OPEN-ISH GOVERNMENT LAWS AND THE DISPARATE RACIAL IMPACT OF CRIMINAL JUSTICE TECHNOLOGIES

BEN WINTERS[*]

ABSTRACT

Automated decision-making systems are used widely and opaquely in and around the U.S. criminal justice cycle. There are serious transparency and oversight concerns around the use of these tools in a system that severely disadvantages already marginalized communities. The Intellectual Property exemptions included in open government laws are one key aspect that prevents public understanding of important details of these tools. This paper attempts to explain the harm compounded by the use of these tools as well as the lack of access to meaningful information about them through government transparency mechanisms and analyze various harm-mitigation options.

[*] Equal Justice Works Fellow, Electronic Privacy Information Center (EPIC). The fellowship and research supporting it are primarily funded by the Philip M. Stern Foundation through Equal Justice Works. The article reflects the current state of legal and risk assessment model developments at the time of publication.

I. INTRODUCTION

On a given day, a single individual may interact with several automated decision-making systems[1] that individually and opaquely surveil, collect, store, and analyze data about them. For example, a person may be required to use facial recognition scanners to enter public housing or encounter algorithms that determine benefit eligibility.[2]

[1] *Automated Decision Systems: Examples of Government Use Cases*, AI NOW INST., https://ainowinstitute.org/nycadschart.pdf [https://perma.cc/26GB-2EV3] (defining "Automated Decision Systems" as "technical systems that aim to aid or replace human decision making").
[2] *See, e.g.*, Lola Fadulu, *Facial Recognition Technology in Public Housing. Prompts Backlash*, N.Y. TIMES, (Sept. 24, 2019), https://www.nytimes.com/2019/09/24/us/politics/facial-recognition-

Predictive policing systems feed in past arrest data to help determine where police should use enforcement resources,[3] leading to added police presence in already over-policed neighborhoods.[4] Simultaneously, property surveillance systems like Ring doorbells on houses across many neighborhoods may be connected to the local police.[5] If a person is stopped by a police officer and arrested, that individual's information is entered into a process of risk assessment tools that informs determinations of bail and detention.[6] Those who have additional medical needs, apply for benefits, or have family and friends interacting with the

technology-housing.html [https://perma.cc/2ZMU-XXNF]; Colin Lecher, *What Happens When an Algorithm Cuts Your Health Care*, VERGE (Mar. 21, 2018), https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy [https://perma.cc/Y2AY-A768].

[3] Andrew G. Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1109, 1113 (2017).

[4] *Id.* at 1148.

[5] *See* Kim Lyons, *Amazon's Ring Now Reportedly Partners with More Than 2,000 US Police and Fire Departments*, VERGE (Jan. 31, 2021), https://www.theverge.com/2021/1/31/22258856/amazon-ring-partners-police-fire-security-privacy-cameras [https://perma.cc/ETF5-PUEB]; Drew Harwell, *Doorbell-Camera Firm Ring Has Partnered with 400 Police Forces, Extending Surveillance Concerns*, WASH. POST (Aug. 28, 2019), https://www.washingtonpost.com/technology/2019/08/28/doorbell-camera-firm-ring-has-partnered-with-police-forces-extending-surveillance-reach/ [https://perma.cc/ZG4Z-JZH6]; Jane Wakefield, *Ring Doorbells to Send Live Video to Mississippi Police*, BBC NEWS (Nov. 5, 2020), https://www.bbc.com/news/technology-54809228 [https://perma.cc/KGL9-A9XL].

[6] *See generally AI and Human Rights: Criminal Justice System*, ELEC. PRIV. INFO. CTR. (EPIC) [hereinafter *AI and Human Rights*], https://epic.org/ai/criminal-justice/#foia [https://perma.cc/3BSG-MQ7D] (showing various documents from public records requests on this page and what inputs go into various risk assessments at pre-trial and in parole processes).

criminal justice system, may also be labeled as riskier by a given actuarial tool.[7]

If applying for a job, a resume scanning algorithm might determine if someone's quality of education is deemed inadequate, or there may be an inaccurate data set and/or misguided values embedded into the system that identifies that they are unfit for a certain job. Another possibility in the job application process is that an algorithm has determined that an applicant does not have the right mood, facial expressions, or tone to match the culture of a job.[8] The automated decision-making around job applications critically connects to the risk assessment tools used pre-trial, at trial, and during parole which use details around a person's job status to make determinations about their risk level.[9] This piece will show that these tools are all connected

[7] *Level of Service Inventory – Revised*, IDAHO DEP'T CORR., https://epic.org/EPIC-19-11-21-ID-FOIA-20191206-ID-lsi-paper-scoresheet-tips-and-hints.pdf [https://perma.cc/HY9W-YPGT] (annotated by the Idaho Department of Corrections).

[8] *See In re HireVue*, ELEC. PRIV. INFO. CTR., https://epic.org/privacy/ftc/hirevue/ [https://perma.cc/QAU5-3527]; Aarti Shahani, *Now Algorithms Are Deciding Whom to Hire, Based on Voice*, NPR (Mar. 23, 2015), https://www.npr.org/sections/alltechconsidered/2015/03/23/394827451/now-algorithms-are-deciding-whom-to-hire-based-on-voice [https://perma.cc/2N3G-52YF]; *see also* Fiona J McEvoy, *3 Reasons to Question the Use of Emotion-Tracking AI in Recruiting*, VENTUREBEAT (Mar. 12, 2018), https://venturebeat.com/2018/03/12/3-reasons-to-question-the-use-of-emotion-tracking-ai-in-recruiting/ [https://perma.cc/GC2W-6X9K] (critiquing the use of AI in hiring).

[9] *See*, e.g., *Re: FOIA Request - 20-FOIA*-00095, PRETRIAL SERVS. AGENCY FOR D.C. OFF. PLAN., POL'Y & ANALYSIS (Feb. 21, 2020), at 6, https://epic.org/EPIC-20-01-08-DC-FOIA-20200308-DCPSA-Factors-Change-2015-2019.pdf [https://perma.cc/GG6B-5M6U] (showing employment status as part of pre-trial tool, which is also used as part of the trial); *Nevada Parole Risk Assessment*, NEV. PAROLE DEP'T, http://parole.nv.gov/uploadedFiles/parolenvgov/content/Information/NV_ParoleRiskAssessmentForm.pdf [https://perma.cc/3NU4-E3X9] (showing employment history as part of parole risk determination).

and operating in the same criminal justice cycle, and that they almost all fail to be meaningfully transparent on several levels. There are various legal and factual levers that keep (1) the existence and use of the tools, and (2) key information such as the factors, the weights, the policies surrounding use and the developer hidden away from public view and public scrutiny.

Automated decision-making tools are being used significantly and opaquely in the U.S. Criminal Justice System.[10] Although the function of these tools vary, many encode series of judgments about the likelihood an individual will, for example, be arrested based on a series of defined factors.[11] Race, gender, socioeconomic status, and age, in addition to proxies of these factors, are often included as factors in these loaded predictions.[12] When advocates and community members aim to understand the full scope of where and how this technology is being used, they are often stifled in part by overbroad trade secret exemptions that are in open government laws invoked by the contractors that

[10] *See Liberty at Risk: Pre-Trial Risk Assessment Tools in the U.S.*, ELEC. PRIV. INFO. CTR., at 5–8 (Sept. 2020) [hereinafter *Liberty at Risk*], https://epic.org/LibertyAtRiskReport.pdf [https://perma.cc/DN34-SLJQ] (showing the status of the usage of pre-trial risk assessment tools in every state); *Overview*, PREDPOL, https://www.predpol.com/about/ [https://perma.cc/9239-A6NR] (stating that "PredPol is currently being used to help protect one out of every 33 people in the United States." This shows that one of the Predictive Policing tools covers many jurisdictions, while other tools in the market also exist); *Automatic License Plate Reader Documents: Interactive Map*, ACLU, https://www.aclu.org/issues/privacy-technology/location-tracking/automatic-license-plate-reader-documents-interactive-map?redirect=maps/automatic-license-plate-reader-documents-interactive-map [https://perma.cc/NBE6-66UU] (showing how widely automated license plate readers are used in the United States).

