AI Ethics in Predictive Policing

WIKIMEDIA/BRETT GUSTAFSON

From Models of Threat to an Ethics of Care

Peter M. Asaro

Digital Object Identifier 10.1109/MTS.2019.2915154
Date of publication: 30 May 2019

The adoption of data-driven organizational management — which includes big data, machine learning, and artificial intelligence (AI) techniques — is growing rapidly across all sectors of the knowledge economy. There is little doubt that the collection, dissemination, analysis, and use of data in government policy formation, strategic planning, decision execution, and the daily performance of duties can improve the functioning of government and the performance of public services. This is as true for law enforcement as any other government service.

Significant concerns have been raised, however, around the use of data-driven algorithms in policing, law enforcement, and judicial proceedings. This includes predictive policing — the use of historic crime data to identify individuals or geographic areas with elevated risks for future crimes, in order to target them for increased policing. Predictive policing has been controversial for multiple reasons, including questions of prejudice and precrime[1] and effectively treating people as guilty of (future) crimes for acts they have not yet committed and may never commit. This central controversy over prejudice and precrime is amplified and exacerbated by concerns over the implicit biases contained in historic data sets, and the obvious implications for racial, gendered, ethnic, religious, class, age, disability, and other forms of discriminatory policing, as well as how the use of predictive information systems shapes the psychology and behavior of police officers.

As more bureaucratic processes are automated, there are growing concerns over the fairness, accountability, and transparency of the algorithms used to make consequential decisions that determine peoples' life opportunities and rights. Less discussed are the ways in which the introduction of data-centric processes and data-driven management have significant consequences on the techno-social and spatio-temporal structure of organizations (1), as well as on the priorities of organization management, the nature of labor, and the quality of results (2). Such is the nature of contemporary technocratic governance (3). Yet neither the increasing collection and reliance on data, nor specific socio-technical and spatio-temporal organization of governmental institutions is determined by the technology alone, nor by the utility of data. Nor is the kind of analysis performed on that data, or the specific problems to which the data is addressed, pre-determined or "natural" in any meaningful sense. Rather, there are myriad social, institutional, and individual values that go into the decisions of which data to collect, when and where to collect it, how to encode it, how to assemble it in databases, how to interpret it, and how to use it to address social, institutional, and individual concerns. It is those values which are the primary concern of ethics in information systems design.

This article outlines a new ethical approach that balances the promising benefits of AI with the realities of how information technologies and AI algorithms are actually adopted, applied, and used. It proposes that AI ethics should be driven by a substantive and systemic Ethics of Care, rather than by narrow Models of Threat based on utilitarian risk and threat models. While it focuses on law enforcement policies and policing practices, it hopes to contribute to the broader discussion over the ethical application of AI technologies in government policy-making and the delivery of public and commercial services more generally. The paper concludes that while data-driven AI techniques could have many socially beneficial applications, actually realizing those benefits requires careful consideration of how systems are embedded in, and shape, existing practices, beyond questions of de-biasing data. Absent such consideration, most applications are likely to have unjust, prejudicial, and discriminatory consequences. This conclusion supports a proposed Ethics of Care in the application of AI, which demands moral attention to those who may be negatively impacted by the use of technology.

Recent Excitement about AI

There is a recent and widespread excitement about the application of artificial intelligence to nearly every aspect of society — from commerce to government. AI, as a scientific research field, has long sought to develop computer programs to perform tasks that were previously thought to require human intelligence. This somewhat abstract and conditional definition has given rise to a wide array of computational techniques, from logical inference to statistical machine learning, that enable computers to process large and complex datasets and quickly provide useful information. Whether through traversing long chains of inference or sifting through vast amounts of data to find patterns, AI aims to provide logically sound and evidence-based insights into datasets. Insofar as these datasets accurately represent phenomena in the world, such AI techniques can potentially provide useful tools for analyzing that data and choosing intelligent actions in response to that analysis, all with far less human labor and effort. This is the traditional approach of AI, or what we might consider artificial specialized intelligence. This type of AI is essentially about creating a customized piece of

[1]"Precrime" is a science fiction concept that first appeared in the writings of Philip K. Dick in a novel (19) that was later turned into a major Hollywood movie (20).

software to address a complex issue or solve a specific problem by automating what would otherwise require human mental effort.[2]

Specialized AI is best seen as an extension of more traditional practices such as software engineering, IT systems design, database management and data science which deploys a range of AI techniques to automate the search for solutions to problems that currently require substantial human mental labor and skill. Much of the current excitement around AI is focused on "deep learning" machine learning techniques that use many-layered "deep" neural networks that can find complex patterns in large datasets ("big data"). Far from artificial sentience, consciousness or general intelligence, we could consider this as enthusiasm for "statistics on steroids." Commercial and governmental institutions have long used statistics to develop representations of the world that can inform future actions and policies. In this sense, the AI revolution is really a continuation, and massive acceleration, of much longer and older trends of datafication and computerization. What is new and unprecedented is the sheer volume of data, the speed at which it can now be effectively processed, the sophistication of the analysis of that data, the degree of automation and the consequent lack of direct human oversight that is possible.

As data-driven organizational management — led by big data, machine learning and AI techniques — continues to accelerate, and more processes are automated, there are growing concerns over the social and ethical implications of this transformation. Machine ethics is concerned with how autonomous systems can be imbued with ethical values. "AI ethics" considers both designing AI to explicitly recognize and solve ethical problems, and the implicit values and ethics of implementing various AI applications and making automated decisions with ethical consequences. This paper will consider the latter, implicit view that corresponds to what is sometimes called "robot ethics," to distinguish it from explicit "machine ethics" [4]. Ideally, the explicit ethics, implicit ethics, and the embedding and regulation of the system in society should all align [5].

The outputs of predictive policing algorithms clearly have ethical consequences, even if the systems under consideration do not try to design systems for explicit ethical reasoning. In the predictive policing systems under consideration, there is little or no effort to design the systems to frame their analysis or results as ethical decisions or perform ethical analyses. What is of concern to the public, and in this paper, is how well the systems are designed, and the ethical implications of introducing them into police practices.

There is a growing body of research examining the ways in which data-driven algorithms are being used in an increasing number of critical decision processes, often with little or no accountability [6]–[9], and sometimes with little or no real understanding of how they function in the real world or why they reach the results they do in particular cases [10]–[12]. Consequently, there are many ways for such systems to "go wrong." Sometimes this is due to a well-intentioned but mathematically naive understanding of how such systems work. This includes the failure to understand how statistical outliers may be mishandled or misrepresented, or how historical data patterns can be self-reinforcing — such as denying credit and charging higher interest rates to poorer individuals and communities, thus systematically denying them opportunities to escape poverty. Sometimes this is due to the intended desire to transfer responsibility and blame to an automated process, and relieve human agents of their responsibility. And sometimes there may be malevolent motives behind using data in obviously discriminatory ways — such as purging voter rolls to deny eligible voters to an opposing political party. But these are ultimately "narrow" views of AI ethics, which look to improving accuracy and performance of the technology, while largely ignoring the context of use. It has also been argued that the focus of AI ethics on "solving" the bias problem is a distraction from other and more important ethical and social issues [13]. Without discounting the value of such narrow approaches, this paper will examine the importance of taking a broader ethical perspective on AI, and the problems that will not be fixed through fairness, accountability and transparency alone.

