictive-police-technology-contribute-to-bias ("Early

predictive policing or facial recognition software sometimes focus on their increased surveillance capacity and are described as new surveillance technologies.[23] All of these terms are sometimes used interchangeably. For simplicity's sake, we can use the term "AI systems" here. All of these technologies introduce some degree of automated decision-making into what had traditionally been the entirely human process of police work.

In theory, AI systems can introduce efficiency and innovation to a field that is as much about the management of risk and the processing of information as it is about stops and arrests. Identifying patterns in large sets of data can help the police prioritize where their officers and dollars go.[24]

In policing, the AI systems that have received the most attention are probably facial recognition software and predictive policing. Predictive policing software can take a variety of forms, but at their most basic they rely on past information to make forecasts about the future: whether crimes are likely to occur in particular places, or whether people are likely to engage in some kinds of crimes or become victims of crime.[25] In 2011, the police department in Santa Cruz, California became one of the first in the United States to pilot a predictive policing program, one developed by the private company PredPol (now Geolitica).[26] That program assessed historical crime data and directed its client, the Santa Cruz police, to those five hundred square foot areas where crime was

---

versions of data-driven policing were used in the 1990s, but it has grown more popular and the technology more sophisticated over the last decade.").

[23] *See, e.g.*, Andrew G. Ferguson, *Surveillance and the Tyrant Test*, 110 GEO. L. J. 205, 210 (2021) (characterizing tools like facial recognition and license plate readers as "new surveillance technologies").

[24] *See, e.g.*, CDEI, *supra* note 14, at 64 ("In theory, tools which help spot patterns of activity and potential crime, should lead to more effective prioritization and allocation of scarce police resources.").

[25] *See, e.g.*, IMPACT ON FUNDAMENTAL RIGHTS, *supra* note 1, at 22 (defining predictive policing as "the algorithmic processing of data sets . . . to reveal patterns of probable future offending and victimization, which can thus be interdicted before they happen"). Examples of predictive software about persons include Chicago's "Strategic Subjects List," which identified persons at high risk of being involved in future gun violence as perpetrators or victims. *See, e.g.*, Mick Dumke & Frank Main, *A Look Inside the Watch List Chicago Police Fought to Keep Secret*, CHI. SUN TIMES (May 18, 2017), https://chicago.suntimes.com/2017/5/18/18386116/a-look-inside-the-watch-list-chicago-police-fought-to-keep-secret (describing risk assessment that scored individuals and listed 398,000 entries in 2017). Another example is the UK Metropolitan Police's use of the Gangs Violence Matrix, a tool to identify those at risk of gang violence as perpetrators or victims. *See* IMPACT ON FUNDAMENTAL RIGHTS, *supra* note 1, at 24.

[26] *See* Erica Goode, *Sending the Police Before There's a Crime*, N.Y. TIMES (Aug. 15, 2011), https://www.nytimes.com/2011/08/16/us/16police.html (describing Santa Cruz's "unusual experiment" to test a prediction method for property crimes).

likely to occur.[27] Dozens of police departments piloted and adopted similar programs in the following years.[28]

Like predictive policing, facial recognition technology is a broad term. The technology uses an algorithm to see if one image can be matched against another in an existing database of images. To deliver results, a facial recognition program must collect images, classify them, train that data, and test these training sets.[29] These comparisons can be used in many ways. For instance, face verification confirms your identity against a stored image.[30] Face identification involves matching a suspect's face to a database of existing images, like a driver's license records.[31] Or, the technology might be used for generalized surveillance, to identify many people in places like airports or public streets.[32]

Predictive policing and facial recognition have received the most public attention in policing, and for good reason. Predictive policing threatens to replace the seemingly unique skill of human police expertise. The assessments of suspicious persons and places by police officers poses its own problems, of course, but turning over some of this decision-making to machines preys on people's suspicions about how trustworthy these assessments are.[33] And the potential of facial recognition to

---

[27] *See id.*

[28] The Electronic Frontier Foundation's Atlas of Surveillance has identified at least 160 agencies using predictive policing as of January 2022. *See Atlas of Surveillance*, ELEC. FRONTIER FOUND. (Jan. 11, 2022), https://atlasofsurveillance.org/atlas.

[29] *See* Andrew Ferguson, Facial Recognition and the Fourth Amendment, 105 MINN. L. REV. 1105, 1112 (2021). Face recognition algorithms learn to identify important facial features by being trained through the comparison of data. An algorithm might be given pairs of face images of the same person; over time, it recognizes that some features act as reliable identifying signals about the same person. *See* CLARE GARVIE, ALVARO M. BEDOYA & JONATHAN FRANKLE, GEORGETOWN L. CTR. ON PRIV.&TECH.,THE PERPETUAL LINE-UP:UNREGULATED POLICE FACE RECOGNITION IN AMERICA 1 (2016), https://www.perpetuallineup.org/sites/default/files/2016-12/The%20 Perpetual%20Line-Up%20-%20Center%20on%20Privacy%20and%20Technology% 20at%20Georgetown%20Law%20-%20121616.pdf.

[30] *See id.* at 109.

[31] *See id.* at 108 (discussing this as "face identification").

[32] *See id.* (discussing this as "face surveillance").

[33] For example, research from DeepMind and the U.K.'s RSA found that sixty percent of survey respondents opposed or strongly opposed the use of automated decision-making in the criminal justice system and the workplace. *See*, BRHMIE BALARAM ET AL., ROYAL SOCIETY FOR THE ENCOURAGEMENT OF ARTS, MANUFACTURERS, AND COMMERCE, ARTIFICIAL INTELLIGENCE: REAL PUBLIC ENGAGEMENT 4 (2018). Similarly, a 2018 Pew Research report found that majorities of Americans surveyed found it "unacceptable" for algorithms to make decisions with "real-world consequences for humans," including criminal risk assessments for people considered for parole. *See Public Attitudes Toward Computer Algorithms*, PEW RSCH. CTR. (Nov. 16, 2018), https://www.pewresearch.org/internet/2018/11/16/public-attitudes-toward-computer-algorithms/.

identify hundreds, even thousands of people in moments sparks concerns about unchecked surveillance power.[34]

But AI systems have other roles in policing as well. While many police departments had been using remote-controlled drones, newer versions are more similar to autonomous cars than radio-controlled toy cars.[35] An autonomous police drone can respond quickly to a 911 call and provide the police with details as they assemble their human response.[36] A police drone can also fly into enclosed spaces for surveillance where the police are concerned about unknown threats.[37] Similarly, the inevitable introduction of autonomous cars will mean not just autonomous police cars, but also the possibility of remote stops of cars by the police.[38]