[11] *See, e.g. Level of Service Inventory – Revised*, *supra* note 7.

[12] Chelsea Barabas et al., *Interventions over Predictions: Reframing the Ethical Debate for Actual Risk Assessment*, 81 PROC. MACH. LEARNING RSCH. 1, 3 (2018).

made the system.[13] The limitations prevent seamless public understanding about the tools their government is adopting, and consequently limit advocacy around tools with accuracy, bias, or privacy concerns. Although the adoption and use of these tools requires significant reform, the time and cost of battling trade secret claims on open government requests to even understand *what* is being used by their police department or courts is a clear place to start reform that could result in improved transparency in this space.

These tools allow agencies to evade accountability and perpetuate, rather than confront, racial, ethnic, and socioeconomic bias. The developers of these tools conceal the inner workings of their programs and embrace trade

[13] *See, e.g.*, Letter from Jeanean West, FOIA Officer, Office of General Counsel, to author (Feb. 11, 2020), https://epic.org/PDN%20Respsone%20to%20FOIA%20Requester.pdf [https://perma.cc/PEV9-CH7H] (notifying a FOIA requestor of the required redisclosure notification being provided to the trade secret holder); Letter from Andrea Barnes, Staff Attorney, Miss. Dep't of Corrections, to author (Dec. 5, 2019), https://epic.org/Winters.Ben.EPIC.Response.11.25.19.12.03.19.pdf [https://perma.cc/7UTF-G6G5] (notifying requestor that a third-party has been put on notice due to trade secret concerns). *See generally Open Gov't Guide*, REPS. COMM. FOR FREEDOM PRESS, https://www.rcfp.org/open-government-guide [https://perma.cc/5R42-B3ZA (showing freedom of information laws in 50 states and includes the trade secret/commercial protections). The Reporters Committee for Freedom of the Press resource highlights several examples of trade secret protection. In Wisconsin, for example, "Trade secrets, as defined in the Uniform Trade Secrets Act, Wis. Stat. § 134.90(1)(c), may be closed" and, when it comes to contracts, proposals, and bids, they "are subject to the balancing test, but may be closed if competitive or bargaining reasons require." *Id.* (citing WIS. STAT. §§ 19.36(5), 19.85(1)(e), and 19.35(1)(a)). In New Mexico, "[t]rade secrets are exempt from disclosure." *Id.* (citing N.M. STAT. ANN § 14-2-1(F) (2019)). In Louisiana, "proprietary or trade secret information provided to public bodies by the developer, owner, or manufacturer of a code, pattern, formula, design, device, method or process in order to obtain approval for sale or use in the state is specifically exempted from the Public Records Act." *Id.* (citing LA. STAT. ANN. § 44:3.2).

secret protections at trials against both individual defendants and in response to open government requests.[14] This opacity diminishes accountability, transparency, trust, and the exercise of a complete criminal defense, to the detriment of defendants. From what researchers have accessed, this policy of secrecy embraces, rather than confronts, the reasons these biases and their effects proliferate.[15]

This article will: (1) illustrate the harms that this cycle of using automated decision-making tools in and around the criminal justice system causes and why the opacity exacerbates those harms specifically along racial lines; (2) explain the overbroad commercial intellectual property (IP) protections at both trial and in open government contexts; and (3) survey harm-mitigation strategies for increased opacity when technology implicates high-risk governmental functions.

[14] Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1343 (2018) (stating that "[d]evelopers often claim that details about how their tools work are trade secrets and refuse to disclose that information to criminal defendants or their attorneys."); *id.* at 1366 (saying that "[i]n addition to facilitating law enforcement evasion of judicial scrutiny, trade secret claims may also motivate—or even compel—such evasion; companies may require law enforcement agencies to conceal the use of their products or engage in 'parallel construction,' in which police disguise the actual methods they use by describing alternative ones, in order to protect sensitive information from courtroom disclosure."); *See also EPIC v. DOJ (Criminal Justice Algorithms)*, ELEC. PRIV. INFO. CTR., https://epic.org/foia/doj/criminal-justice-algorithms/ [https://perma.cc/5R5F-JH9C] (showing a FOIA case fighting disclosure of report on the use of predictive analytical tools in the criminal justice system).

[15] *See Liberty at Risk*, *supra* note 10, at 1; *Mapping Pretrial Injustice: A Community-Driven Database*, MOVEMENT ALL. PROJECT, [hereinafter *Mapping Pretrial Injustice*], https://pretrialrisk.com/ [https://perma.cc/R57Q-63KL]; *see generally* sources cited *infra* note 79.

II. A SAMPLING OF THE TECHNOLOGY AND THE PARTICULARLY RACIALIZED HARMS

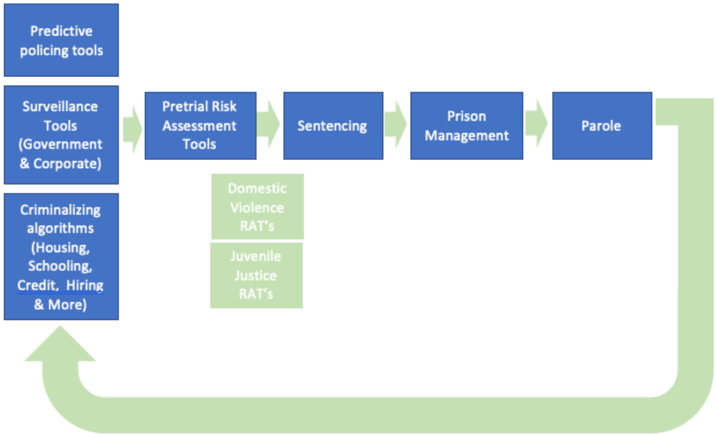
Predictive policing tools
Surveillance Tools (Government & Corporate)
Criminalizing algorithms (Housing, Schooling, Credit, Hiring & More)
Pretrial Risk Assessment Tools
Sentencing
Prison Management
Parole
Domestic Violence RAT's
Juvenile Justice RAT's

Figure 1: An approximate cycle of the different algorithms and automated decision-making systems used in the criminal justice cycle

There are near-endless streams of automated decision-making tools used in and around the U.S. criminal justice system that can be categorized by the different users of the technology. These can take the form of predictive algorithms, synthesized databases, or surveillance tools. Some technologies are used by government entities for criminal justice purposes, like police departments and corrections departments.[16] Other users are government entities outside the criminal justice context, like health departments, that use algorithms to support decision-making, resource allocation, and benefits.[17] Others are

[16] *See Liberty at Risk*, *supra* note 10, at 2–4.

[17] *See* DAVID FREEMAN ENGSTROM ET AL., GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 10 (2020), https://www-cdn.law.stanford.edu/wp-content/uploads/2020/02/ACUS-AI-Report.pdf [https://perma.cc/LCR8-FEVM] (providing examples of algorithm use by United States administrative agencies); Lecher, *supra* note 2

corporations that deploy profiling products themselves such as Clearview AI, and some are corporations that sell surveillance products to consumers such as Ring Doorbell, while often networking these cameras and working with law enforcement.[18] This section will not chronicle every type of technology, as it is nearly impossible to do so partially due to minimal transparency requirements and frequent changes in adoption. Also, significant work mapping out racial harms of technologies around the criminal justice system has been done by scholars and journalists.[19] The next several paragraphs in this section will expand on Figure 1 above, which focuses on algorithmic tools and other technology used in and around the criminal justice system.

Predictive Policing is "any policing strategy or tactic that develops and uses information and advanced analysis to inform forward-thinking crime prevention."[20] Predictive policing comes in two main forms: location-based and person-based.[21] Location-based predictive policing works by identifying places of repeated property crime in an

(discussing using an algorithm to set the required number of hours that a caretaker could visit patients).

[18] *See* Kashmir Hill, *The Secretive Company That Might End Privacy As We Know It*, N.Y. TIMES (Jan. 18, 2020), https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html [https://perma.cc/C7V5-VQAJ]; Rani Molla, *How Amazon's Ring is Creating a Surveillance Network with Video Doorbells,* VOX (Jan. 28, 2020), https://www.vox.com/2019/9/5/20849846/amazon-ring-explainer-video-doorbell-hacks [https://perma.cc/F85P-VVUR].