Two Approaches to AI Ethics

This paper aims to go beyond the ways in which data and AI algorithms might be biased or unaccountable, and consider the ethics of how AI systems are embedded in social practices. Because AI ostensibly automates various forms of human reasoning, consideration, and judgement, the accuracy or fairness of such processes alone do not guarantee that their use will provide just, ethical, and socially desirable results. Rather, careful attention must be paid to the ways in which the implementation of such systems changes the practices of those who use them. In order to redirect attention to the bigger picture of the socio-technical embeddedness of AI when considering ethics, the paper will formulate two broad concepts of AI ethics, which will be named

[2]Some theorists have speculated about the possibility or consequences of an artificial general intelligence (AGI) which might be able to learn with little or no direct instruction from humans, and in some sense recognize problems on its own that are in need of solution, and then adapt itself to solve them. AGI is not technologically feasible for the foreseeable future, and as such it will not be given much consideration here.

"Models of Threat" and an "Ethics of Care."[3] It will first outline these concepts in broad terms. It will then examine two illustrative cases, in the area of predictive policing, which epitomize each approach. It concludes with some observations and reflections on how to design better and more ethical AI through an Ethics of Care approach.

Perhaps the greatest ethical concerns over algorithmic decisions have been raised around the use of data-driven algorithms in policing, law enforcement, and judicial proceedings. One well-researched and much discussed example from the Florida judicial system involves the use of algorithms to predict future recidivism in convicts as a basis for determining the length of their sentences.[4] Another growing application is predictive policing — the use of historic crime data to identify individuals or geographic areas with elevated risks for future crimes, in order to target them for increased policing. Predictive policing has been controversial — as it aspires to prevent crime, it also raises questions of prejudice and precrime and effectively treating individuals and communities as guilty of (future) crimes for acts they have not yet committed and may never commit (21), (22). This central controversy of prejudice and precrime is amplified and exacerbated by more general concerns over the implicit biases contained in historic data sets, and the obvious implications for racial, gendered, ethnic, religious, class, age, disability, and other forms of discriminatory policing.

Predictive policing as a term can refer to a variety of technologies and practices. The technical usage of the term usually refers to algorithmic processes for predicting locations or individuals with high probabilities of being involved in future crime, based upon historical data patterns (23). Recent approaches utilize "big data" techniques and arguably entail forms of mass surveillance of the public (24). However, these recent algorithmic techniques and applications have their roots in much older practices of collecting and utilizing comparative statistics (better known as CompStat) about crimes to manage large police forces, which began in New York City in 1995. While many CompStat programs utilized computer programs to calculate the statistics from crime and accident reports and arrest records and in some cases automatically generate "pin-maps" of crime activity, CompStat was really a set of data collection, analysis, and management practices rather than a piece of software (25). And CompStat has seen its share of criticism, including from former police officers (26).

Moreover, the algorithmic techniques that are increasingly being employed by police forces draw upon data that goes well beyond the digitized crime reports of the CompStat legacy, or automatically generated "heat maps" of areas of high crime activity.[5] In recent years, police departments have begun deploying and integrating large-scale video surveillance systems, traffic cameras, license-plate and face recognition technologies, audio gun-shot locators, cellphone interceptors, aerial surveillance, and a host of other surveillance and data-collection technologies. As these systems become networked and produce large amounts of data, there is increased pressure to analyze, integrate, and utilize this data for improving law enforcement, which leads to increased reliance on automation and algorithms for sorting and sifting through that data and translating it into policing priorities and strategies. As such, the term predictive policing can be taken to refer to a broad class of algorithmic and data-driven practices and software tools utilized by police forces. Predictive policing is also a good example of how AI might be deployed more generally, and the ethical challenges that may arise.

A general approach to AI ethics is characterized here as an "Ethics of Care." Ethics of Care uses predictive policing, and the design of AI-based systems within it, to lay out the framework for an AI Ethics of Care. In particular we look at two recent, but very different, implementations of data-driven interventions on youth gun violence in Chicago, Illinois, U.S.A. Predictive policing is particularly good for this purpose for several reasons. As should be clear from the discussion above, policing is an area the gives rise to a number of critical ethical and legal issues, and has relevance not only to society at large, but to a host of other governmental functions and other industries. It is also an area that has an historical practice of data collection, and recent trials in

[3]Neither term is original, and each is meant to evoke traditions of thought and their general perspective, while not necessarily implying that the specific projects described were conscious of, or directly influenced by, those traditions. "Threat Modeling" has been an important methodology in cybersecurity for identifying, assessing, prioritizing, and mitigating threats and vulnerabilities since at least the early 2000s (14), while "Threat Perception" has been a key concept in international relations and political psychology in assessing military threats and deterrence strategies (15). "Ethics of Care" has been gaining popularity in medical and educational ethics since its introduction by Carol Gilligan to explain moral development in child psychology in the late 1970s and its extension by Nel Noddings into a moral theory based on interpersonal relationships of caregiving and receiving in the early 1980s (16).

[4]In an analysis of 7000 sentencing cases in Broward County, Florida, over the period 2012-2013 that used the COMPAS software, journalists found similar error rates in the assessment and sentencing of white and black convicts, but diametrically opposed in their direction. White convicts were more likely to be erroneously predicted *not to* commit future crimes, while black convicts were more likely to be erroneously predicted *to* commit future crimes, resulting in shorter sentences for white convicts and longer sentences for black convicts (17).

Another study of the same dataset shows that amateur humans are able to make better predictions than the COMPAS software, using the same six factors as the software, and even better predictions can be made using just two factors — defendant's age and number of past convictions (18).

[5]Such "heat maps" have become ubiquitous in the age of big data, and are even reproduced, albeit at lower resolution, on real estate websites such Trulia.com (27).

the application of AI techniques to those practices. Further the algorithms of predictive policing embed values and make designations and decisions with implicit ethical consequences.

The Ethics of Care approach has a history of its own as well, and is similar in some ways to concepts in related fields, including the "Duty to Protect" in policing (28) and the "Duty of Care" in law (29). In contrast, the Models of Threat approach construes the world and the individuals within it as risks and threats which must be managed, mitigated, and eliminated. The later discussion section will consider what it means to implement the Ethics of Care approach, following the examples. First we give a brief sketch of each approach.

The Models of Threat approach begins from the assumption that the world can be classified into clear categories, i.e, threats and non-threats, and that this is the first step in choosing an appropriate action to take.[6] It focuses on capturing and processing increasing amounts and types of data, and processing this data to provide increasingly accurate classifiers of what constitutes a threat, and predictors of the likelihood and risk from that threat. It largely assumes that the actions that will be taken to address threats and risks are independent of the observation, collection, and analysis of data. This approach also assumes that the primary values are in the accuracy, precision, fidelity, and comprehensiveness of the data model, and in the correctness of its classifications and reliability of its predictions. This approach could also be characterized as taking a narrow view, being very detail oriented, atomistic, and deeply analytic.

By contrast, the Ethics of Care approach is holistic, and takes a broad, big-picture view of the values and goals of systems design. It considers the interaction and interrelation between an action or intervention and the nature of classifying things and predicting outcomes within specific contexts. The goals and values of an Ethics of Care approach is to benefit everyone represented by the system as well as those who use the system, and the society as a whole. The Ethics of Care approach recognizes the complexity of social relations and sociotechnical systems, including the organization using the system, and does not expect more and better data to simply solve complex social and institutional problems, but rather to provide opportunities for finding better solutions, better actions, and better policies than what are already considered.

The traditional notion of Ethics of Care is that interpersonal relationships form the basis for normativity, and should be guided by benevolence (16).[7] When it comes to law enforcement, we can see the Models of Threat approach seeking to better identify violations of the law, and to predict when and where violations will occur, so as to better deploy police officers to respond. It might also aim to assist police in identifying perpetrators and bringing them to justice. The Ethics of Care approach, might instead consider the factors that lead people to violate the law, and seek out new interventions that make crimes less likely, thus requiring fewer resources to enforce the law. It would also view the relationship between law enforcement and the community as primary and consider how any new data tool might impact that relationship.