Other AI systems can address police issues that are important but don't generate the same public concern. Most of us don't focus on the administrative parts of policing, but police officers devote enormous amounts of time to pushing paper and filling out forms.[39] The paperwork associated with arrests, for instance, takes up so much time that it can provide a perverse incentive for some officers to use arrests as an excuse for overtime pay.[40] AI systems can make these processes less cumbersome by automating form-filling and aggregating information. Companies like Axon Enterprise and Mark43 offer cloud-based records based management (RMS) systems that automate some of the report-

---

[34] *See, e.g.*, Laura K. Donohue, *Technological Leap, Statutory Gap, and Constitutional Abyss: Remote Biometric Identification Comes of Age*, 97 MINN. L. REV. 407, 415 (2012) (arguing that remotely used biometric technologies like face recognition are "significantly different from that which the government has held at any point in U.S. history").
[35] Cade Metz, *Police Drones Are Starting to Think for Themselves*, N.Y. TIMES (Dec. 5, 2020), https://www.nytimes.com/2020/12/05/technology/police-drones.html.
[36] *See id.*
[37] *See id.*
[38] *See, e.g.*, Elizabeth E. Joh, *Automated Seizures: Police Stops of Self-Driving Cars*, 94 N.Y.U. L. REV. (ONLINE ISSUE) 113 (2019).
[39] *See, e.g.*, Brad W. Smith et al., *Community Policing and the Work Routines of Street-Level Officers*, 26 CRIM. JUST. REV. 17, 31 (2001) (reporting research that "administrative activities consumed a significant portion [of an officer's daily shift].").
[40] *See, e.g.*, EDITH LINN, ARREST DECISIONS: WHAT WORKS FOR THE OFFICER? 1 (2009)(finding that the overtime pay associated with arrest procedures influences police officer behavior).

taking process.[41] Axon's CEO even envisions an entirely automated information flow one day from body camera video to police report.[42]

In sum, AI systems are already a part of ordinary police work. Public attention tends to focus on a few applications that are controversial because they raise the specter of vastly increased police power with new risks and few checks. But AI systems also assume other tasks in policing, including through some seemingly mundane tasks that are nevertheless central to what police do: processing information to investigate crime.

II. THE SECOND WAVE OF AI SYSTEMS IN POLICING

Today AI systems in policing find a very different audience from the one that endorsed predictive policing as one of the fifty "best inventions of the year" in 2011.[43] If the 2010s can be characterized as an enthusiastic embrace of novel police technologies, the 2020s could be deemed a second wave of AI-based systems in policing.[44] It is a second wave not only because there is much more use of AI everywhere, but also because the social and political context has changed as well. Civil rights organizations, policymakers, and scholars have pointed out the shortcomings of those AI systems already in place. And the harms of AI systems in policing are no longer theoretical. People have been mistakenly stopped and arrested because of mistaken AI

---

[41] *See, e.g.*, Thad Rueter, *Mark43 Raises $101M to Expand Police Tech Products*, GOVTECH BIZ (July 12, 2021), https://www.govtech.com/biz/mark43-raises-101m-to-expand-police-tech-products (citing evidence for "increased spending for [cloud based records management] even amid pandemic spending cuts and the broad 'defund the police' movement in the U.S. that calls for government to shift some of law enforcement's responsibilities to other agencies"); Peter Hall, *New Record Keeping Software Will Make It Easier for Lehigh County Police Departments to Share Information*, MORNING CALL (Sept. 13, 2021), https://www.mcall.com/news/local/mc-nws-lehigh-county-police-record-software-upgrade-20210913-ziw73kxxorfsxokseg3w5bjbsy-story.html (describing new $3.6 million dollar three year contract with Mark43 which will provide cloud based report writing software including predictive language use).

[42] *See also* Dana Goodyear, *Can the Manufacturer of Tasers Provide the Answer to Police Abuse?*, THE NEW YORKER (Aug. 20, 2018), https://www.newyorker.com/magazine/2018/08/27/can-the-manufacturer-of-tasers-provide-the-answer-to-police-abuse.

[43] Lev Grossman et al., *The 50 Best Inventions*, TIME MAG. (Nov. 28, 2011), http://content.time.com/time/subscriber/article/0,33009,2099708,00.html.

[44] These terms are loosely based upon Frank Pasquale's description: "While the first wave of algorithmic accountability focuses on improving existing systems, a second wave of research has asked whether they should be used at all—and, if so, who gets to govern them." Frank Pasquale, *The Second Wave of Algorithmic Accountability*, LPE PROJECT (Nov. 25, 2019), https://lpeproject.org/blog/the-second-wave-of-algorithmic-accountability/.

determinations.[45] And there are likely many, many people who received greater police scrutiny short of a physical encounter because an AI system flagged them for attention.

A. Critiques of AI Systems in Policing

There are now some well-established criticisms of the use of AI systems in policing. We can divide them broadly into three categories: bias, privacy, and secrecy. The data used in these systems may be biased.[46] The design of the systems may reflect the biases of the engineers who created them. These biases can in turn amplify biases against marginalized groups, or even create new forms of bias.[47]

As costs for data collection, storage, and analysis become ever cheaper, the police gain the ability to conduct indiscriminate mass surveillance. These capabilities can chill speech, the ability to freely associate with others, and to remain anonymous.[48] Each of these data points, whether collected directly by the police or by third parties like cellphone apps, may seem unworthy of privacy protection. But in the aggregate, they form the ability to create a time machine into our past movements, and sometimes our real-time movements as well.

Discovering how American law enforcement agencies use AI-based systems has been challenging because of their secrecy and opacity. One type of secrecy happens when some AI systems can make determinations about data in ways that even developers cannot completely explain.[49] This black box problem may have few consequences in some applications, like chatbots for recreation. But there are—and increasingly will be—many situations where people feel

---

[45] As of January 2022, there are at least three known cases where facial recognition technology provided a mistaken match. *See* Kashmir Hill, *Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match*, N.Y. TIMES (Dec. 29, 2021), https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html.

[46] CDEI *supra* note 14, at 68 ("Police data can be biased due to it either being unrepresentative of how crime is distributed or in more serious cases reflecting unlawful policing practices.").