[19] For a current piece focusing on the breadth of racially impactful technologies around policing, see Laura Moy, *A Taxonomy of Police Technology's Racial Inequity Problems*, 2021 U. ILL. L. REV. 139.

[20] CRAIG D. UCHIDA, A NATIONAL DISCUSSION ON PREDICTIVE POLICING: DEFINING OUR TERMS AND MAPPING SUCCESSFUL IMPLEMENTATION STRATEGIES 1 (2009), https://www.ojp.gov/pdffiles1/nij/grants/230404.pdf [https://perma.cc/AW8R-8T2X].

[21] *See* Ferguson, *supra* note 3, at 1114.

attempt to predict *where* similar crimes will occur next.[22] Person-based predictive policing aims to pinpoint *who* might be committing a crime (i.e. trying to measure the risk that a given individual will be arrested for allegedly committing a crime.)[23] Both types of predictive policing are used in different jurisdictions and rely on past policing data as the main input for these predictions, creating a cycle of arresting resources.[24] The Bureau of Justice Assistance has and continues to give grants to police departments around the country to create and pilot these programs.[25] Simultaneously, two high profile jurisdictions recently stopped using predictive policing tools limited effectiveness and significant demonstrated bias, and some jurisdictions are banning the use of the technology.[26]

[22] *See id.* at 1127.

[23] *See id.* at 1137.

[24] *Id.* at 1149.

[25] U.S. DEP'T. JUST., PREDICTIVE ANALYTICS IN LAW ENFORCEMENT: A REPORT BY THE DEPARTMENT OF JUSTICE 4–5 , Department of Justice (2014), https://epic.org/foia/doj/criminal-justice-algorithms/EPIC-16-06-15-DOJ-FOIA-20200319-Settlement-Production-pt1.pdf [https://perma.cc/B39T-ZJXW].

[26] Caroline Haskins, *The Los Angeles Police Department Says It is Dumping A Controversial Predictive Policing Tool*, BUZZFEED NEWS (Apr. 21, 2020), https://www.buzzfeednews.com/article/carolinehaskins1/los-angeles-police-department-dumping-predpol-predictive [https://perma.cc/P9K2-FVU4]; Jeremy Gorner & Annie Sweeney, *For Years Chicago Police Rated the Risk of Tens of Thousands Being Caught Up in Violence. That Controversial Effort Has Quietly Been Ended*, CHI. TRIBUNE (Jan. 24, 2020), https://www.chicagotribune.com/news/criminal-justice/ct-chicago-police-strategic-subject-list-ended-20200125-spn4kjmrxrh4tmktdjckhtox4i-story.html [https://perma.cc/EMK3-KZYY]; 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 [https://perma.cc/8QW4-UWDK]; Ryan Johnston, *Oakland, Calif., Set to Ban Predictive Policing, Biometric Surveillance Tools*, STATESCOOP (Dec. 17, 2020), https://statescoop.com/oakland-calif-set-to-ban-

Surveillance Tools encompass a large swath of technologies and functions that can be used to track and store information about a person. This ranges from Ring doorbells and Clearview AI, which partners with law enforcement, to facial recognition systems at the border and in U.S. cities.[27]

"Criminalizing algorithms" include algorithms used in housing, credit determinations, healthcare, hiring, and school choice.[28] Many of these have been shown to make recommendations and decisions that negatively affect marginalized communities, encode systemic racism, and improperly lead people through Criminal Justice system.[29] The results of these criminalizing algorithms and the data points collected in their use can lead to higher determinations of riskiness and greater interaction with the criminal justice system.

predictive-policing-biometric-surveillance-tools/ [https://perma.cc/EV3P-8TB9].

[27] *See generally* Wakefield, *supra* note 5; Zach Whittaker, *Amazon's Ring Neighbors App Exposed Users' Precise Locations and Home Addresses*, TECHCRUNCH (Jan. 14, 2021, 10:00 AM), https://techcrunch.com/2021/01/14/ring-neighbors-exposed-locations-addresses/ [https://perma.cc/723J-VJSG] (discussing how the Ring's Neighbors app exposed the private location data of its users); Oscar Williams, *Clearview AI Facial Recognition Startup Partners with "600" Law Enforcement Agencies*, NEW STATESMEN (Jan. 20, 2020), https://tech.newstatesman.com/security/clearview-ai-facial-recognition-startup [https://perma.cc/SSS4-FD8M]; Abrar Al-Heeti, *US Border Protection Used Facial Recognition on 23 Million Travelers in 2020*, CNET (Feb. 11, 2021, 3:15 PM), https://www.cnet.com/news/us-border-patrol-used-facial-recognition-on-23-million-travelers-in-2020/ [https://perma.cc/TU2C-ZV5T]; Shirin Ghaffary & Rani Molla, *Here's Where the US Government is Using Facial Recognition Technology to Surveil Americans*, VOX (Dec. 10, 2019, 8:00 AM), https://www.vox.com/recode/2019/7/18/20698307/facial-recognition-technology-us-government-fight-for-the-future [https://perma.cc/AD2W-NSXT].

[28] *AI and Human Rights*, *supra* note 6.

[29] *See generally* sources cited *infra* note 79.

Risk Assessment Tools are used in almost every state in the U.S.—and many use them in a pre-trial setting, although they are also used at sentencing, in prison management, and for parole determinations.[30] There are also specific risk assessment tools used to assess risk for particular purposes such as in domestic violence or juvenile justice cases, with the understanding that they are designed to predict behavior more specific than general criminal risk or violent criminal risk of rearrest or re-offense.[31]

Pretrial Risk Assessment Tools are designed to attempt to predict future behavior by defendants and incarcerated persons and quantify the associated risk.[32] The tools vary, but make estimates using "actuarial assessments" like (1) "the likelihood that the defendant will re-offend before trial" ("recidivism risk") and (2) "the likelihood the defendant will fail to appear at trial" ("FTA").[33] These Pretrial Risk Assessment Tools use factors such as socioeconomic status, family background, neighborhood crime, employment status, as well as other considerations to reach a supposed prediction of an individual's criminal risk and report the risk using a simplified metric.[34]

Significant empirical research has shown disparate impacts of risk assessment tools on criminal justice outcomes based on the race, ethnicity, and age of the

[30] *Liberty at Risk*, *supra* note 10, at 1, 5–8.

[31] Chris Baird et al., *A Comparison on Risk Assessment Instruments in Juvenile Justice*, NAT'L COUNCIL ON CRIME AND DELINQ. (Aug. 2013), http://www.evidentchange.org/sites/default/files/publication_pdf/nccd_fire_report.pdf [https://perma.cc/2SS8-49C9]; *Ontario Domestic Assault Risk Assessment*, MENTAL HEALTH CTR. PENETANGUISHENE RSCH. DEP'T (last visited Apr. 9, 2021), https://grcounseling.com/wp-content/uploads/2016/08/domestic-violence-risk-assessment.pdf [https://perma.cc/CP8N-XQZS].

[32] *Liberty at Risk*, supra note 10, at 1.

[33] *Id.*

[34] *Id.* at 22–24.

accused.[35] The concerns with the use of these tools do not stop there.

Over the last several years, prominent groups such as Pretrial Justice Institute (PJI) strongly advocated for the broad introduction of these Pretrial Risk Assessment Tools.[36] However, in February 2020, PJI reversed their stance on this position, specifically stating that they "now see that pretrial risk assessment tools, designed to predict an individual's appearance in court without a new arrest, can no longer be a part of our solution for building equitable pretrial justice systems."[37] One week later, Public Safety Assessment, a widely used pretrial risk assessment tool, developed by the Laura and John Arnold Foundation, released a statement in which they clarified that "implementing a [risk] assessment alone cannot and will not result in the pretrial justice goals we seek to achieve."[38]

The perils of these risk-calculation tools and their inter-relatedness is aptly described as part of "a cycle of injustice," in a report by Our Data Bodies, the community-led technology resistance group led by Tawana Petty et al.:[39]

> [T]he collection, storage, sharing, and analysis of data as part of a looping cycle of injustice that results in diversion from shared public resources, surveillance of families and communities, and violations of basic

[35] *See* sources cited *infra* note 79.