A Note on "Precrime"

Beyond the practical socio-technical meanings of predictive policing, there is also a deeply troubling connotation to the term, captured in the concept of "precrime." This notion is more philosophical in nature, and draws upon our concepts of guilt, responsibility, agency, causality, and their temporality, as well as the means and ultimate aims of law enforcement in the regulation of society. The term is also mentioned extensively by nearly every press article about predictive policing, and the commercial software startup PredPol, which supplies Los Angeles and many other police departments with data analysis software, states prominently on their "About" page that they are not selling "Minority Report" technology (30). Yet, the notion of precrime has powerful cultural meanings for good reasons beyond the popularity of sci-fi.

The basic idea of precrime stems from the idea that the goal of policing is the reduction and, ultimately, the elimination of crime altogether. While investigating crimes after they occur and responding to crimes-in-action are good, it would be even better to prevent crimes before they happen, or so this line of thinking goes. This view tends to emphasize deterrence over other key elements of criminal justice — retribution and reformation. The goal is to disrupt or dissuade criminality before it manifests. While crime prevention could focus on eliminating the means of committing

[6]This is not to say that the world, or its representation in a computational model, is necessarily discrete. One could represent the likelihood that an individual or area might present a threat or risk as a continuous variable. And while the scale and threshold for action on the basis of that variable might not be predetermined, or determined by the system, it is expected that such metrics will influence the decisions and actions of police officers with respect to those individuals and areas — i.e., that the threat or risk represented by the calculation can and should result in actions.

[7]According to the *Internet Encyclopedia of Philosophy*, "Normatively, care ethics seeks to maintain relationships by contextualizing and promoting the wellbeing of caregivers and care receivers in a network of social relations. Most often defined as a practice or virtue rather than a theory as such, "care" involves maintaining the world of, and meeting the needs of, our self and others. It builds on the motivation to care for those who are dependent and vulnerable, and it is inspired by both memories of being cared for and the idealizations of self. Following in the sentimentalist tradition of moral theory, care ethics affirms the importance of caring motivation, emotion and the body in moral deliberation, as well as reasoning from particulars." (16).

crimes,[8] it more often focuses on motives, and as such employs psychological theories of choice and sociological theories of behavior, and generally focuses on maximizing the likelihood and cost of penalties for wrongdoing by stricter enforcement and harsher penalties.[9] The temporality also becomes deeply problematic here. There is an obvious utility in preventing crimes before they occur, but our notions of individual responsibility, guilt, and punishment rest on the commission of acts — of actually doing certain things that constitute crimes — rather than imagining, desiring, or simply being psychologically pre-disposed or circumstantially inclined toward doing things which would be criminal. In some instances, planning or discussing criminal acts with others are acts that can themselves constitute a lesser crime, such as conspiracy or solicitation to commit a crime, and a failed attempt, e.g., to kill someone, can still constitute the crime of attempted murder even if nobody is actually hurt. But there are, and should be, different standards for citizens who have committed no crime, those in the act of committing a crime, those suspected of a crime, those convicted of a crime, and those who have served their sentences for a crime. How should law enforcement treat "those 'likely' to commit a crime"? And does the epistemic basis for that likelihood determination matter?

The classification of individuals also becomes critical here. When we say that an individual is "likely to commit a crime" is that based on their individual behavior and actions, or because of membership in a certain demographic group? "Profiling" becomes problematic in the latter case, when individuals are classified according to population-level statistics and biases. Statistics are notorious for not distinguishing correlations in data from causal reasons, and it would be unjust to treat people with suspicion for coincidental correlations when the underlying causal mechanisms for criminal behavior are absent. This kind of profiling becomes deeply problematic when it becomes prejudicial, and the correlation is taken as itself constitutive of guilt, or warranting a presumption of guilt, rather than a presumption of innocence.[10]

According to the U.S. legal system, criminal liability and guilt depend upon a combination of *actus reus* (the "guilty act") and *mens rea* ("the guilty mind"). That is, one must actually commit the act for which one is held responsible, and one must have had in mind the intention, or at least the awareness, that one was doing something wrong, or should have known (as mere ignorance of the law is not a suitable defense). From this perspective, one cannot be guilty of a crime before actually committing the act, and should not be held liable for a crime not committed. And this is where precrime clashes with fundamental concepts of justice. If society, and police, act upon precrimes, and those suspected of them, in the same way as already committed crimes, then they are treating *as guilty*, or at the very least *as suspect*, those who have not yet, and not actually, committed a crime. This is a profound form of prejudice, in which judgments are made not only before relevant evidence of a criminal act can be obtained and analyzed, but before such evidence can even exist. Rather, judgement is passed on information derived from statistical inference, patterns, trends and probabilities. But a statistical likelihood of an event is neither an event nor an act.[11] And it is fundamentally unjust to treat someone as guilty of a crime they did not commit. Moreover, it is powerfully felt as an injustice when individuals and communities are treated "as if" they are guilty of doing something they have not yet, or not individually, done, based simply on their being members of a category or demographic group. Indeed, the imposition of social categories can even give rise to new social identities (35) — and thus machine-generated categories are likely to create new types of people. This makes the creation and designation of a "criminal type" deeply problematic.

Still, there is a practical concern that law enforcement cannot ignore information about likely crimes without sacrificing their duty to prevent crime. While the scope and nature of that duty are themselves contested, this is a powerful intuition. Indeed, it is the same intuition that motivates much data-driven management. That is, if we *can* use historical data to predict future trends and events, and thus better allocate valuable resources towards fulfilling a mission or goal, then we *should* do so. While not incorrect — certainly better use of information can improve policing in many ways — if pursued without careful consideration, caution, and

[8]For instance, adding better locks to protect property, such as ignition immobilizers on cars, or making it more difficult to resell stolen goods (31). In some cases, increasing the policing of crimes may actually have counter-intuitive effects of increasing crime, according to an economic analysis of the theft of art works (32).

[9]Rarely do these approaches take into account the outright irrationality or the failure of individuals to actually think about committing crimes in rational terms. This is because cognition in the wild follows other lines of reason and risk assessment, from inflamed passions, to rational biases, to human necessity.

[10]For example, if one is worried about a copycat bombing like the Boston Marathon bombing, it might make sense to flag individuals who shop for pressure cookers and backpacks. However, one should still presume there is a reasonable explanation for this rather than presuming they must be terrorists for doing so (32).

[11]Just consider gambling on horse races, which historically gave rise to modern statistics (33). Oddsmakers go to great lengths to provide accurate statistical predictions of the chances for each horse in a race. Yet, whichever horse is the favorite to win does not necessarily win — the actual outcome of the race matters. The favorite only wins about 1/3 of the time (34). Gambling would not make sense if this were not the case — though in many games of chance it can be argued that it is mathematically irrational to place bets at all.

sensitivity to its various implications and specific implementations, pursuing such intuitions blindly can quickly lead to problems. Unfortunately, the strength of this intuition and its simple logic make it an easy policy argument to make in many institutional and bureaucratic settings. One might even argue that this is the "default" policy argument in the age of data, and thus Models of Threat is the default approach to predictive policing. And it is safe to assume that without critical reflection and active awareness on the part of systems designers, something similar will be the likely default goal of most AI systems. To better understand how the design of systems can mitigate or exacerbate the problems inherent in data-driven management, we now turn to two examples of predictive policing.