[47] *Cf.* CDEI, *supra* note 14, at 21 ("There is clear evidence that algorithmic bias can occur, whether through entrenching previous human biases or introducing new ones.").

[48] *See Perpetual Line-Up*, *supra* note 4, at 41-44 ("Despite the fact that leading law enforcement agencies . . . have explicitly recognized the potential chilling effect of face recognition on free speech, we found that almost none of the agencies using face recognition have adopted express prohibitions against using the technology to track political or other First Amendment activity.").

[49] *See, e.g.*, Calo, *supra* note 19, at 414 (observing that deep learning AI systems "can say what will happen but not why").

the real impacts of such inscrutable decisions. They will be turned down for a loan, or stopped by the police. That is why "explainability" is a widely shared principle from AI-guidance proposals from around the world.[50]

Another type of secrecy in many AI systems, particularly in the United States, stems from companies making claims that disclosure will harm their intellectual property rights.[51] This means that trying to find out about the AI-based system—even one that directly impacted your life in some way—may be nearly impossible to find out. The company that developed it may claim that providing important information might divulge a trade secret.[52] A public agency that uses the AI system might also claim that it is bound by a non-disclosure agreement entered into with that same company.[53]

The response to these issues has been uneven. There is widespread agreement that the increasing use of AI systems needs guidance.[54] A survey of more than thirty documents stating AI principles from around the world identified several shared themes.[55] These included values important for policing: privacy,[56] accountability,[57]

---

[50] It's also true that the field of "explainable AI" (XAI) has not achieved consensus on how exactly this value can be implemented in practice. *See, e.g.*, Jessica Newman, *Explainability Won't Save AI*, BROOKINGS (May 19, 2021), https://www.brookings.edu/techstream/explainability-wont-save-ai/ (noting that "the XAI field has generally struggled to realize the goals of understandable, trustworthy, and controllable AI in practice.").

[51] *See generally* Elizabeth E. Joh, *The Undue Influence of Surveillance Technology Companies on Policing*, 92 N.Y.U. L. Rev. (ONLINE ISSUE) 101 (2017) [hereinafter *Undue Influence*]; Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343 (2018).

[52] *Undue Influence*, *supra* note 51, at 125-126 (discussing TrueAllelle's citing of trade secrets for non-disclosure).

[53] *Id.* at 104-08 (discussing use of non-disclosure agreements to shield details of cell site simulator technology).

[54] *See, e.g.*, Calo, *supra* note 19, at 411 ("Perhaps the most visible and developed area of AI policy to date
involves the capacity of algorithms or trained systems to reflect human values such as fairness, accountability, and transparency ("FAT")).

[55] Berkman Klein Report, *supra* note 17, at 4-5.

[56] "Privacy" is defined as referring to the idea that "AI systems should respect individuals' privacy, both in the use of data for the development of technological systems and by providing impacted people with agency over their data and decisions made with it." *Id.* at 4.

[57] "Accountability" is defined as including mechanisms to ensure that those impacted by AI systems have appropriate remedies and that AI's effects are appropriately distributed. *Id.*

transparency and explainability,[58] fairness and non-discrimination,[59] human control of technology,[60] and the promotion of human values.[61] Most of us would agree that these are worthy goals.

B. Attempts to Regulate Police AI Systems

How these values have translated into practice is another matter. The United States has no national legislation on the use of AI-based systems, in any field. What has occurred in this absence is a patchwork of solutions. This section discusses some of the most prominent efforts to regulate AI in policing and their shortcomings.

First, there have been attempts to regulate the police use of surveillance technologies, of which AI-based systems are a part, by enacting local ordinances at the city or county level. In 2016, the ACLU launched an initiative to help local communities pass laws requiring oversight and transparency about the police use of new technologies.[62] Its Community Control Over Police Surveillance (CCOPs) campaign, supported by many civil rights groups,[63] published a model ordinance to serve as a template for local governments to follow.[64] Key features of the model act include the requirement of explicit approval for the purchase or use of new surveillance technologies,[65] the requirement of surveillance

---

[58] "Transparency" and "explainability" include the translation of "operations into intelligible outputs and the provision of information about where, when, and how they are being used." *Id.*

[59] "Fairness" and "non-discrimination" are defined as designing AI "to maximize fairness and [to] promote inclusivity." *Id.*

[60] "Human control of technology" refers to a requirement that "important decisions remain subject to human review." *Id.*

[61] "Promotion of human values" refers to the idea that "the ends to which AI is devote . . . should correspond with our core values and generally promote humanity's well-being." *Id.*

[62] *Community Control over Police Surveillance (CCOPS)*, ACLU, https://www.aclu.org/issues/privacy-technology/surveillance-technologies/community-control-over-police-surveillance?redirect=feature/community-control-over-police-surveillance (last visited Feb. 27, 2022).

[63] *See* Dave Maass, *Join the Movement for Community Control Over Police Surveillance*, ELEC. FRONTIER FOUND. (Sept. 21, 2016), https://www.eff.org/deeplinks/2016/09/join-movement-community-control-over-police-surveillance.

[64] *Community Control Over Police Surveillance (CCOPS) Model Bill*, ACLU (Apr. 2021) [hereinafter CCOPS Model Bill], https://www.aclu.org/legal-document/community-control-over-police-surveillance-ccops-model-bill.

[65] *Id.* at Section 1.

impact reports and other surveillance data,[66] and the creation of community advisory committees.[67]

In practice, this model has not found wide adoption. A 2020 study found that only fourteen local governments had passed local ordinances regulating police use of new surveillance technologies.[68] In other cities, proposals have been defeated or stalled. The reasons are varied, but this kind of intensive local oversight of police can be a difficult political project.[69] Their slow place and infrequent adoption thus far means that local administrative regulations are unlikely to provide significant constraints or guidance soon.