[36] *Updated Position on Pretrial Risk Assessment Tools*, PRETRIAL JUST. INST. (Feb. 7, 2020), https://www.pretrial.org/wp-content/uploads/Risk-Statement-PJI-2020.pdf [https://perma.cc/M45P-XWZX].

[37] *Id.*

[38] Madeline Carter & Alison Shames, *APPR Statement on Pretrial Justice and Pretrial Assessment*, ADVANCING PRETRIAL POL'Y & RSCH. (Feb. 2020), https://mailchi.mp/7f49d0c94263/our-statement-on-pretrial-justice?e=a01efafabd [https://perma.cc/Q7EP-ZNEN].

[39] Tawana Petty et al., *Our Data Bodies: Reclaiming Our Data*, OUR DATA BODIES PROJECT, at 1 (June 15, 2018), https://www.odbproject.org/wp-content/uploads/2016/12/ODB.InterimReport.FINAL_.7.16.2018.pdf [https://perma.cc/6W55-JJMC].

> human rights. Connected to the experience of power and powerlessness, the theme of "set-up" concerns [the ways in which] data collection and data-driven systems often purport to help but neglect and fail Angelinos. Interviewees described these set-ups as "traps" or moments in their lives of being forced or cornered into making decisions where human rights and needs are on a chopping board. When using social services to meet basic needs or expecting that a 9-1-1 call in an emergency will bring health and/or safety support into their homes or communities, our interviewees spoke about systems that confuse, stigmatize, divert, repel, or harm. These systems—or the data they require—give people the impression of helping, but they achieve the opposite. They ask or collect, but rarely give, and that leads to mistrust, disengagement, or avoidance. Furthermore, systems perpetuate violent cycles when they are designed to harm, criminalize, maintain forced engagement.[40]

III. The Problems with These Tools and How They are Racially Exacerbated

This section will outline key issues with the suite of tools used throughout the criminal justice cycle. These issues include: opacity, accuracy, lack of clear purpose or evaluation metrics, validation studies, and bias. This is an inexhaustive list of issues with these tools, but outlines how these tools being used simultaneously harm people of color and socioeconomically disadvantaged people. This section samples the technologies used that carry different levels of risks, but this section is not comprehensive. Therefore, although the concerns that will be discussed herein resonate for most predictive technology used throughout the criminal justice cycle, one limitation of this paper will be that there is

[40] *Id.* at 20.

limited discussion of mapping specific concerns onto specific technologies.[41]

A. Opacity

Different laws and norms govern the transparency around: (1) private companies using private software or other technological techniques; (2) public entities using publicly developed software or other technological techniques;-and (3) public entities using privately contracted software or other technical tools.[42] Functionally, people are subject to tools from each of these categories that impact and can compound each other. There is both opacity in fact and opacity in law; while not mutually exclusive, it can be helpful to identify separate strands of the opacity problems.

1. Opacity in fact

Opacity in fact is multifaceted, but represents the dynamic in which people are unaware that a tool is being used on them. An individual might not know if a given camera leads to a database, if that database uses facial recognition software, and/or if that database is shared with the police or a company. A starker version of factual opacity is when there are invisible automated decision-making

[41] See generally Moy, *supra* note 19, for a more in depth exploration of concerns that arise out of the police's use of predictive technology.

[42] Regarding situation (1), there are no known laws limiting these tools. They are market-controlled. Regarding (2) and (3), see generally Hannah Bloch-Wehba, *Access to Algorithms*, 88 FORDHAM L. REV. 1265 (2020) (discussing the role of the Freedom of Information Act and the First Amendment in providing legal support for algorithmic transparency).

systems in applications such as credit scoring,[43] health determinations,[44] and within the criminal justice system.[45]

Without specific requirements for transparency on both private and public forms of automated decision-making, there is significant opacity in fact. Jurisdictions often fail to publish key information about the automated decision-making tools they use within the criminal justice cycle, leaving startling news stories of severe algorithmic harm to fill the gap in knowledge.[46] An example of where it is made obvious that a system is being used, while still not making meaningful disclosures about data collection use and error rates, as well as several other requirements that would be included in algorithmic impact assessments, is in airports that offer facial recognition for airplane boarding.[47]

In other locations, feeds from cameras in public places are later synthesized and analyzed using facial

[43] FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 31 (2015); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. U. L. REV. 1, 17 (2014).

[44] Rebecca Robbins & Erin Brodwin, *An Invisible Hand: Patients Aren't Being Told About the AI Systems Advising Their Care*, STAT (July 15, 2020), https://www.statnews.com/2020/07/15/artificial-intelligence-patient-consent-hospitals/ [https://perma.cc/92ST-9V9H].

[45] *See Liberty at Risk*, *supra* note 10, at Executive Summary.

[46] *See, e.g.*, Kashmir Hill, *Wrongfully Accused By an Algorithm*, N.Y. TIMES (June 24, 2020), https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html [https://perma.cc/K7UL-UF5L]; Julia Angwin et al., *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 [https://perma.cc/7P7E-5ZSL].

[47] *See* Kathryn Steele, *Delta Expands Optional Facial Recognition Boarding to New Airports, More Customers*, DELTA AIR LINES (Dec. 8, 2019, 1:39 AM), https://news.delta.com/delta-expands-optional-facial-recognition-boarding-new-airports-more-customers [https://perma.cc/V9G6-SZ4S].

recognition systems or combined with other datasets.[48] In this situation, you do not know what picture is being captured of you, how that picture is being used, or how accurately it can be matched to a given person. One of the reasons that this can be problematic is borne out with one study that illustrates how misidentification can be dangerous and damaging along racial lines. Amazon's Rekognition algorithm misidentified members of Congress as criminals when running their faces against a criminal database and did so disproportionally for black members.[49] Transparency, here, is not a panacea but a starting point for advocates and community members.[50] The bias of automated decision-making systems will be discussed at length in subsection B.

Another issue is the accuracy rate in risk assessments.[51] One risk assessment published through a

[48] Ayyan Zubair, *Domain Awareness System*, SURVEILLANCE TECH. OVERSIGHT PROJECT (Sep. 26, 2019), https://www.stopspying.org/latest-news/2019/9/26/domain-awareness-system [https://perma.cc/9NWY-N8NL].

[49] Jacob Snow, *Amazon's Face Recognition Falsely Matched 28 Members of Congress with Mugshots*, ACLU (July 26, 2018, 8:00 AM), https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28 [https://perma.cc/4VPN-4Q3T].

[50] *See* Ari Ezra Waldman, *Power, Process, and Automated Decision-Making*, 88 FORDHAM L. REV. 613, 617 (2019) (saying, "[i]ndeed, the very characteristics that make automated decision-making systems so attractive—predictive abilities, complexity, power, and independence—are also what make them so problematic for the rule of law and legal legitimacy. Proposals aimed at making algorithms accountable to the law are attempts to address these problems. And yet. . . the proposals are bandages that ignore the underlying incompatibility between algorithmic decision-making and a society based on normative values like equality and fairness.").

[51] *Validation*, MAPPING PRETRIAL INJUSTICE, https://pretrialrisk.com/the-basics/pretrial-risk-assessment-instruments-prai/validation/ [https://perma.cc/5C65-9WNL] (pointing out that "[a] common statistical way of measuring accuracy and predictive validity is through the 'area under the curve,' or AUC. The AUC score is supposed

public records request was a validation study for pre-trial risk assessment tools which identified that the desired threshold for statistical validity was that of a coin flip or better.[52] Even for these systems that are adopted and given a significant amount of credence without transparency disclosures, audits, impact assessments, or other requirements, even if everything is going according to "plan," many will only accurately predict if someone gets rearrested around 55-65% of the time.[53] If the accuracy rates were more publicly available for specific tools in specific jurisdictions, public oversight might influence more thoughtful procurement.

to show how well the tool balances its correct and incorrect predictions—how often it correctly answers the question at hand (like how 'risky' someone is), and how often it gets the prediction wrong. The closer an AUC score is to 1, the more accurate a tool is said to be. **An AUC score of 0.5 is no better than chance in predicting risk**: a 50/50 shot. Some RATs have AUC scores as low as 0.55, barely more accurate than random chance or a coin toss. Several common tools have scores around 0.65, which is considered 'good' in criminology research but "poor" in other fields; a score of 0.65 means over one third of those judged by these tools are being mislabeled. And unlike many other fields, there is a lack of independent evaluation of these validation studies, which severely limits any claims that pretrial RATs are truly predictive. Sarah Desmarais and Evan Lowder point out that 'demonstrating predictive validity does not equate with research demonstrating implementation success.' Even if a tool is considered highly 'accurate' by these standards, it doesn't mean that RATs are being implemented as intended or in a decarceral or racially unbiased way. The predictions they make are not always accurate, not always listened to even if accurate, and are applied inconsistently and in structurally racist ways.").