One City, Two Cases of Predictive Policing

The City of Chicago, IL, has seen a spike in gun violence in recent years. The city has led the United States in the number of shootings and gun homicides, peaking with 758 total homicides and more than 4300 shootings in 2016, and down slightly in 2017 [36]. This has led to a serious effort by the Chicago Police Department (CPD) to address this spike by focusing on neighborhoods and individuals most likely to become involved in gun violence. A number of studies, experiments, and policies have been tested and implemented in recent years. By comparing different applications of data-driven interventions occurring in the same city at the same time period, we can develop insights into the implications of data for shaping policing practices.

Two such experiments, in particular, offer a good insight into the ways in which data can be applied to address gun violence, and also into the ways that the implementation and utilization of those insights can have radically different social and ethical implications. One has been the subject of critical scrutiny by journalists and researchers, called the Strategic Subjects List. More often called the "heat list" by police officers, it was first used by CPD in 2012, and its use continues, though under a revised set of guidelines following criticism of the early uses described here. The other started in the summer of 2011 as a pilot research program implemented by the City of Chicago, and was studied the following year by University of Chicago researchers. Called One Summer, it has since been adopted as an annual program by the City of Chicago. While both started out as academic research projects, both were analyzed by outside researchers in 2012, and both utilized data to assess and identify youth who are at-risk of being involved in gun violence, in most other ways the two programs are very different.

The two projects can best be characterized as illustrative case studies, embodying two different philosophies of predictive policing, and perhaps two extremes thereof. They accordingly have very different ways of thinking about what being an "at-risk" youth means, and consequently pursue very different approaches to intervening so as to reduce that risk. More importantly, they also had very different outcomes in terms of their effectiveness in reducing gun violence and in influencing the life outcomes for those identified as "at-risk" in each program. In short, the Strategic Subjects List can be described as taking a "Models of Threat" approach to at-risk youth. That is, at-risk youth in that project are primarily viewed as threats to the community because they are at-risk, and interventions are targeted at increased police scrutiny and enforcement against those individuals. Whereas the One Summer program takes an "Ethics of Care" approach to at-risk youth, in which at-risk youth are given access to social services and resources aimed at reducing their risks of becoming involved in violence.[12] Like their philosophies, their outcomes were also dramatically different, despite resting on similar data-driven assessments of being "at-risk."

The Heat List

The Strategic Subject List (SSL) algorithm was developed as an experiment by a researcher at the Illinois Institute of Technology, and was utilized by CPD starting in 2012 and continuing until today. In its early iterations and implementations, it took data about individuals from CPD arrest records, taking into account some 48 factors, including number of arrests, convictions, drug arrests, gang affiliations, and being the victim of crimes or violence [38]. The SSL then went further, taking into account these factors for the individual's social network as determined by who was arrested together with an individual [39]. These factors were weighted and compiled into an overall SSL score from 1–500. The initial implementation contained over 398 000 individuals drawn from police arrest records, and identified 1400 as being at "high-risk" of being involved in violence. While some 258 received the top score of 500 points, only 48% of these had previously been arrested for a gun crime, and many people on the list had never themselves been arrested, but rather were victims or were in the social networks of victims or perpetrators [39]. Many police officers reported that they were not fully informed of how the list was compiled. They assumed, or were led to believe, that everyone on the list was a perpetrator of violence and was likely to commit more violence, whereas the SSL scores combined those at-risk of being victims with those at-risk of being perpetrators in a single metric of "being involved in violence."

[12]The slogan of the One Summer program is "Nothing Stops a Bullet Like a Job" [37].

The practical use of the SSL list and scores was somewhat haphazard in its early years.[13] While there was no official policy regarding its use, it did feature in some CompStat reports (40), and was used by police officers in some more controversial ways. The first of these, called "custom notification," involved police officers making personal visits to high-risk individuals, informing them of their presence on the list, and further, informing them that they would be subjected to additional police scrutiny (41). In other words, they were told that the police were "watching them" more carefully, and they should expect more police encounters. The other, and more common use of the SSL was as a "heat list" following a violent crime, in order to round-up the "usual suspects" from the list for questioning, in this case people in the vicinity of the crime who had high scores on the list. As a result, people on the list were far more likely to be detained and arrested by police, simply for being on the list. A detailed RAND study showed that the use of heat list in this way had no statistical impact on the likelihood of individuals on the list being involved in gun violence, nor on the overall gun violence in their communities (42). It did, however, radically increase the likelihood of being arrested and convicted of a crime for those people on the list.

Further, the data and algorithm behind the SSL was not shared publicly, making it difficult to determine whether the list simply replicated long-standing racial and class discrimination. The CPD told the *Chicago Tribune* that,

> "(The SSL) is not based on race, ethnicity or geographic location..We don't use it to target certain individuals other than we pay a visit to their residence to offer them services to get out of the (gang)."

But a California-based group that defends civil liberties in the digital world raised concern that the arrest data that goes into the SSL could be inherently biased against African-American and other minorities:

> "Until they show us the algorithm and the exhaustive factors of what goes into the algorithm, the public should be concerned about whether the program further replicates racial disparities in the criminal justice system,"

said Adam Schwartz, a staff attorney for the Electronic Frontier Foundation (41).

The Ethics of Care approach might instead consider factors that lead people to violate the law, and seek out new interventions that make crimes less likely.

That same *Chicago Tribune* article indicates that 85% of the 2100 shooting victims so far that year had been on the SSL, but does not indicate how they scored or whether they were all in the list of 1400 high-risk individuals, or the longer list of 398000 individuals included in the dataset.

Both of the main applications of the SSL, the "custom notification" warnings and using the "heat list" to bring people in for questioning, contain elements of precrime. In the warnings, there is a sense in which the police still cannot arrest an individual before a crime, but they do attempt to intimidate and threaten an individual who, in the majority of cases, has never been arrested for a violent crime. While the police do offer to "help individuals to leave gangs," it is not clear what specific services they offered, or whether those services are effective in either helping individuals get out of gangs or in avoiding future violence. Similarly, rounding up people in the area who appear on the "heat list" may be an expedient tool, but it is no substitute for doing the policework of a real investigation, or following leads from witnesses and suspects. Indeed, it may impede or undermine community-oriented policing strategies. While police may complain that witnesses, and even victims, are often unwilling to cooperate with police, these heavy-handed tactics of rounding up suspects based on data-driven lists only further breaks down trust between communities and the police. As such, these uses of SSL actually work against confidence-building efforts by police, while offering little or no demonstrative positive results (42), (43).

Both applications also appear to engage in victim-blaming. In some cases literally so, insofar as the SSL combines victims and perpetrators in a single category of "being a party to violence" or at-risk of being "involved in violence." It makes little sense to show up at someone's door to tell them that they may be the victims of violence,[14] and less sense to threaten them with

[13]It is also worth noting that the SSL, and the data and algorithms upon which it was based, was kept private by the CPD. It was only after a long legal battle that the *Chicago Sun-Times* newspaper was able to force the CPD to make the SSL and its data public (39).