A second important development can also be found in cities around the country. In 2019, the Board of Supervisors voted to ban the use of facial recognition by its police and other public agencies.[70] In 2020, the city of Santa Cruz, California became the first American city to ban the use of predictive policing software.[71] A few dozen other local governments have followed their lead in considering bans or moratoria on the use of specific technologies, particularly facial recognition software.[72]

While civil liberties organizations have lauded these measures as successes, they have limits. On the one hand, bans are blunt tools with an intuitive appeal. They impose easy-to-understand total embargoes. But these bans are problematic. Technology-specific bans can simultaneously be both blunt but too narrow. They address only one specific system, such as facial recognition technology in body cameras, without addressing other AI-based systems that might pose similar

---

[66] *Id.* at Sections 6-7.
[67] *Id.* at Section 8.
[68] Mailyn Fidler, *Local Police Surveillance and the Administrative Fourth Amendment*, 36 SANTA CLARA HIGH TECH. L. J. 481, 545 (2020).
[69] Mailyn Fidler & Lily Liu, *Four Obstacles to Local Surveillance Ordnances*, LAWFARE (Sept. 4, 2020), https://www.lawfareblog.com/four-obstacles-local-surveillance-ordinances (identifying objections from politically strong mayors, police lobbying, an overemphasis on surveillance cameras, and concerns about public safety and overregulation as obstacles that stalled attempts at local oversight of police technologies).
[70] Kate Conger et al., *San Francisco Bans Facial Recognition Technology*, N.Y. TIMES (May 14, 2019), https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html.
[71] Kristi Sturgill, *Santa Cruz Becomes the First U.S. City to Ban Predictive Policing*, L.A. TIMES (June 26, 2020), https://www.latimes.com/california/story/2020-06-26/santa-cruz-becomes-first-u-s-city-to-ban-predictive-policing.
[72] Kashmir Hill, *How One State Managed to Actually Write Rules on Facial Recognition*, N.Y. TIMES (Feb. 27, 2021), https://www.nytimes.com/2021/02/27/technology/Massachusetts-facial-recognition-rules.html?searchResultPosition=3 (noting that Oakland, Portland, San Francisco, and Minneapolis have banned use of facial recognition technology).

harms.[73] There is a larger problem with advocating bans, too. Why should the police be barred from the potential benefits of the digital world, as every other sector of society moves in this direction?[74] Whatever the potential risks that arise from police use of AI systems, it would be strange to conclude that the solution would be a total prohibition on their use in law enforcement.

Third, some courts are beginning to consider the harms of AI-based systems with seriousness. These issues have been considered as traditional criminal procedure claims, such as the Wisconsin Supreme Court's 2016 decision that a proprietary risk assessment tool used to sentence the defendant did not violate his due process rights.[75]

But there are limits to this approach as well. Let's look at the framework of constitutional criminal procedure. By raising claims about the police tactics used in their own cases, defendants help define all of our rights. But defendants are inadequate proxies in the case of AI systems. In order to raise a criminal procedure claim, a defendant has to identify the evidence that came about as a result. But the police might rely on an AI system for the early stages of an investigation without collecting evidence. Or, the police might use an AI system for indiscriminate surveillance that only sometimes leads to the prosecution of individuals. At the same time, most of us would probably agree that the police should not use AI systems without any rules at all.

To be sure, the pursuit of local surveillance oversight mechanisms, the passage of bans for demonstrably flawed AI systems, and increasing judicial awareness of their pitfalls have made progress. Such measures have made the procurement of these tools and their costs more transparent, and thus more amenable to oversight. But ethical guidelines can address a broader set of issues in policing, including those situations where there may be not be harms in a traditional legal sense.

---

[73] *Cf.* Bruce Schneier, *We're Banning Facial Recognition. We're Missing the Point.*, N.Y. TIMES (Jan. 20, 2020), https://www.nytimes.com/2020/01/20/opinion/facial-recognition-ban-privacy.html?smid=url-share ("A ban on facial recognition won't make any difference if, in response, surveillance systems switch to identifying people by smartphone MAC addresses. The problem is that we are being identified without our knowledge or consent, and society needs rules about when that is permissible.").

[74] Andrew Ferguson makes a similar observation about what he characterizes as the "trap lens" with regard to new surveillance technologies. Ferguson, *supra* note 23, at 241 (noting that police abolitionists and advocates of bans "need to make an argument about why policing does not deserve to evolve in a digital world" when "every other professional enterprise has benefited from technological innovation").

[75] State v. Loomis, 881 N.W.2d 749, 770 (Wis. 2016) (holding that "if used properly with an awareness of the limitations and cautions, a circuits court's consideration of a COMPAS risk assessment at sentencing does not violate a defendant's right to due process").

III. THE CONTEXT OF AMERICAN POLICING

AI systems are already being used in American policing. Yet few police departments face any significant oversight or regulation. This is where a discussion of ethical and responsible policing might provide further guidance. Guidelines for responsible use of AI systems in policing are already a topic of public debate elsewhere.[76] Any conversation about such guidelines, however, should consider the specific context of American policing. In particular, we should highlight 1) the highly decentralized nature of American policing and 2) the longstanding racial tensions that are part of American police history.

A. Decentralization of Policing

One of the most distinctive aspects of American policing is its extreme decentralization.[77] To speak of "the police" in the United States is really to refer to the more than 18,000 individual law enforcement agencies, most of which are organized at the city and county levels.[78] There are more than 12,000 local police departments alone.[79] The most common type of agency is a small one, with ten or fewer offices, significantly smaller than the 40,000 officers in the New York Police Department.[80] And because most of these agencies are organized at the city or county level, they are controlled at the local level. States can and do impose rules on what police departments do within their borders, but not on every subject, and little has been done to control the police use of AI systems. Although the federal government can regulate, for instance, the private companies that design, sell, and use AI systems, it cannot regulate directly how states control their police agencies.[81] While the

---

[76] The Toronto Police Department, for instance, is currently developing an ethics policy for its use of AI systems. *See* Toronto Police Services Board, *supra* note 15.
[77] In fact, policing is so decentralized we have hard time counting how many agencies even exist. DUREN BANKS ET AL., U.S. DEP'T OF JUSTICE, NAT'L SOURCES OF LAW ENFORCEMENT EMPLOYMENT DATA 1 (2016) ("The decentralized, fragmented, and local nature of law enforcement in the United States makes it challenging to accurately count the number of agencies and officers.").
[78] *See id.*
[79] SHELLEY S. HYLAND & ELIZABETH DAVIS, U.S. DEP'T OF JUSTICE, LOCAL POLICE DEPARTMENTS, 2016: PERSONNEL 1 (2019).
[80] *See* BANKS ET AL., *supra* note 77. *See also* HYLAND & DAVIS, supra note 79, at 2 (observing that 48% of all local police departments employed less than 10 full time officers).
[81] *See, e.g.*, Printz v. United States, 521 U.S. 898, 936 (1997) ("The Federal Government may neither issue directives requiring the States to address particular

federal government can condition federal grants on changes in police conduct, there are few signs that funds used for technology purchases have any real constraints.[82]

When it comes to the tools police use, local officials like mayors and city councils are often the ones with the ability to impose conditions and requirements. Here, local communities can provide input, but as we saw earlier, local control of AI systems in policing has enjoyed limited success. Of the thousands of local governments, fewer than twenty have imposed any sort of regulations or requirements over how police can acquire or use these technologies.[83] While the pandemic has shown that communities can be engaged in and vocal about issues of local government, AI systems generate far less local engagement. This may be for a variety of reasons. People may readily accept police justifications that these systems are necessary innovations for criminal investigations. And many of these AI systems, including any potential for the harms or risks they pose, may be hard to explain and understand.