[52] Email from Zachary K. Hamilton, Director, Washington State Institute for Criminal Justice, to Doug Koebernick, Nebraska Department of Correctional Services (July 14, 2016, 5:34 PM) [hereinafter Hamilton-Koebernick Email], https://epic.org/EPIC-19-11-08-NEDCS-FOIA-20191112-D-Koebernick-Z-Hamilton-Email.pdf [https://perma.cc/RZ49-8KRR].

[53] *Validation*, *supra* note 51.

In addition to the very real concerns about the details of the use and results of a given system, there are broader trust issues caused by a lack of proactive disclosure from law enforcement entities. The feeling of powerlessness could increase, especially among communities in which these technologies have been shown to have significant error disparities specifically on racial lines.[54]

Legal protections, such as trade secret exemptions in open government laws cover the automated decision-making tools adopted around the U.S. criminal justice system, creating significant legal barriers between citizens and the actions of their government. Minimizing these additional protections for contractors developing these systems may lead to more thoughtful adoption of tools and an increased quality of the systems that is reflective of the serious decisions they help make. In Part IV this article will explore some solutions that can be used by jurisdictions looking to hold contractors and themselves more accountable. One through-line for this and many other aspects of technology regulation, however, is that there needs to be curbing of economic incentives in the short term for the tradeoff of a higher-quality demand in the long term, prioritizing human rights and constitutional rights over adoption of new technologies.

2. Opacity in law

Opacity in law starts with trade secrets and other commercial protection. This is borne out in both open government laws for the general public, and in court for specific defendants already subject to a given tool in a cognizable way.[55] After a significant fight in court, defendants can sometimes gain access to nonpublic

[54] *See* Petty et al., *supra* note 39, at 20; *see also* sources cited *infra* note 79.

[55] Wexler, *supra* note 14, at 1351. *See generally* Bloch-Wehba, *supra* note 42.

information about a Criminal Justice technology through a one-time agreement typically only giving access to the specific defendant via an often expansive protective order.[56] Agreements for protective orders are not guaranteed for defendants, but the practice functionally recognizes the need for access to details about the tools as part of adequate representation.

In addition to concerns of fairness and equity, Natalie Ram has illustrated both criminal due process and confrontation clause concerns associated with secrecy for criminal justice tools.[57] The limits in both scope of information shared and who it is shared with disadvantages the public in their oversight function, as well as the communities most commonly subject to these tools.[58]

There are exemptions in open government laws that extend trade secret protection through explicit statutory language.[59] While state analogues vary in exact wording of what is protected, an exemption from the federal Freedom of Information Act ("FOIA") provides that trade secrets or "information which is (1) commercial or financial, (2)

[56] Bloch-Wehba, *supra* note 42, at 1287.

[57] Natalie Ram, *Innovating Criminal Justice*, 112 NW. U. L. REV. 659, 692 (2018) (saying that "[t]rade secret assertion in the context of criminal justice tools also raises constitutional concerns. The secrecy surrounding the existence, use, and function of criminal justice tools interfere with defendants' and courts' efforts to ensure that the government does not engage in unreasonable searches. Such secrecy is also at least in tension with, if not in violation of, defendants' ability to vindicate their due process interests throughout the criminal justice process, as well as their confrontation rights at trial.").

[58] Hannah Bloch-Wehba, *supra* note 42 at 1272 (saying, "compromises between the private vendors' commercial interests and the liberty interests of those affected by algorithmic governance overlook the public's separate and independent interest in oversight and monitoring of government decision-making.").

[59] *See generally Open Gov't Guide*, *supra* note 13 (showing freedom of information laws in 50 states, including the trade secret/commercial protections).

obtained from a person, and (3) privileged or confidential" is exempted.[60]

In *Food Marketing Institute v. Argus*, the trade secret exemption was expanded, reversing decades of precedent requiring a showing of competitive harm if the "trade secret" were to be released under FOIA.[61] Instead, now the entity must only prove either that it (1) treats a piece of information as confidential or, (2) if it is the type of information that is usually kept confidential, that there is either express or implied assurance by the government that it will maintain confidentiality.[62] The Department of Justice issued guidance after the *Food Marketing Institute* decision and updated practitioner guidance.[63] When the government has not made any "express or implied indications at the time the information was submitted that the government would publicly disclose this information," there is a presumption of valid trade secrecy if the entity customarily held the information as private.[64]

To varying extents, state open government laws across the country have similar commercial protections for trade secrets.[65] The justifications of trade secret protection

[60] 5 U.S.C. § 552(b)(4) (2016); ELEC. PRIV. INFO. CTR., LITIGATION UNDER THE FEDERAL OPEN GOVERNMENT LAWS 113 (Marc Rotenberg et al. eds., 25th ed. 2010) (citing Pub. Citizen Health Rsch. Grp. v. F.D.A., 185 F.3d 898, 903 (D.C. Cir. 1999); Nat'l Parks & Conservation Ass'n v. Morton (I), 498 F.2d 765, 766 (D.C. Cir. 1974); Brockway v. Dep't of Air Force, 518 F.2d 1184,1188 (8th Cir. 1975)).

[61] Food Mktg. Inst. v. Argus Leader Media, 139 S. Ct. 2356, 2361 (2019).

[62] *Step-by-Step Guide for Determining if Commercial or Financial Information Obtained from a Person is Confidential Under Exemption 4 of the FOIA*, U.S. DEP'T OF JUST., https://www.justice.gov/oip/step-step-guide-determining-if-commercial-or-financial-information-obtained-person-confidential[https://perma.cc/F2BM-H4TL].

[63] *Id.*

[64] *Id.*

[65] *See generally Open Gov't Guide*, REPS. COMM. FOR FREEDOM PRESS, https://www.rcfp.org/open-government-guide [https://perma.cc/5R42-B3ZA].

including competition, innovation, and labor ownership,[66] should be weighed against the interests at stake. Trade secrets being applied to the criminal setting is a relatively recent legal development, but it is becoming more common.[67] Preservation of commercial viability and promoting commercial innovation for policing tools and tools that directly affect bail and sentencing decisions risks people's liberty while maximizing profit and minimizing accountability. In criminal cases, as Rebecca Wexler explains, civil trade secret protection applied to criminal cases is dangerous because it "will almost certainly lead to systemic overclaiming and wrongful exclusion of relevant evidence; impose an unreasonable burden on defendants' discovery and subpoena rights; and undermine the legitimacy of criminal proceedings by implying that the government values intellectual property owners more than other groups affected by criminal proceedings."[68]

In the open government context, the evidentiary mechanisms are not present like they are in criminal cases. But, as Hannah Bloch-Wehba articulates, they "codif[y] expectations regarding the government's disclosure of information to the *public*, … [and] operat[e] both to protect the balance of power between the public and the government and to ensure that key information regarding government decision-making is open to public scrutiny."[69] Open government laws provide a right to government records without having to be personally affected by a tool, and policies about disclosure exemptions should reflect that. In applications where documents related to automated

[66] *See generally* Michael Risch, *Why Do We Have Trade Secrets?*, 11 MARQ. INTELL. PROP. L. REV. 1, 26–37 (2007) (explaining the justifications for trade secrets).

[67] *See* Rebecca Wexler, *supra* note 14 at 1388–94 (summarizing the history of the trade secret privilege in criminal proceedings).

[68] *Id.* at 1395.

[69] Bloch-Wehba, *supra* note 42, at 1268.

decision-making tools used by the government to help make decisions around bail, policing resources, parole and investigations are statutorily exempt from public view, the interest in liberty and public scrutiny outweighs the interest in competition or innovation.