[14]Making someone aware of a specific threat against them would be helpful, but people are usually aware of the fact that they live in a violent neighborhood. Nonspecific warnings are of little help, as has been seen with color-coded threat risks from the Department of Homeland Security, which do not specify any particular location or type of activity to be on the lookout for.

increased surveillance, or to round them up for questioning after a violent crime. Detailed analysis of the effects of these practices bear out the futility of these interventions. Accordingly, this approach can best be characterized as "Models of Threat." Individuals on the SSL are seen as threats, and are themselves threatened and subjected to additional police attention, and are much more likely to be questioned and arrested. Indeed, from a crime statistics perspective, the success of a police department rests on the number of violent crimes, and many gun crimes are the result of and/or give rise to retaliation, so it makes sense to combine the victims and perpetrators of violence in a single metric. In other words, individuals likely to be involved in violence are a "threat" to the department's CompStat numbers, regardless of whether they are victims or perpetrators. Thus, in a Models of Threat approach, even a victim is viewed as a "threat." Yet, in any commonsense approach to violence there should be a difference in how one approaches or intervenes with an individual who is likely to be a victim from someone likely to be a perpetrator.[15] It would be difficult to argue this approach has improved policing — for instance by making police work more efficient according to its own metrics — when it has been proven to have no effect on violent crime on either an individual or community level. And while conflating victims and perpetrators is poor data practice, it is not clear that "getting the data right" would actually improve the results of SSL. It is hoped that an AI ethic would be able to avoid such ineffectual and counterproductive applications. But to do so, it must look beyond the numbers and datasets, to understand how data and information systems are embedded in communities and policing practices.

Nothing Stops a Bullet Like a Job

The Ethics of Care approach offers a stark contrast to the Models of Threat. One Summer started as a pilot program in the summer of 2011 by the City of Chicago. In 2012 it became part of a controlled study (One Summer Plus) by researchers at the University of Chicago Crime Lab. The basic idea was to intervene with at-risk youth by providing them with summer jobs, for 8 weeks and 25 hours a week at minimum wage, mostly working for organizations focused on their local communities. According to the City's press release about the program, "at-risk" was defined by a combination of attending an at-risk school and a review of individual applications as follows.

> More than 700 youth ages 14–21 were selected to participate in One Summer Plus in 2012 from an open application process available at thirteen Chicago public schools located in high-violence and low-income neighborhoods. Applicants faced a number of challenges; the year before they entered the program, they had missed an average of six weeks of school and about 20 percent had been arrested (44).

As a data-driven technique, it was largely the schools that were identified through historical data. While the methodology used to identify the 13 schools is not discussed in detail, presumably it was based on the geographic location of historical incidence of violence, and the proximity of those schools to violent areas, in combination with demographic income data. But it is important to note that individual students were initially identified only by virtue of attending a designated school. The accepted applicants may have been further screened for factors such as school attendance, previous arrests, or other factors. But it is worth noting that this was not a highly sophisticated data-driven technique for identifying which individual youth were "at-risk." As far as the program was concerned, anyone living in a low-income, high-violence area was "at-risk," and more detailed or nuanced classifications were not essential to participation or effectiveness.

Researchers studying One Summer found a 51% reduction in involvement in violence-related arrests among youth who participated in the program compared to the control group that did not participate.[16] Their analysis of the data from the initial study, and of subsequent years, demonstrates that this was not simply the result of getting them off the streets for 25 hours per week, but that there were significant changes in their cognitive and behavioral approaches to school, work and becoming involved in violence (46). Much of this was attributed to improved impulse control, learned both through their employment and through training sessions they received as part of the program. There were also economic benefits resulting from the additional income received by the participants and their families, and participants were much more likely to seek and get jobs after participating in the program.

The One Summer program provides a good illustration of an Ethics of Care approach insofar as it focuses on the contextual manifestations of violence, and seeks a means of directly intervening to change that context. Rather than focusing on the metric or individual "threat," Ethics of Care focuses on the system. Ethics of

[15]The assumption made by researchers in doing this appears to be that there is significant overlap in the categories of victims and perpetrators. This is especially true given the cyclical nature of gun violence in Chicago, driven by rivalries and revenge killings that beget further revenge killings. Still, associating with people connected to violence might make you more likely to become a victim of violence without becoming more likely to commit violence.

[16]Subsequent research places the figure at a 43% reduction in violent arrests (45).

Care also starts from respecting people and maintains a focus on the duties and responsibilities to the individuals it deals with. By contrast, a Models of Threat approach sees people as statistics, and treats the individuals on a list as threats, whether they have done anything or not, and regardless of whether they are victims or perpetrators–thereby undermining their humanity. Ethics of Care sees the individual as having rights and deserving of respect, and sees those at risk as being in need of care. An Ethics of Care does not disregard data, but rather utilizes data in the service of performing a duty in a manner that respects everyone involved. That respect extends to taking the effort and care to understand a situation from multiple perspectives, including that of citizens and working police — and how data gets used and how it relates to the lived world. Indeed, as the RAND researcher who studied the SSL says, data and AI ethics is less about sophisticated data analysis techniques and more about understanding context:

> The biggest issue for those agencies considering predictive policing is not the statistical model or tool used to make forecasts. Getting predictions that are somewhat reasonable in identifying where or who is at greater risk of crime is fairly easy. Instead, agencies should be most concerned about what they plan to do as a result (47).

There is a deeper lesson in this observation — the possibility of action, and the types of interventions envisioned, can strongly shape data representations, and the value of various kinds of data. While the current fashion is to collect any and all available data, in the hope that something useful might be inferable from it, there is still value in considering what actions are available to address a problem. This also means using data to find new means of acting and intervening, and better understanding the problem, rather than simply making the current means of addressing a problem more efficient. Indeed, many AI ethicists concerned about AGI worry that a hyper-efficient AGI might be so good at achieving a set goal, or maximizing a certain value, that it does so to the great detriment of other human values.[17] In the case of policing, many current policies and tactical goals of policing could be dangerous, unjust, and counter-productive if executed with complete accuracy and efficiency. And most people would not be happy living in a society where every violation of the law was detected and punished strictly and with perfect efficiency. At least this would require rethinking many laws, policies and punishments (48). In order to better appreciate how actions and practice could or should shape data, particularly for AI ethics, we turn now to a discussion of what the framework for AI ethics drawn from an Ethics of Care would look like.

[17]Nick Bostrum's infamous paperclip maximizer, which quickly and efficiently turns the world into paperclips at the expense of everyone and everything else, is an example of this.

AI Ethics of Care: From Data to Models to Implementation

The Ethics of Care has its own history, coming out of feminist thought. As a general normative theory, it has been criticized for failing to question what is right to do, in favor of seeking what is best to do in the circumstances. But as an approach to practical applied ethics, it has proven illuminating in areas such as educational and healthcare ethics (49), (50). It is proposed that policing, like education and healthcare, aims to "serve and protect" the community with limited resources,[18] and as such is also a good candidate for an Ethics of Care. It is further proposed that in trying to improve the management of a broad variety of governmental, non-profit and commercial organizations with data-driven techniques, AI ethics can also draw upon the Ethics of Care, as robot ethics has done (53). In this section we look at how an Ethics of Care can be applied to data science and AI, from data collection, to data modeling, to

[18]The motto of the Los Angeles Police Department, "To Protect and To Serve," was introduced in1955 following a contest at their police academy, won by Officer Joseph S. Dorobek (28). It, and its variants, have since been adopted as the motto of numerous police departments across the United States. But what do these words really mean? The topic has been much discussed within police departments. In 1998, an Ohio police officer offered his views in *Police Magazine*:

> While what constitutes "protect" may be open to some debate, it seems to be more clear-cut than does the word "serve." It's obvious that we protect the citizens and their property from the criminal element. The word "serve" on the other hand is somewhat ambiguous. What "to serve" may mean to one law enforcement agency it may mean quite the opposite to another. "To serve" also takes on a different meaning depending upon department size. For example, I know a chief in a small village not far from the city where I work. He recently had a call to "assist the woman." We all get these types of calls, but his was to assist the woman in re-hanging her draperies! To serve? Is that what people want? A tax supported drapery service? (51).