B. Racial Bias and Inequality

Concerns about bias are, of course, present in policing systems around the world.[84] However, the use of AI systems in American policing should be sensitive to our own particular context, history, and experiences. To raise the concern that AI systems used by the police might harbor bias or exhibit discriminatory behavior is to miss the point. Even as the murder of George Floyd while in police custody provoked

---

problems, nor command the States' officers, or those of their political subdivisions, to administer or enforce a federal regulatory program."). Accordingly, a proposed Algorithmic Accountability Act would direct Federal Trade Commission to require *companies* to reduce bias and improve privacy protections in the algorithms they produce. *See* Press Release, Office of Sen. Ron Wyden, Wyden, Booker, Clarke Introduce Bill Requiring Companies to Target Bias in Corporate Algorithms (Apr. 10, 2019), https://www.wyden.senate.gov/news/press-releases/wyden-booker-clarke-introduce-bill-requiring-companies-to-target-bias-in-corporate-algorithms-.

[82] The millions distributed by the federal government for body cameras is a good example. Regulation over body camera use has been left up to states, cities, and individual departments. *See, e.g.,* Urban Institute, Police Body-Worn Camera Legislation Tracker (2018), at https://apps.urban.org/features/body-camera-update/ (noting that laws "governing how and when police body-worn cameras can be used and whether the footage is releasted vary considerably across the country").

[83] *See* Fidler, *supra* note 68.

[84] For instance, an important 2017 review of deaths in police custody commissioned by the UK Home Secretary stated that "Deaths of people from BAME communities, in particular young Black men, resonate with the Black community's experience of systemic racism." *See* ELISH ANGIOLINI, REPORT OF THE INDEPENDENT REVIEW OF DEATHS AND SERIOUS INCIDENTS IN POLICE CUSTODY 84 (2017).

national, and even global calls for greater accountability in policing, unequal and discriminatory policing is part of the history of American policing.[85] Before George Floyd, there were the deaths of Freddie Gray and Michael Brown. And before that was the death of Amadou Diallo and the abuse of Abner Louima. And before that, the beating of Rodney King. We could add to these individual cases the systematic reporting of racially biased policing against Black and Hispanic drivers,[86] Black and Hispanic pedestrians,[87] and even Black and Hispanic bicyclists.[88] The impacts of inequitable policing, then, are by definition unevenly experienced. Such experiences have left those most vulnerable to over-policing and discriminatory practices "legally estranged" from their own

---

[85] Former Minneapolis police officer Derek Chauvin was convicted of second-degree unintentional murder, third-degree murder and second-degree manslaughter after a widely circulated video captured him pressing his knee against George Floyd's neck on May 25, 2020. Police had responded to a call that Floyd had used a counterfeit twenty-dollar bill to buy cigarettes. *See* John Eligon et al., *Derek Chauvin Verdict Brings a Rare Rebuke of Police Misconduct*, N.Y. TIMES (Apr. 20, 2021), https://www.nytimes.com/2021/04/20/us/george-floyd-chauvin-verdict.html; Amy Forliti, *Explainer: What Next After Chauvin's Conviction on 3 Counts*?, ASSOCIATED PRESS, (Apr. 20, 2021), https://apnews.com/article/derek-chauvin-trial-charges-716fa235ecf6212f0ee4993110d959df.

[86] There are numerous studies here, dating back to the 1990s. A pioneering observational study by John Lamberth found that African Americans made up 13.5% of the population on the New Jersey turnpike and 15% of speeders but represented 35% of those pulled over by the police. In other words, African Americans were 4.85 times as likely to be stopped as others. *See* John Lamberth, *Driving While Black*, WASH. POST (Aug. 16, 1998), https://www.washingtonpost.com/archive/opinions/1998/08/16/driving-while-black/23ecdf90-7317-44b5-ac43-4c9d7b874e3d/ (summarizing his study's methodology and findings); *see also* David A. Harris, *The Stories, the Statistics, and the Law: Why "Driving While Black" Matters*, 84 MINN. L. REV. 265 (1999). The nonpartisan Public Policy Institute of California provides a recent example of similar findings. *See* Magnus Lofstrom et al., *African Americans Are Notably Overrepresented in Police Stops*, PPIC (Aug. 13, 2020), https://www.ppic.org/blog/african-americans-are-notably-overrepresented-in-police-stops/ (finding in review of 1.8 million police stops, "the data clearly shows that African-Americans make up a much larger share of interactions with law enforcement relative to their populations [sic] share than any other racial/ethnic group in California").

[87] *See, e.g.*, Lyndsay Winkley & Teri Figueroa, *Another Report Finds Deep Racial Disparities in Sheriff's Departments Stop Data*, SAN DIEGO UNION-TRIB. (Dec. 9, 2021), https://www.sandiegouniontribune.com/news/public-safety/story/2021-12-09/another-report-finds-deep-racial-disparities-in-sheriffs-department-data (citing Center for Policing Equity study finding "Black pedestrians were stopped by sheriff's deputies 3.5 times as often" compared to Whites).

[88] *See, e.g.*, Alene Tchekmedyian et al., *L.A. Sheriff's Deputies Use Minor Stops to Search Bicyclist, With Latinos Hit Hardest*, L.A. TIMES (Nov. 4, 2021), https://www.latimes.com/projects/la-county-sheriff-bike-stops-analysis/ (documenting more than 44,000 bike stops logged by the Sheriff's Department and finding 7 of 10 stops involved Latino cyclists).

police departments.[89] These and countless other incidents in American policing have generated countless commission reports, lawsuits, and calls for reform for nearly a century.[90]

Thus the risks of AI in policing arise in the context of an institution that has a long history of meting out justice unequally and in a discriminatory way. What follows? First, bias in AI systems can perpetuate existing biases or introduce new ones, but it does so in the context of a social institution with a long history of discrimination, especially against African-Americans. We should not be surprised, then, if the use of an AI system in a community in longstanding tension with its local police department meets skepticism, resistance, or calls for prohibition.