B. Performance Issues – Accuracy and Bias

Inaccuracy plagues many forms of predictive automated decision-making. Many pre-trial risk assessment tools use the measure of about 55– 65% predictive validity, barely better than the chance of a coin flip outcome, when trying to predict who will be arrested again.[70] Those tools are trying to predict "criminality," a concept that cannot be clearly defined nor predicted, posing additional accuracy challenges. For example, a public records request to the Nebraska Department of Corrections yielded emails between a developer of a risk assessment algorithm and administrators in the Department of Corrections where the developer explained that the rate at which they test if something is statistically valid is if it is more than 50% likely to be accurate.[71]

> The primary criterion for creating a validated tool to improve the prediction of recidivism beyond random chance (i.e. a coin flip). . . one should not simply be concerned that the tool improves beyond random chance but that its prediction is more accurate than any other tool under consideration. Again, I cannot argue that the YLS/CMI has been identified to provide a better prediction than random chance in more places than any other tool. However, we attempted to create the STRONG-R to be more accurate than the

[70] *Validation*, *supra* note 51; *see* Hamilton-Koebernick Email, *supra* note 52.
[71] Hamilton-Koebernick Email, *supra* note 52.

> YLS/CMI and to customize the prediction for the specific population it is being used to assess.[72]

This is under-acknowledged by the entities adopting the tool and, often by design, is unknown by those affected by the tool. A lack of widespread regulations requiring regular, independently done validation studies on the population that it is used on combined with the limited access to this data due to commercial protections yield only limited knowledge of accuracy rates.[73] Transparency can help dispel the myth that automated decision-making systems, just because they use computers or are based off of statistical analysis, are more accurate or useful than other tools or strategies to make the justice system more equitable.

Validation studies will be covered under section IV.c, but, as a start, validation studies are processes by which statistical analysis is done to evaluate a given automated decision-making system to check its predictive validity by comparing predictions against actual outcomes in a given jurisdiction.[74] Mapping Pretrial Injustice surveyed jurisdictions using pretrial risk assessment tools about their validation practices and found that 21% of the jurisdictions performed validation checks 5-10 years ago, 21% of the validation checks used nonlocal data, 9% of the checks used validation studies from over 10 years ago, and only 28%

[72] *Id.*

[73] *See Validation,* supra note 51 (finding that "[m]any jurisdictions are using tools that have not been validated with their local population or have not been validated at all."). One example of a regulation addressing validation is §2(e) Miss. H.B. 585 (2014), https://www.mdoc.ms.gov/Documents/House%20Bill%20585%20as%20approved%20by%20the%20Governor.pdf [https://perma.cc/U87Y-3MSL] (stating, "'Risk and needs assessment' means the use of an actuarial assessment tool validated on a Mississippi corrections population to determine a person's risk to reoffend and the characteristics that, if addressed, reduce the risk to reoffend.").

[74] *Validation*, supra note 51.

used validation studies with local data within 5 years.[75] Using local data to validate a tool is important because, due to different populations, different police forces, different crime histories, and different goals—a factor that predicts criminal risk in one jurisdiction may not accurately predict the same measure in another. Performing proper validation studies frequently can help track whether the tool used by a given jurisdiction is actually helping to achieve its goals. For many validation studies, it is also key to point out that they have thresholds as low as a 55% accuracy[76] benchmark to determine accuracy—a worrisomely low bar.

Evaluating bias is a necessary complement to the evaluation of accuracy and other metrics of a given system. Even when a system is technically accurate, systems can encode or reinforce systemic biases or be inherently dangerous systems. One example of inaccuracy and bias is an analysis of a pre-trial risk assessment tool, done by ProPublica in 2016. The analysis showed that nearly twice as many black defendants were labeled as high risk to reoffend, but did not actually reoffend, as white defendants.[77] The inverse was also true—twice as many white defendants were labeled low risk but ended up reoffending compared to black defendants.[78] Several other studies of risk assessment tools, as well as predictive policing tools, have shown disparate scoring and ineffectiveness based on ethnicity, age, zip code, and more.[79]

[75] *Id.*

[76] *Id.*

[77] Angwin et al., *supra* note 46. 23.5% of white defendants were labeled higher risk but did not re-offend. 44.9% of black defendants were labeled higher risk but did not re-offend. *Id.*

[78] *Id*. 47.7% of white defendants were labeled lower risk but did re-offend. 28% of black defendants were labeled lower risk but did re-offend. *Id.*

[79] *See, e.g,* Megan T. Stevenson, *Assessing Risk Assessment in Action*, 103 MINN. L. REV. 303, 329 (2019); Melissa Hamilton, *The Biased Algorithm: Evidence of Disparate Impact on Hispanics*, 56 AM. CRIM.

The risk of bias is not limited to risk assessment-type tools. A study from the National Institute of Standards and Technology ("NIST") analyzed the facial recognition algorithms of a "majority of the industry" and found that some software was up to 100 times more likely to return a false positive of a non-white individual than it was for white individuals.[80] Specifically, NIST found "for one-to-many matching, the team saw higher rates of false positives for African American females," which they highlight "are particularly important because the consequences could include false accusations."[81] As of now, a well-funded and powerful entity like NIST has to both choose to do a study like this and be limited to systems they have access to. Audits or validation studies are often done by the company itself or an outside tester that they hire, without independent evaluation, resulting in conflicts of interest.[82] This conflict

L. REV. 1553, 1560–61 (2019); Megan T. Stevenson & Christopher Slobogin, *Algorithmic Risk Assessments and the Double-Edged Sword of Youth*, 96 WASH. U. L. REV. 681, 681 (2018); Songül Tolan et al., *Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia*, INT'L CONF. ON AI AND L. (2019), https://chato.cl/papers/miron_tolan_gomez_castillo_2019_machine_learning_risk_assessment_savry.pdf [https://perma.cc/9WJC-W3T9]; Will Douglas Heaven, *Predictive Policing Algorithms are Racist. They Need to be Dismantled*, MIT TECH. REV. (Jul. 17, 2020), https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/ [https://perma.cc/Q79W-BQM6].

[80] *NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software*, NAT'L INST. STANDARDS & TECH. (Dec. 19, 2019), https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software [https://perma.cc/L7S7-62JK].

[81] *Id.*

[82] *See* Mona Sloane, *The Algorithmic Auditing Trap*, ONEZERO (Mar. 17, 2021) https://onezero.medium.com/the-algorithmic-auditing-trap-9a6f2d4d461d [https://perma.cc/4G2E-ZUNY]; *Validation*, *supra* note 51 ("And unlike many other fields, there is a lack of independent evaluation of these validation studies, which severely limits any claims

arises in part because it is against a company's interest to publish information illustrating that their software is inaccurate or biased. The legal and practical infrastructure supporting the lack of transparency in these systems directly obfuscates the opacity and bias in these systems. With more transparency, more robust, independent testing can be done, leading to increased oversight.

C. Process Issues

In addition to the performance issues of risk assessment tools and other criminal justice automated decision-making systems, there are significant process issues in the procurement and execution process that enable and exacerbate the negative effects articulated above. Process issues result in a lack of transparency around who is developing a tool, the stated purpose of a tool, input data, logic of a tool, decision-making matrix, and data sharing and retention policies.[83]

Procurement regulations differ greatly between states and regulate how governments contracts services.[84] As of now, without complementary regulation of automated

that pretrial RATs are truly predictive."); *see also, e.g.,* Northpointe, *Results from a Psychometric Study Conducted for the Wisconsin Department of Corrections Division of Adult Institutions*, EPIC (Feb. 11, 2014), https://epic.org/EPIC-19-11-08-NEDCS-FOIA-20191112-Northpointe-Self-Validation.pdf [https://perma.cc/KVP5-HH2M]; Northpointe, *Predictive Validity of the COMPAS Reentry Risk Scales*: *An Outcomes Study Conducted for the Michigan Department of Corrections: Updated Results on an Expanded Release Sample*, EPIC (Aug. 22, 2013), https://epic.org/EPIC-19-11-08-NE-DCS-FOIA-20191112-Northpointe-Self-Validation-2.pdf [https://perma.cc/9ZBU-ANG7].

[83] *See Liberty at Risk*, *supra* note 10, at 15.