There are two striking aspects to this passage and the article, which also seems representative of the views of many police officers, and much of the public. The first striking aspect is the extent to which "service" is framed as a question of resources. Of course, the police are public servants, as are other agents and officers of government. But they also have a specific function, and should have priorities within that function. Indeed, the rest of the article is devoted to discussing the way nonemergency calls are overloading 9-1-1 operators and keeping police from getting to real emergencies. "In many small cities, the police are the only visible and accessible arm of the local government available after 5 p.m. and on weekends. Because of that we become the water department, the street department, the dog warden, etc. — and people begin to expect it from us." (51).

Of course, the "public" within the concept of public servant should be understood to include everyone in the community, not just "citizens" or "taxpayers" or even just "law abiding" people. Police have a duty to serve everyone, including the "criminal element."

Following several court and Supreme Court decisions in the United States, there is now a legal precedent that police do not have a specific legal duty to protect, or even to enforce the law or court orders. At least in terms of having a duty to lend aid or to protect a particular individual, a police officer is not compelled by the law to intervene, put themselves at risk, or act to enforce applicable laws. The court has upheld the discretion of police to decide when and where to enforce the law or protect individuals from danger (52).

data-driven policies and actions, drawing upon practical examples from data-driven policing.

Predictive policing, as the application of AI techniques to policing data, has its roots in much older practices of collecting crime data. Yet it also has the potential to draw upon data from other sources in increasingly networked police departments, and increasingly digitally surveilled communities. Ethical questions arise at almost every stage of data collection and analysis, from where data is collected and sensors are placed, to how data is encoded, to existing biases in segregated communities and policing practices, to the ways data is used in police management and police encounters with the public. For building a more general approach to AI ethics, it is useful to separate these problems out and identify the key ethical issues, and how AI researchers and system designers might think about and address them.

Data: From CompStat to Critical Data Science

Information and communication technologies (ICT) have long been central to policing. From the keeping of criminal records and crime statistics and their collection in databases, to the use of police boxes, telephones, radio dispatching, and 9-1-1 emergency call centers, many ICT technologies have become as closely associated with policing as badges and handcuffs. Initially, these technologies were analog–paper records, photographs and inked fingerprints, dedicated police telephone boxes, and wireless radios. With the computerization of businesses and government agencies from the 1960s to 1990s, many aspects of police work also became digitized and computerized. Police patrol cars began getting computers in the early 1980s, which allowed officers to check vehicle license plates, and eventually check individuals for outstanding warrants. The transition from paper to digital records for crime reports soon led to interest in compiling crime statistics at a local level for use in guiding the management of patrols and policing priorities. CompStat, short for Comparative Statistics, was the result. Initially adopted by the New York City police department in 1995, similar practices have been adopted across the country, especially in large urban departments.

CompStat as a mere data gathering and management practice has not been without its critics. In 2010, John Eterno and Eli Silverman, retired New York police captains turned university professor and criminology professor, respectively, published a book-length criticism of CompStat practices in the NYPD (54). The book argues that there was widespread misreporting of crimes across NYPD precincts, which took the form of downgrading the seriousness of reported crimes in an effort to show annual improvements in serious crime statistics. They argued that this systematic downgrading of crime statistics was the result of pressure from police leadership and administration. They further argued that pressures to increase police stops, especially in the era of "stop and frisk" in New York City, was highly racially discriminatory. The book caused enough controversy and embarrassment for the NYPD that the Police Commissioner ordered an independent study to review CompStat (55). That review did indeed find serious systemic reporting errors. It did not, however, find evidence that this was the result of administrative pressure, though the review did not investigate that point exhaustively, nor did it seriously assess systemic racism within CompStat's data collection practices.

What emerges from the investigations and reports into CompStat, from a data science and AI ethics perspective, is the susceptibility of data to political and bureaucratic pressure. While it may be convenient to assume that a given dataset offers an accurate representation of the world, this should not be taken for granted. In this case there were widespread and systematic errors in the reported data. If that data were to be used by predictive policing algorithms, those errors could have a significant impact on policing practices. And if that data is indeed racially biased, as it most likely is, it could further bias policing practices. But without an awareness of these issues, and the potential for inaccurate data or latent bias within data, the designers of those AI algorithms may be creating garbage-in-garbage-out systems, believing that they are producing quality systems (as measured by their available data). The lesson for AI ethics is to never take for granted the accuracy of given data, but to be suspicious, to seek out likely ways in which political, economic, or social pressures may have influenced historical datasets, to consider how it may be shaping current data collection practices, and to be sensitive to the ways in which new data practices may transform social practices and how that relates to the communities and individuals a system aims to care for.

With the growing popularity of AI, and increasing concerns about its impact on society, universities and professional organizations have recognized the problem and taken up the challenge of teaching ethics to the next generation of AI designers. Today, many undergraduate and graduate programs teaching AI include ethical training, but its adoption has been uneven and more could be done. Many online and professional training programs still lack critical design and ethical thinking in favor of teaching the latest techniques and tools over good design. Professional organizations including IEEE, ACM, and AAAI have also led initiatives to develop ethical standards, codes of ethics, and organize a growing number of conferences and workshops on AI ethics. These are all positive developments, and it is hoped that

this paper will contribute to the discussion of the ethical design of AI, especially as AI comes to be applied in an increasing number of socially significant and ethically consequential decisions.

While not every AI system developer can become an expert in the application domain of their techniques, the basics of critical data analysis should be taught alongside statistical techniques and machine learning techniques. In particular, system designers should be adept at recognizing the necessary characteristics of an adequate dataset, and what can and cannot be reasonably drawn from a given dataset. In many cases, only domain experts will have the kind of cultural knowledge to identify exogenous influences. This fact supports a systems design approach that includes domain experts as well as critical social scientists as members of design teams, and recognizes and respects the necessity of their expertise in shaping the ultimate system design (56).

Models Matter

A dataset on its own is just a collection of numbers delimited by some kind of file structure. Even decisions as to how to represent a data field with a number — binary, integer, real, pointer, formula — can have consequences for how that data gets processed. Numbers are abstract values, which are then represented by digital numerals within computational systems. How they are numerically represented can matter. But often it is far more important how we choose to represent the world through numbers. Even when we are simply "counting" things in the world, we are also engaged in processes of classification and categorization. The data "model" that a system employs involves myriad representational choices, and seeks to serve various purposes (57).

The most obvious case in law enforcement is to characterize the law, and represent violations of the law. But there are many possible computational models of any given set of legal rules and codes, and they may not always represent the same mappings of events in the world as computational encodings.

Consider the case of CompStat crime underreporting discussed above. We could look to New York Penal Law §155.05 and §155.25 for a definition of "Petite Larceny" which is theft or withholding of property valued at less than $1000 (and not a firearm, automobile, or credit card) (58). What if a bike has been stolen, which cost a little more than $1000 when it was new, but it is used and would likely not sell for that much, nor would an insurance company compensate its loss for more than $1000? Determining the appropriate crime requires estimating the value of the property. This is a non-trivial categorization — an auction might determine the current market value, or a bike sales expert might be able to give an appraisal, but these may not agree on the price, nor be available means for a law enforcement officer. To some extent there is discretion on the part of law enforcement, prosecutors, and judges as to how to appraise and categorize such a crime — and they may take factors into account other than the strict value of the property. But once categorized, that discretionary nature tends to be erased — the crime becomes defined through its given category, documented and entered into data collection systems. AI systems designers need to be sensitive these types of processes. Indeed, understanding data collection, and critical data representation issues should be integral to computer and information science education. Taking care in the design of AI means being able to determine what an adequate dataset is, and being able to think critically about how to define it, and what the implications of various choices of categorization are. How best to do this, in general, is a matter for further research.