Second, crafting AI ethics for policing requires speaking to two different audiences. Each is important but distinct. One audience is engaged primarily in "tech policy": the drafting and decision-making of rules and policies that engage in the use of technologies across industries and institutions. Advocacy organizations and policymakers engaged in AI policy often address the use of AI in matters that can include online speech, advertising, healthcare, lending, and employment. Policing is only one subject, and subsumed under criminal justice policy, at that. And even when policing is a concern, this tech policy lens tends towards a focus on individual privacy and the harms of mass surveillance.

On the other hand, the Black Lives Matter movement and related campaigns have focused on police violence and addressing longstanding structural problems in the relationship between the police and marginalized communities. Young African-American men make up an overwhelming number of those killed by police, year after year.[91] Many

---

[89] Monica Bell's theory of legal estrangement describes this problem well: one that captures "both legal cynicism-the subjective "cultural orientation" among groups 'in which the law and the agents of its enforcement, such as the police and courts, are viewed as illegitimate, unresponsive, and ill equipped to ensure public safety' and the objective structural conditions (including officer behaviors and the substantive criminal law) that give birth to this subjective orientation." Monica C. Bell, *Police Reform and the Dismantling of Legal Estrangement*, 126 YALE L. J. 2054, 2066-67 (2017).

[90] President Hoover's commission of the Report of the Enforcement of the Prohibition Laws, better known as the Wickersham Report, was among the first national reports focusing on problems in policing. *See* NATIONAL COMMISSION ON LAW OBSERVANCE AND ENFORCEMENT, REPORT NO. 2, REPORT ON THE ENFORCEMENT OF THE PROHIBITION LAWS OF THE UNITED STATES (1931).

[91] Starting in 2015, the Washington Post has tracked every fatal shooting by a police officer in the United States. Among its findings is the observation that African Americans are killed by the police at more than twice the rate of Whites. *See Fatal Force: 1022 People Have Been Shot and Killed by Police in the Past Year*, WASH. POST

Black and Hispanic communities are simultaneously over-policed and under-policed.[92] We know from federal investigations that some municipal budgets literally depend on fines and fees, almost always imposed on the poor, and always meted out by local police.[93] These problems have been rightly identified as reasons for desperately needed police reforms.

To be sure, there are groups and voices that have brought these two concerns together. Some civil rights groups have made explicit the disproportionately borne harms of unregulated AI systems on marginalized communities.[94] This has led, for instance, to a coalition of civil rights groups to publish "civil rights principles for the era of big data."[95]

IV. ETHICAL COMMITMENTS IN AI-SYSTEMS IN POLICING

What then, do we mean by the ethical use of AI in American policing? Police departments should make prior public commitments to the values they adopt as they rely on AI systems of all types. Ethical commitments can serve as meaningful guides, even if they lack penalties or enforcement consequences.[96] These commitments should embody

---

(Feb. 2, 2022), https://www.washingtonpost.com/graphics/investigations/police-shootings-database/.

[92] Alexandra Natapoff, *Underenforcement*, 75 FORDHAM L. REV. 1715, 1775 (2006) ("Our criminal system is rife with inegalitarian enforcement failures—pervasive, yet little-noticed way that the state predictably abandons its constituents by failing to enforce the rules.").

[93] *See e.g.*, U. S. DEP'T OF JUSTICE, C.R. DIV., INVESTIGATION OF THE FERGUSON POLICE DEPARTMENT 2 (2015) ("The City budgets for sizeable increases in municipal fines and fees each year, exhorts police and court staff to deliver those revenue increases, and closely monitors whether those increase are achieved.").

[94] *See, e.g.*, Letter from Am. Civ. Liberties Union et al., to Dr. Eric S. Lander, Dir., White House Office of Sci. & Tech. Pol'y, Exec. Off. of the President, et al., Centering Civil Rights in AI Policy (July 13, 2021) (available at https://www.upturn.org/static/files/2021-07-13%20Coalition%20Letter%20to%20OSTP%20on%20Centering%20Civil%20Rights%20in%20AI%20Policy.pdf) (urging White House Office of Science & Technology Policy to "bring civil rights and racial justice to the forefront of AI policy across the board in areas beyond national security—in housing, in employment, in criminal legal issues, and more.").

[95] *See Civil Rights Principles for the Era of Big Data*, LEADERSHIP CONF. ON CIV. & HUM. RTS. (Feb. 27, 2014), https://civilrights.org/2014/02/27/civil-rights-principles-era-big-data/ (urging that "it is vitally important that these technologies be designed and used in ways that respect the values of equal opportunity and equal justice").

[96] This is the principle underlying soft law: "instruments or arrangements that create substantive expectations that are not directly enforceable, unlike 'hard law' requirements such as treaties and statutes." *See* Gary E. Marchant & Brad Allenby, *Soft Law: New Tools for Governing Emerging Technologies*, 73 BULL. ATOMIC SCIENTISTS 108, 112 (2017) (arguing that one "soft-law category of potential relevance

social values, not just legal or technocratic concerns. This section identifies four ethical commitments we can embrace in policing. These propositions are not meant to be exclusive, but rather a starting point for further development.

A. Transparency and Oversight Mean Little Without Broad Explainability

Think of this principle of the "why" of AI. We can begin with the narrower definition of explainability in AI policy discussions. Explainability refers to the idea that a person subjected to a decision or outcome informed by an AI system should be able to understand how the system works, and why a particular decision was reached in their case.[97] This specific sense of explainability matters because AI systems can be both difficult to explain and understand, and yet also have direct impacts on people's lives.[98]

We can find this call for explainability in AI policy discussions across many fields. That is because explainability can serve multiple goals, including giving users confidence in AI systems, reducing bias, meeting regulatory standards, and helping to improve the AI system itself.[99] But these differing goals mean that the requirement of explainability means different things to different audiences. For developers, explainability might include actions like publishing the algorithm or creating systems that are inherently interpretable rather than creating models that are difficult to understand.[100] For individuals facing an adverse decision made by an AI system, that might mean having the decision-making process made understandable to a layperson.[101]

For the police, explainability matters in several senses. First, there is the individual affected by an adverse decision. In other fields, that might mean the person turned down for a loan or a person who is skipped over for a job interview because of an automated decision. In

---

to many emerging technologies includes various types of private standards, guidelines, codes of conduct, and principles").

[97] *See, e.g.*, FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015).