[84] See generally *2018 Survey of State Procurement Practices*, NAT'L ASS'N STATE PROCUREMENT OFFS. (2018), https://www.naspo.org/wp-content/uploads/2019/12/2018-FINAL-Survey-Report_6-14-18.pdf [https://perma.cc/X3HE-XLLB].

decision-making, procurement processes can be levied to meet the current need for transparency and oversight. It is a field ripe for updating given the increasing automation of the administrative state.[85] It is through the procurement process that agreements with contractors that give this level of deference are accepted, where hundreds of thousands of dollars are spent on a given automated decision-making system, and where simple changes can be made to ensure transparency and other forms of public oversight.

Particularly, the lack of a requirement for the entity adopting an automated decision-making tool to articulate the purpose for adopting the tool, benchmarks for evaluation of the effectiveness of a tool, and regular, independent, and localized evaluation and validation studies of the purpose of a tool diminishes thoughtful procurement and accountability.[86]

In terms of data privacy and security, minimum baseline standards and policies should be instituted that can help limit the improper sale of data, introduce safeguards for accuracy, require data collection and use to be directly proportional to the needs of a system, and empower an oversight body.

Additional opportunities to improve process issues come in creating rights for people to understand exactly what data of theirs is being used in a given system, what system is being used against them, and how they can contest and understand an algorithmically supported decision that might be erroneous. This model is used in a limited way for consumer credit reporting in the U.S.[87] and more generally

[85] *See* ENGSTROM ET AL., *supra* note 17, at 9–10; *Liberty at Risk*, *supra* note 10, at Executive Summary.

[86] *See Liberty at Risk*, *supra* note 10, at 15.

[87] Fair Credit Reporting Act, 15 U.S.C. §§ 1681–1681x (providing rights to examine credit reports, dispute incomplete or inaccurate information, and giving obligations to consumer reporting agencies to delete inaccurate information).

in Europe through the General Data Protection Regulation,[88] and should be a minimum for people subject to these systems throughout the criminal justice cycle. Targeted surveillance oversight laws, discussed more in Section IV, is one path towards more transparency and accountability.

IV. HARM-MITIGATION APPROACHES

Jurisdictions should improve opacity, alleviating some of the harms discussed above, by constraining protections in open government regulations for trade secret and commercial protections. Additionally, while sweeping algorithmic transparency and accountability bills are currently difficult to pass in legislatures,[89] this section explores different targeted improvements that legislators can make as well as how administrators can improve the quality of procurement for many government-contracted criminal justice technologies.

A. Constrain trade secret protections within open government laws

For automated decision-making systems in and around the criminal justice system, exemptions within the trade secret and other commercial protections in open

[88] *See generally* Regulation 2016/679, of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), §§ 1–5, 2016 O.J. (L 119) 1, 39–47.

[89] For bills that did not get passed, see, e.g., Algorithmic Accountability Act of 2019, H.R. 2231, 116th Cong. (2019); An Act Relating to Establishing Guidelines for Government Procurement and Use of Automated Decision Systems in Order to Protect Consumers, Improve Transparency, and Create More Market Predictability, Wash. S.B. 5116, 67th Legislature (2021) (Wash.), https://app.leg.wa.gov/billsummary?BillNumber=5116&Initiative=false&Year=2021 [https://perma.cc/5YLQ-6XG9].

government statutes should be expressly given. This allows jurisdictions to take a risk-based approach, not tearing down trade secrets for all purposes, but giving the public the power to oversee important decisions made by their government. At the federal level, Congress should pass legislation increasing amend FOIA following the expansion of trade secret protections in *Food Marketing Institute*[90] to increase citizen access.

B. Procurement policies and decisions

Governments contracting with the companies that develop these tools should procure more transparently and purposefully, raising the standard of quality, accuracy, and disclosure. Following a report that the CEO of a surveillance company contracting with the state of Utah had ties to the Ku Klux Klan,[91] the state auditor released a set of recommended guidelines for the procurement or development of software for the state.[92] The "Software Application Procurement Principles for Utah Government Entities" include:

(1) "Limit Sharing of Sensitive Data";

(2) "Minimize Sensitive Data Collection and Accumulation";

(3) "Validate Technology Claims – including Capability Review[,]" particularly "Asserted use of AI

[90] *See* Food Mktg. Inst., 139 S. Ct. at 2361.

[91] Matt Stroud, *CEO of Surveillance Firm Banjo Once Helped KKK Leader Shoot Up a Synagogue*, ONEZERO (Apr. 28, 2020), https://onezero.medium.com/ceo-of-surveillance-firm-banjo-once-helped-kkk-leader-shoot-up-synagogue-fdba4ad32829 [https://perma.cc/W85A-6MVN].

[92] *Application Procurement Principles for Utah Government Entities*, OFF. STATE AUDITOR (Feb. 1, 2021) (Utah), https://auditor.utah.gov/wp-content/uploads/sites/6/2021/02/Office-of-the-State-Auditor-Software-Application-Procurement-Principles-Privacy-and-Anti-Discrimination-Feb-2-2021-Final.pdf [https://perma.cc/Q8GJ-AVWF].

[Artificial Intelligence)] or ML [Machine Learning)]. . . proposed use of disparate data sources, especially social media or integration of government and private sources, and. . . Real-time capabilities. . .";

(4) "Perform In-Depth Review of. . . Algorithms";

(5) "Review Steps Taken to Mitigate Discrimination";

(6) "Determine Ongoing Validation Procedures"; and

(7) "Require Vendor to Obtain Consent of Individuals Contained with Training Datasets. . . ."[93]

This list is a starting point of how contracting agencies can procure with higher standards and help to protect their citizens, regardless of if these steps are required by law. According to an analysis of surveillance oversight laws throughout the country, ten of the sixteen jurisdictions surveyed require surveillance tools to be approved through an oversight process, but many empower communities and increase the levels of transparency.[94] Not all of the automated decision-making systems used throughout the criminal justice cycle would be covered by surveillance technology laws or the principles proposed by the Utah auditor, but it's an important blueprint when beginning to regulate procurement and transparency into automated decision-making.

C. Targeted legislation about specific technologies

In March 2019, Idaho became the first state to enact a law specifically promoting transparency, accountability,

[93] *Id.*

[94] Rebecca Williams, *Everything Local Surveillance Laws Are Missing in One Post*, HARV. KENNEDY SCH.: BELFER CTR. FOR SCI. & INT'L AFFS. (Apr. 26, 2021), https://www.belfercenter.org/publication/everything-local-surveillance-laws-are-missing-one-post [https://perma.cc/WWF7-ZU68].

and explain-ability in pre-trial risk assessment tools.[95] The Idaho law prevents trade secrecy or IP defenses in criminal cases, requires public availability of "all documents, data, records, and information used by the builder to build or validate the pretrial risk assessment tool," and empowers defendants to review all calculations and data that went into their risk score.[96]

Other direct approaches simply respond to a particularly problematic technology by banning or placing a moratorium on its use. For example, over a dozen jurisdictions in the United States have banned face recognition for one purpose or another. There have been bans or moratoriums on the use of Facial Surveillance systems in Alameda, CA; Berkeley, CA; Boston, MA; Brookline, MA; Cambridge, MA; Jackson, MS; Northampton, MA; Oakland, CA; Portland, ME; Portland, OR; San Francisco, CA; Somerville, MA; and Springfield, MA.[97] Although important for certain technologies, it forces jurisdictions to continually play catch-up and many only regulate police use. In order to increase the likelihood of enforcement, significant penalties and a private right of action for violations should be included in any bans or moratoriums on specific technology.

Another, although highly imperfect, strategy towards achieving transparency around some automated decision-making systems used by the government is to utilize a task force or commission in a government entity specifically set up to understand how an automated decision-making system is used throughout the state. These exist in Alabama, Vermont, New York State, and New York City, among

[95] IDAHO CODE § 19-1910 (2019).