Putting AI Into Practice

The discussion so far has focused on input — how data is structured and collected. But the presentation of data analysis, and its impact on individual and institutional practices must also be taken in account. A good example of such an issue can be seen in the use of the SSL by Chicago police. In principle, the SSL could have been used to recruit youth for the One Summer program. The choice by precincts and officers to use the list for "custom notification" and for "heat lists" following crimes is not disconnected from the design of a system like SSL. While data scientists and software engineers may wish to wash their hands of responsibility for how officers actually use their tools, they cannot. At the very least this constitutes a sort of negligence and failure to warn. Many officers were not properly or fully informed of how the list was put together, and held mistaken and problematic understandings of what it was and how it worked. The officers also lacked training, guidance, and direction on how to use the system, if indeed there ever was a comprehensive plan as to how to deploy and use the system. These factors surely contributed to its misuse, and all but guaranteed its ineffectual use.

An Ethics of Care approach ought to ensure that the operators of AI systems and users of data they generate are aware of the scope and limitations of those systems. It may be too much to expect them to fully understand the computational techniques — indeed even AI experts may find the performance of certain machine learning systems inscrutable. But this does not mean that people who use these systems can be ignorant of what the system can and cannot do, how reliable it is, and what its limitations in representing the world are.

Designers also need to be aware of the context in which AI systems will be deployed and used. It should

not be hard to predict what police might do with a "heat list," if one has a realistic sense of police work and the pressures operating within precincts and departments. This again points to the need for domain experts and participatory design (56). One imagines that a police sergeant on the design team of the SSL would have pointed out the likely misuses of the system. Prototyping and testing could also help reveal such tendencies, as well as short-term and long-term evaluations of the system implementation.

Transparency over the algorithms, data, and practices of implementation are also necessary. While the Chicago Police Department sought to avoid embarrassment from releasing the details of the SSL, it would be impossible for independent outside researchers to evaluate its impacts — positive and negative — without access to the data and algorithms. It should not take a prolonged lawsuit from a newspaper for government agencies to share public data. Of course, as more and more commercial systems, like PredPol,[19] make the algorithms and even the data proprietary, they will fall under intellectual property protections. This means private companies will be processing the data, and will not be required to reveal their algorithms, or subject them to independent outside scrutiny. In some cases, private companies are even withholding crime data from the cities that produced it because they have formatted it in a database for their system and even encrypted it such that it cannot be used if the city changes to another software platform (59).

Central Issues Facing AI Predictive Policing

It is hoped that this article has shed light upon some of the central issues facing AI ethics in general and predictive policing in particular. While the use of data and AI in policing is not intrinsically or necessarily unethical, it must be done with care to avoid unjust and unethical impacts. First among these issues is that while AI ethics needs to understand the computational techniques it deploys, it also needs a critical understanding of the datasets it operates on, how data is collected, and the social organizations and the biases that those datasets may represent. This requires understanding how data practices are embedded within socio-technical systems, and not blindly analyzing data assuming that it is without bias. It is also important to understand how the use of AI tools and techniques will impact the beliefs and practices of those who engage with them. Datasets and their computational analysis have the power to "makeup people" in the sense of Hacking (36), and also to prejudge them according to statistical patterns and categories. Even when statistically justified, such categories, and the actions of government agents on the basis of those categories, may disrespect individual rights, human dignity, and undermine justice.

By taking an Ethics of Care approach to AI systems design and ethics, designers should have a greater awareness and respect for these issues. While any design approach is ultimately limited in its ability to mitigate all possible failures and harms, an Ethics of Care can help mitigate the most significant and widespread flaws in AI systems that will impact people's lives in consequential ways. An AI Ethics of Care has the potential to apply to areas far beyond predictive policing, and can inform many applications of AI for consequential decisions.

[19]PredPol is a commercial software company developing data management and predictive data systems for police departments (30).

Acknowledgment

This work was supported in part by a Beneficial AI research grant from the Future of Life Institute.

Author Information

Peter Asaro is Associate Professor and Director of Graduate Studies in the School of Media Studies at The New School, New York, NY. He is also Visiting Professor at the Munich Center for Technology in Society at TU Munich, and Affiliate Scholar at Stanford Law School's Center for Internet and Society. Email: asarop@newschool.edu.

References

(1) P. Miller and T.O'Leary, "Accounting, 'economic citizenship' and the spatial reordering of manufacture," *Accounting, Organizations and Society*, vol. 19, no. 1, pp. 15-43, 1994.

(2) S. Zuboff, *In the Age of the Smart Machine: The Future of Work and Power*. Basic, 1988.

(3) L. Winner, Langdon, *Autonomous Technology*. Cambridge, MA: M.I.T. Press, 1977.

(4) P. Asaro and W. Wallach, "An introduction to machine ethics and robot ethics," in *Machine Ethics and Robot Ethics* (The Library of Essays on the Ethics of Emerging Technologies), W. Wallach and P. Asaro, Eds. Routledge, 2017; http://peterasaro.org/writing/WALLACH%20ASARO%20(Machine%20Ethics%20Robot%20Ethics)%20.pdf.

(5) P. Asaro, "What should we want from a robot ethic?," *Int. Rev. Information Ethics*, vol. 6, no. 12, pp. 9-16, 2006.

(6) D.K. Citron, "Technological due process," University of Maryland Legal Studies, Res. Pap. no. 2007-26; *Washington Univ. Law Rev.*, vol. 85, pp. 1249-1313, 2007; https://ssrn.com/abstract=1012360.

(7) F. Pasquale, *The Black Box Society*. Cambridge, MA: Harvard Univ. Press, 2015.

(8) A. Selbst and S. Barocas, "Regulating inscrutable systems," presented at WeRobot 2017; 2017; http://www.werobot2017.com/wp-content/uploads/2017/03/Selbst-and-Barocas-Regulating-Inscrutable-Systems-1.pdf.

(9) R. Caplan, J. Donovan, L. Hanson, and J. Matthews, "Algorithmic accountability: A primer," *Data & Society Tech. Rep.*, Apr. 18, 2018; https://datasociety.net/output/algorithmic-accountability-a-primer/.

(10) V. Eubanks, *Automating Inequality: How High-tech Tools Profile, Police and Punish the Poor*. St. Martin's, 2017.

(11) S.U. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: New York Univ. Press, 2018.

(12) C. O'Neil, *Weapons of Math Destruction*. Crown Random House, 2017.