[98] *See, e.g.*, THE ROYAL SOCIETY, EXPLAINABLE AI: THE BASICS 5 (2019) ("There has, for some time, been growing discussion in research and policy communities about the extent to which individual developing AI, or subject to an AI-enabled decision, are able to understand how AI works, and why a particular decision was reached.").

[99] *See id.* at 9-10 (discussing justifications for explainability requirement).

[100] *See id.* at 12-13 (explaining how different explainability needs require different actions).

[101] An example of this would be an explanation of why an applicant was turned for a loan through an automated process. *See id.* at 14.

policing, that adverse decision might include decisions like a purported facial recognition match, an assessment of risk during a traffic stop, or a prediction of violent behavior leading to a further investigation. Individuals routinely contest even traditional policing actions. Explainability helps people understand how an automated process came to a particular decision, whether it might contain errors, and thus provide a possible basis for contestation and appeal.

Second, the community being policed is owed a different form of explainability. The responsible use of AI in policing also requires a clear explanation of why any particular AI system is worth adoption. Why should a particular risk assessment tool, for instance, be favored over other approaches to identify persons or places in need of intervention? Why would any AI system be preferred over the existing policing approach? A dominant theory in policing studies focuses on procedural justice: that people view the police as legitimate when they have been treated with fairness and respect.[102] Legitimacy matters in this perspective because it, rather than the risk of punishment, is the basis for why people obey and follow the law. The hasty and secretive introduction of AI systems for policing can only detract from a community's perception of how fairly its police conduct themselves.

Third, there are the police themselves. Artificial intelligence married with robotics may one day lead to nearly total automation in policing. Today, though, police typically *implement* decisions suggested by AI. Whether the police receive forecasts, threat assessments, or image matches, explainability means that officers should understand how these systems work, and their limitations. Without this kind of explainability, police officers face risks. They may blindly follow the assessment of an AI system without taking further steps to verify or confirm.[103] Alternatively, they might balk at a prediction they cannot explain, and follow through with their own intuitive decision.[104]

---

[102] Tom Tyler's scholarship is most closely associated with these insights. *See, e.g.*, Jason Sunshine & Tom R. Tyler, *The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing*, 37 L. & SOC'Y REV. 513 (2003); Tom R. Tyler, *Procedural Justice, Legitimacy, and the Effective Rule of Law*, 30 CRIME & JUST. 283 (2003).

[103] CDEI, *supra* note 14, at 68 ("One possibility is that the decisionmaker over-relies on the automated output, without applying their professional judgement to the information.").

[104] *See id.* (noting possibility that a "human decision-maker [may feel] inherently uncomfortable with taking insights from an algorithm to the point where they are nervous to use it at all").

B. Fairness is Not Just the Reduction of Bias in AI Systems Used for Policing

Fairness concerns are common across many AI policy documents, and most discussions consider fairness to mean the impartial and equitable treatment of persons.[105] In its survey of more than thirty source materials from around the world, the Berkman Klein Center found that some principle of "fairness and non-discrimination" was the most highly represented theme.[106] Fairness in AI systems can mean many things, including considerations of how an AI system might visit disproportionate harms, inclusiveness in AI design, and fair representation in data sets used for training models.[107] Though they differ in detail, proposals to reduce bias in AI systems are motivated by the need to establish and increase trust and legitimacy in the public.

But the adoption of AI systems poses a unique challenge for American policing. Although the Obama administration's Twenty-First Policing Report offered hopeful predictions for the future of policing, American policing today finds itself embroiled in crises, along race, class, and political lines.[108] In this context, a narrow concept of fairness is ill-suited to AI systems in policing. Instead, the principle of fairness should consider how the AI system contributes to an improvement in the provision of policing services.

A broader view of fairness includes both attention to specific issues of bias in AI systems, as well as how these systems fit into the broader delivery of fair policing, especially to marginalized communities.[109] We can use facial recognition as an example. Much attention has been given to the high rates of erroneous matches for non-whites. A narrow view of fairness would recognize that this problem stems from the underrepresentation of non-whites in the training data of

---

[105] *See, e.g.*, Berkman Klein Report, *supra* note 17, at 49; CDEI, *supra* note 14, at 3 (noting "urgent need for the world to do better in using algorithms in the right way: to promote fairness, not undermine it").
[106] Berkman Klein Report, *supra* note 17, at 47.
[107] *See id.*
[108] *Cf.* Cynthia Lum & Daniel S. Nagin, *Reinventing American Policing*, 46 CRIME & JUST. 339, 339-340 (2017) (observing that American policing is experiencing a "tumultuous period" and suggesting that new strategies must focus on crime prevention and citizen reaction).
[109] *See* European Commission Community Research and Development Information Service, *Shaping the Ethical Dimensions of Smart Information Systems (SIS) – A European Perspective* (SHERPA) Deliverable No. 1.4, 41 (2019), https://cordis.europa.eu/project/id/786641 ("One of the reasons why the rise of datafication and algorithmic decision-making has an effect on issues of justice is its burden on predominantly poorer members of society").

a facial recognition program and would seek to address it. A broader view of fairness in policing would ask whether certain uses in the community would be unfair, even if the software's identification rates were made fairer. Even a more accurate facial recognition system used indiscriminately during traffic stops would be unlikely to satisfy broad fairness concerns.[110]

C. Privacy and Fairness Represent Different Values

Privacy and fairness are commonly used terms in discussions of ethical AI system use, but they represent different values. We can think of privacy as a form of shielding or controlling individual information from unwanted exposure.[111] It is, at its core, an individual protection. Policing scholars and civil rights advocates have focused on the harms posed by increasingly powerful and ubiquitous surveillance technologies like facial recognition, license plate readers, and a generation before that, closed-caption television cameras. They target these technologies because they collect enormous amounts of data and impact privacy and its associated individual constitutional rights, like free expression and anonymity.

Fairness, however, is different. Fairness can be a value for individuals and communities. And fairness in the use of AI systems can have multiple meanings as well. Fairness might mean that an individual subjected to, say, a facial recognition match is assured that the software has been designed and assessed to minimize bias for race, ethnicity, and gender. But fairness also means where, when, and how that facial recognition technology is used as a policing practice in the community. What is more, because fairness is a principle of police reform outside of AI tech policy, all of these forms of fairness should be compatible with one another.