[96] *Id.*

[97] *Map*, BAN FACIAL RECOGNITION, https://www.banfacialrecognition.com/map/ [https://perma.cc/7TSK-EGPA].

others.[98] The results of these task forces are more likely to alleviate the factual opacity of what systems are being used, rather than to publish details such as what factors are used in a given tool that are important to understand and make sure are correct when an automated decision-making system is used.[99] If sufficiently empowered, task forces can be an important transparency function for public automated

[98] *See* S.J. Res. 71 (May 15, 2019) (Ala.), http://alisondb.legislature.state.al.us/ALISON/SearchableInstruments/2019RS/PrintFiles/SJR71-int.pdf [https://perma.cc/B8NX-PNCX]; H. 378 (May 21, 2018) (Vt.), https://legislature.vermont.gov/Documents/2018/Docs/ACTS/ACT137/ACT137%20As%20Enacted.pdf [https://perma.cc/SE9N-NU44]; S. 3971B (Feb. 22, 2019) (N.Y.), https://www.nysenate.gov/legislation/bills/2019/s3971 [https://perma.cc/DZH2-PD2Z]; N.Y.C. Local L. 49 (Jan. 11, 2018) (N.Y.C.) https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0# [https://perma.cc/VEP2-HGNZ].

[99] *See Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force,* AI NOW INST., at 94 (Dec. 2019), https://ainowinstitute.org/ads-shadowreport-2019.pdf [https://perma.cc/U4Z5-8GY4 (saying, "New York City Automated Decision Systems Task Force members repeatedly requested information about ADS currently used because the local context was necessary to fulfill the statutory mandate, but many agencies resisted cooperating or only provided selective information about one system. To avoid similar problems, similar government bodies or processes must be given authority to request and access information about all existing ADS, without special exemptions or carveouts that can undermine necessary analysis and subsequent recommendations. While it may be difficult for a task force or government process to undertake a thorough analysis of each ADS system, a task force or government process should be empowered to select representative ADS that reflect the variety of ways these systems can impact human welfare." This illustrates the fact that even being able to discover and publish the uses of automated decision-making systems by the government is not something these task forces are routinely empowered to do effectively.).

decision-making and prime legislators for broader regulatory proposals.[100]

D. Assessments and audits: alternatives to blanket transparency of source code

One approach that does not require voluminous source code and other developmental documents to be made fully public is to require some sort of required assessment or audit that is designed to ensure key aspects of an automated decision-making system are considered, purposeful, and public. One potential benefit to this approach is that it could solve some transparency and accountability issues while allowing IP holders to avoid disclosure of a large swath of their source code and other proprietary information.

The content of assessments varies, mostly because they are not very widely deployed yet. However, one widely deployed assessment is for public entities in Canada.[101] The assessment guides users through questions about why they are adopting a given system, what capabilities their system holds, how explainable it is, what kind of decisions it helps make, how much intervention is involved, how sensitive their data is, how synthesized the data is, who the adopting agency is consulting about the adoption, mitigating measures, procedural fairness, and more.[102] Depending on

[100] *See, e.g.*, N.Y. S. A6042, 2020-2021 Legis. Session (2021) (N.Y.), https://www.nysenate.gov/legislation/bills/2021/A6042 [https://perma.cc/PGR7-T685]; Vt. H. 263 (2021) (Vt.), https://legislature.vermont.gov/bill/status/2022/H.263 [https://perma.cc/72AZ-EGQD]. These illustrate strong proposed bills in jurisdictions following the establishment and proceedings of task forces.

[101] *See Directive on Automated Decision-Making*, GOV'T CAN. (last modified May 2, 2019), https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592 [https://perma.cc/YFL7-RSD3].

[102] *Algorithmic Impact Assessment*, GOV'T CAN. (last modified Mar. 22, 2021), https://open.canada.ca/aia-eia-js/?lang=en [https://perma.cc/9MKL-MXE8].

the results, the agency is required to take mitigating measures, provide more information, or use a different system.[103]

For this option to maximize trust and accountability, assessments should be mandatory, robust, public, and part of an infrastructure that legitimizes it. A landmark report of how Algorithmic Impact Assessments can be operationalized describes a robust process of requiring a pre-acquisition review, initial agency disclosure requirements, comment period, due process challenge period, and renewal.[104] In this report, AI Now addresses the trade secret barrier for assessments by saying:

> While there are certainly some core aspects of systems that have competitive commercial value, it is unlikely that these extend to information such as the existence of the system, the purpose for which it was acquired, or the results of the agency's internal impact assessment. Nor should trade secret claims stand as an obstacle to ensuring meaningful external research on such systems. AIAs provide an opportunity for agencies to raise any questions or concerns about trade secret claims in the pre-acquisition period, before entering into any contractual obligations. If a vendor objects to meaningful external review, this would signal a conflict between that vendor's system and public accountability.[105]

[103] *See id.*; *Framework for the Management of Compliance*, GOV'T CAN. (last modified Aug. 27, 2010), https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=17151&section=html at 9.10-9.12 (explaining the range of consequences of not complying with certain government directives, including *Directive on Automated Decision-Making*) [https://perma.cc/XP67-6CVW].

[104] Dillon Reisman et al., *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, AI NOW INST. (Apr. 2018), at 7–10, https://ainowinstitute.org/aiareport2018.pdf [https://perma.cc/8Z6U-QVYV].

[105] *Id.* at 14.

Requiring assessments in this regular, overseen fashion could be a helpful alternative to the status quo without jeopardizing substantial competitive harm, since they would require disclosure of basic operations about their tools and the impacts they cause, rather than the source code.

E. The value and limits of transparency

Transparency is not, itself, the end goal for advocates trying to ensure equity in systems used by governments and corporations. However, transparency can go hand-in-hand with accountability, and, without any transparency, the hope for change is depleted. Without knowledge, citizens and advocates cannot exercise their rights in the democratic environment effectively because they do not have the resources to understand how the automated decision-making systems might be affecting them or their communities. Improving transparency can help alleviate the outweighed pressure and negative effect of automated decision-making systems on communities of color. It can allow third-party researchers to test datasets as well as the algorithms themselves to expose inequities.

Automated decision-making systems require an entity to articulate which factors they want to include in helping determine outcomes like how likely someone is going to be arrested again, receive a second interview for a job, and more.[106] These outcomes are determined by entities now adopting automated decision-making tools to help automate the continued determination of those outcomes. Adopting a tool and passing the burden of justifying complicated decisions to a third-party contractor who can

[106] *See* Miranda Bogen, *All the Ways Hiring Algorithms Can Introduce Bias*, HARV. BUS. REV. (May 6, 2019), https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias [https://perma.cc/736N-CTBB]; Heaven, *supra* note 79; *see generally Liberty at Risk*, *supra* note 10.

protect any disclosure with the overbroad commercial and trade secret protections in open government laws leaves people with little hope for recourse.[107] Each decision about what factor is included in a tool, and how much weight it will hold is a decision that is not simple and is not objective. Increasing the transparency about what automated decision-making systems are used by their government and how they work is necessary for public engagement and input about how their government is operating. The reason that this is particularly salient for the uses of automated decision-making in the criminal justice system is because the stakes are extremely high, where an automated decision-making system influences the length of a prison sentence or the likelihood of a police encounter.

Although transparency is not a panacea and will not stop either the use of these tools by themselves or the harm they cause, it is helpful to have identifiable legal and organizational forces that can be improved upon.

V. CONCLUSION

There is a wide variety of automated decision-making tools used by government entities and corporations alike which operate in and around the criminal justice system in the United States. There is a huge variety of the type, quality, and frequency of these tools, but they all hold immense power. The tools discussed in this piece have an outsized impact on communities of color and communities that are lower income. Transparency in this particular field is very elusive, which is especially damaging to communities who burden the harm most. Trade secrets and

[107] *See, e.g.,* State v. Loomis, 881 N.W.2d 749, 761–62 (Wis. 2016) (stating, "[a]dditionally, this is not a situation in which portions of a PSI are considered by the circuit court, but not released to the defendant. The circuit court and Loomis had access to the same copy of the risk assessment.").

other commercial protections included in open government laws, combined with a lack of procurement regulations, contribute to this and must be changed.

Moving forward, a more transparent approach towards adoption of automated decision-making tools can allow equity to be built, and more thoughtful adoption of these tools to take hold. By minimizing commercial protections for the details concerning these tools as part of a suite of improvements around this set of very sensitive government uses, jurisdictions can prioritize the health, safety, and equity of their citizens over supporting a criminal justice technology "industry."