[13] J. Powles and H. Nissenbaum, "The seductive diversion of 'solving' bias in artificial intelligence," *Medium*, Dec. 7, 2018; https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53.
[14] Wikipedia, "Threat model," https://en.wikipedia.org/wiki/Threat_model; accessed May 2019.
[15] J.G. Stein, "Threat perception in international relations," *The Oxford Handbook of Political Psychology*, 2nd ed., L. Huddy, D.O. Sears, and J.S. Levy, Eds. Oxford, U.K. Oxford Univ. Press, 2013.
[16] M. Sander-Staud, "Care ethics," *Internet Encyclopedia of Philosophy,* https://www.iep.utm.edu/care-eth/, accessed May 2019.
[17] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, "Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks," *ProPublica*, May 23, 2016; https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
[18] J. Dressel and H. Farid, "The accuracy, fairness, and limits of predicting recidivism," *Science Advances*, vol. 4, no. 1, Jan. 17, 2018; http://advances.sciencemag.org/content/4/1/eaao5580/tab-pdf.
[19] P.K. Dick, *The Minority Report*, 1956.
[20] *Minority Report*, Stephen Spielberg, dir., 2002.
[21] A. Shapiro, "Reform predictive policing," *Nature*, Jan. 25, 2017; https://www.nature.com/news/reform-predictive-policing-1.21338.
[22] A.G. Ferguson, *The Rise of Big Data Policing: Surveillance, Race and the Future of Law Enforcement*. New York, NY: New York Univ. Press, 2017.
[23] H. Kerrigan, "Data-driven policing," *Governing the States and Localities*, May 2011; http://www.governing.com/topics/public-justice-safety/Data-driven-Policing.html.
[24] S. Brayne, "Big Data surveillance: The case of policing," *American Sociological Rev.*, vol. 82, no. 5, pp. 977-1008, 2017.
[25] "CompStat," *Wikipedia*, https://en.wikipedia.org/wiki/CompStat, accessed May 2019.
[26] G. Rayman, "NYPD commanders critique Comp Stat and the reviews aren't good," *Village Voice*, Oct. 18, 2010; https://www.villagevoice.com/2010/10/18/nypd-commanders-critique-comp-stat-and-the-reviews-arent-good/.
[27] "Crime data in Chicago," *Trulia.com*; https://www.trulia.com/real_estate/Chicago-Illinois/crime/, accessed May 2019.
[28] "The origin of the LAPD motto," *BEAT Mag.*, Dec. 1963; http://www.lapdonline.org/history_of_the_lapd/content_basic_view/1128.
[29] Wikipedia, "Duty of care;" https://en.wikipedia.org/wiki/Duty_of_care, accessed May 2019.
[30] "Overview," *PredPol*, 2018; http://www.predpol.com/about/.
[31] J. Barro, "Here's why stealing cars went out of fashion," *NYTimes*, Aug. 11, 2014; https://www.nytimes.com/2014/08/12/upshot/heres-why-stealing-cars-went-out-of-fashion.html.
[32] F. Chen and R. Regan, "Arts and craftiness: An economic analysis of art heists," *J. Cultural Economics*, vol. 41, no. 3, pp. 283-307, Aug. 2017; https://economiststalkart.org/2016/05/31/why-are-there-so-many-art-thefts-and-what-can-be-done-about-them/.
[33] A. Gabbatt, "New York woman visited by police after researching pressure cookers online," *The Guardian*, Aug. 1, 2013; https://www.theguardian.com/world/2013/aug/01/new-york-police-terrorism-pressure-cooker.
[34] I. Hacking, *The Emergence of Probability: A Philosophical Study of Early Ideas About Probability, Induction and Statistical Inference,* 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2006 (1975).
[35] R. Nilsen, "How well do horse racing favorites perform?," Feb. 12, 2012; http://agameofskill.com/how-well-do-horse-racing-favorites-perform/.
[36] I. Hacking, "Making up people," in *Reconstructing Individualism: Autonomy, Individuality and the Self in Western Thought*, T.C. Heller, Ed. Stanford Univ. Press, 1986, pp. 222-236.
[37] Y. Romanyshyn, "Chicago homicide rate compared: Most big cities don't recover from spikes right away," *Chicago Tribune*, Sept. 26, 2017; http://www.chicagotribune.com/news/data/ct-homicide-spikes-comparison-htmlstory.html.
[38] University of Chicago, "One Summer Project," *Urban Labs*, https://urbanlabs.uchicago.edu/projects/one-summer-chicago-plus-nothing-stops-a-bullet-like-a-job, accessed May 2019.
[39] "Strategic Subject List," *Chicago Data Portal*; https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List/4aki-r3np, accessed May 2019.
[40] M. Dumke and F. Main, "A look inside the watch list Chicago Police fought to keep secret," *Chicago Sun-Times*, May 18, 2017; https://chicago.suntimes.com/politics/what-gets-people-on-watch-list-chicago-police-fought-to-keep-secret-watchdogs/.
[41] Y. Kunichoff and P. Sier, "The contradictions of Chicago Police's secretive list," *Chicago Mag.*, Aug. 2017; http://www.chicagomag.com/city-life/August-2017/Chicago-Police-Strategic-Subject-List/.
[42] J. Gorner, "With violence up, Chicago Police focus on a list of likeliest to kill, be killed," *Chicago Tribune*, July 22, 2016; http://www.chicagotribune.com/news/ct-chicago-police-violence-strategy-met-20160722-story.html.
[43] J. Saundrs, P. Hunt, and J.S. Hollywood, "Predictions put into practice: A auasi-experimental evaluation of Chicago's Predictive Policing pilot," *J. Experimental Criminology,* vol. 12, no. 3, pp 347–371, Sept. 2016.
[44] M.K. Sparrow, *Handcuffed: What Holds Policing Back, and the Keys to Reform*. Brookings Inst. Press, 2016.
[45] Office of the Mayor, "Study: Chicago's One Summer Plus Youth Employment Program cuts violent crime arrests in half," Press Release, City of Chicago, Chicago, IL, Aug. 6, 2013, https://www.cityofchicago.org/city/en/depts/mayor/press_room/press_releases/2013/august_2013/study_chicago_s_onesummerplusyouthemploymentprogramcutsviolentcr.html.
[46] S.B. Heller, "Summer jobs reduce violence among disadvantaged youth," *Science*, vol. 346, no. 6214, pp. 1219-1223, Dec. 5, 2014.
[47] J. Hollywood, "CPD's 'Heat List' and the dilemma of Predictive Policing," *RAND Blog*, Sept. 2016; https://www.rand.org/blog/2016/09/cpds-heat-list-and-the-dilemma-of-predictive-policing.html.
[48] W. Hartzog, G. Conti, J. Nelson, and L.A. Shay, "Inefficiently automated law enforcement," *Michigan State Law Rev.*, pp. 1763-1796, 2015; https://pdfs.semanticscholar.org/ec71/95d72b4ea51c9c6cc5d6a0e153448bbf702e.pdf.
[49] "Ethics of Care" *Wikipedia*, https://en.wikipedia.org/wiki/Ethics_of_care, accessed May 2019.
[50] V. Held, *Ethics of Care: Personal, Political and Global*, 2nd ed. Oxford Univ. Press, 2006.
[51] M. Burg, "To serve and protect?" *Police Mag.*, 1998; http://www.policemag.com/channel/patrol/articles/1998/12/to-serve-and-protect.aspx.
[52] K. Keopong, "The police are not required to protect you," *Barnes Law*, June 26, 2016; http://www.barneslawllp.com/police-not-required-protect/.
[53] A. Van Wynsberghe, "Designing robots for care: Care centered value-sensitive design," *Science and Engineering Ethics*, vol. 19, no. 2, pp. 407-433, 2013.
[54] J. Eterno and E. Silverman, *The Crime Numbers Game: Management by Manipulation*. CRC, 2010.
[55] D.N. Kelley and S.L. McCarthy, "The Report of the Crime Reporting Review Committee to Commissioner Raymond W. Kelley concerning CompStat Auditing," *NYPD*, Apr. 8, 2013 (released July 2013); http://www.nyc.gov/html/nypd/downloads/pdf/public_information/crime_reporting_review_committee_final_report_2013.pdf.
[56] P. Asaro, "Transforming society by transforming technology: The science and politics of participatory design," *Accounting, Management and Information Technologies*, vol. 10, no. 4, pp. 257-290, 2000; http://peterasaro.org/writing/Asaro%20PD.pdf.
[57] G.C. Bowker and S.L. Star, *Sorting Things Out: Classification and its Consequences*. Cambridge, MA: M.I.T. Press, 2000.
[58] "New York State Penal Code," *New York Laws*, 2019; http://ypdcrime.com/penal.law/article155.htm?#p155.05.
[59] E. Joh, "The undue influence of surveillance technology companies on policing," *New York Univ. Law Rev.*, 2017.

TS