And we might also imagine instances where privacy and fairness values might exist in conflict. Consider this hypothetical. When autonomous driving technology becomes widespread, should police be

---

[110] *See* Caroline Haskins, *A Popular Workshop for Police Encouraged Cops to use Face Scans to ID People They Pull Over at Traffic Stops*, BUS. INSIDER (Feb. 2, 2022), https://www.businessinsider.com/police-workshop-street-cop-training-podcast-facial-recognition-traffic-stops-2022-2 (describing police instructor advising police "to use facial recognition at traffic stops in order to find out a person's identity and if they have a warrant out for their arrest, even if it's unclear whether that person committed a crime").

[111] There is an enormous literature on privacy and the law and many definitions of privacy. *See generally* DANIEL J. SOLOVE, UNDERSTANDING PRIVACY (2010) (arguing that there is no single workable definition of privacy).

able to conduct remote traffic stops or remote enforcement? Privacy advocates might object that such stops and enforcement actions might collect unnecessarily large amounts of data, might be subject to security breaches, and might intrude upon perfectly lawful behavior. On the other hand, a remote police action vastly reduces the potential for police violence. For some, the reduction in potential violence may outweigh concerns about individual privacy. There may be other reasons communities would object to this increased automation, but this example suggests that privacy and fairness considerations do not always coincide.

D. Responsible AI Use Factors in the Nature and Degree of Private Sector Reliance

Finally, the responsible use of AI systems in policing should consider the risks inherent in privately developed tools. In the U.S., most of the AI systems used in policing are products developed by private companies.[112] Whether a predictive policing tool or a records management system, these tools are marketed to the police who are customers. Police departments may purchase these tools, but increasingly common are subscription-based models in which the public agencies never own either software or hardware.[113] Just like retail customers, police departments may be enticed by the promise of future upgrades, but these newly important relationships may strain a model of responsible policing.

These customer-vendor relationships hold the potential to pose obstacles to responsible policing. Not only is there is an algorithmic "black box" problem that makes it difficult for even developers to explain the AI systems that they have designed, there is the added complication of corporate secrecy. The invocation of trade secrets and non-disclosure agreements, and general claims of proprietary information are common in the commercial world, but unusual in traditional policing. These claims also mean that there is another layer of secrecy around these AI systems.

---

[112] *Cf.* Hannah Bloch-Wehba, *Visible Policing: Technology, Transparency, and Democratic Control,* 109 CAL. L. REV. 917, 919 (2021) (noting new police technologies are "often procured from or otherwise reliant on the private sector").

[113] Axon is increasingly focused on offering a SaaS (Software as a Service) to law enforcement agencies. Brett Schafer, *How the Company Behind TASER Guns is Becoming a SaaS Powerhouse,* MOTLEY FOOL (Mar. 3, 2021), https://www.fool.com/investing/2021/03/03/how-company-behind-taser-becoming-saas-power/.

The more that policing outsources its functions, from the development of suspicion to the most mundane information processing, the more it relies on the judgments of private companies about what responsible AI systems will do and how they will behave. In the United States, Axon Enterprise is a dominant provider of policing platforms. The public may associate Axon with the body cameras it licenses to police departments around the country, but an increasingly larger share of its revenue is invisible to the public. Police department customers pay Axon yearly recurring subscription fees for data storage and software access stored in Axon's cloud servers.[114] As police increasingly must rely on private platforms to collect, store, and analyze the information they process, they become beholden to these companies' decisions.

The need to impose public oversight and enact regulations to curb the influence of these private companies on policing has been recognized by scholars[115] and has been the subject of some local government action.[116] Framing this as an ethical concern, in addition to pushing for traditional regulatory concerns, can help communities in their oversight of their own police departments.

E. AI Systems in Policing Don't Need to End with Policing

The promise of AI systems is that we can sift through the vast amounts of digitized data to identify patterns: patterns of financial irresponsibility, ill health, job unsuitability, and crime. Even if we could successfully address the concerns raised by the current use of AI systems—bias, opacity, and so on—we would still be left with what to do with these insights. In other words, implementation is still a human decision.

Implementation too can be part of an ethical framework for the use of AI systems in policing.[117] If we can forecast crime, is the

---

[114] Dana Goodyear, *Can the Manufacturer of Tasers Provide the Answer to Police Abuse?*, THE NEW YORKER (Aug. 20, 2018), https://www.newyorker.com/magazine/2018/08/27/can-the-manufacturer-of-tasers-provide-the-answer-to-police-abuse (describing Axon as having "an iPod/iTunes opportunity—a chance to pair a hardware business with an endlessly recurring and expanding data-storage subscription plan.").

[115] *See, e.g.*, Catherine Crump, *Surveillance Policy Making by Procurement*, 91 WASH. L. REV. 1591 (2016) ("Surveillance policy making by procurement can short-circuit [the process of local control] when elected officials and the public are left without a meaningful understanding of what technologies their law enforcement agency is acquiring.").

[116] *See supra* part II (discussing local surveillance technology ordinances).

[117] *Cf.* Calo, *supra* note 19, at 412 (noting danger that AI systems can be "selectively applied to . . . marginalized populations").

responsible approach one of increased police presence? If we can identify who might be at high risk for offending or victimization, are police interventions the appropriate consideration?

Such questions speak to a broader audience than those engaged in AI policy. The movement to "abolish the police" is a reaction to distrust and to the call for social solutions beyond traditional law enforcement. Asking mental health specialists to respond to mental health crises is a way of responding to these concerns. So too is asking whether the assessments of AI systems in policing should be met with novel responses rather than traditional police investigations.

CONCLUSION

AI systems are everywhere. Most people are used to them in their daily lives, and they are increasingly important decision-making mechanisms in social services, healthcare, finance, and criminal justice. In this sense, the use of AI systems in policing is part of a larger social transformation.

And just as in many of these fields, the regulation and oversight of AI systems in policing is woefully inadequate. We have no real national standards in the United States. Existing efforts are piecemeal and slow going. One way to address this gap is to introduce ethical principles. Many non-profits and governmental bodies around the world are in the process of drafting ethical guidelines. These guidance documents are not binding or enforceable, but they are far preferable to no standards at all.

The use of AI in policing stands at the intersection of two distinct discussions: the widely acknowledged need for ethical principles in the use of AI systems, and the renewed attention to inequality and bias in American policing. Just as in lending, employment, and healthcare, the use of AI systems in policing needs not just greater regulation, but also a set of principles to guide their use with responsibility. In this way, ethical considerations can contribute to the larger project of police reform and even conversations about envisioning policing entirely differently.