AI AND THE ADMINISTRATION OF JUSTICE IN THE UNITED STATES OF AMERICA: PREDICTIVE POLICING AND PREDICTIVE JUSTICE

By Emily Silverman *

1 Predictive Policing

1.1 National practices

1.1.1 Definition of 'predictive policing' and introductory remarks

There seem to be no legal definitions of predictive policing in federal- or state-level legislation in the United States. There are, however, legal definitions in a number of local ordinances, including Santa Cruz, California; Pittsburgh, Pennsylvania; New Orleans, Louisiana; and Oakland, California.

In Santa Cruz, California, 'Predictive Policing Technology' is defined in a city ordinance enacted in 2020 as

> [S]oftware that is used to predict information or trends about crime or criminality in the past or future, including but not limited to the characteristics or profile of any person(s) likely to commit a crime, the identity of any person(s) likely to commit crime, the locations or frequency of crime, or the person(s) impacted by predicted crime.[1]

As of 2020, Pittsburgh, Pennsylvania, defines predictive policing technology as

> Any fully or partially-automated computational application of programs, devices, hardware, or software based on machine learning or artificial intelligence that is, independent of a user, used to predict information or trends on crime or criminality that has or has yet to occur, including, but not limited to, the characteristics or profile of any individual(s) likely to commit a crime, the identity of any individuals likely to commit crime, the locations or frequency of crime, or the individuals affected by predicted crime or criminality.[2]

New Orleans, Louisiana, defines predictive policing technology since 2021 as

* Emily Silverman, JD (Berkeley Law), LLM, Senior Researcher, Max Planck Institute for the Study of Crime, Security and Law (Freiburg im Br/Germany), e.silverman@csl.mpg.de.

[1] Santa Cruz, Cal, City Ord. § 9.85.020(C) (Ord No 2020-17, 11 August 2020).

[2] Pittsburg, Pa, City Ord § 116.15(b)(3) (Ord No 31-2020, § 1, eff 24 September 2020).

> [T]he usage of predictive analytics software in law enforcement to predict information or trends about criminality, including but not limited to the perpetrator(s), victim(s), locations or frequency of future crime. It does not include, for example, software used to collect or display historic crime statistics for informational purposes.[3]

In Oakland, California, the following definition was added to the city's municipal code in 2021:

> 'Predictive Policing Technology' means computer algorithms that use preexisting data to forecast or predict places or times that have a high risk of crime, or individuals or groups who are likely to be connected to a crime. This definition does not include computer algorithms used solely to visualize, chart, or map past criminal activity (e.g. heat maps).[4]

These and similar AI-based systems are or have been used primarily by the police. It is possible that these types of predictive technology will migrate into the national security sphere as part of the military's effort to predict who and where its enemies are. In 2018, however, there was no evidence that the military was seeking inspiration from law enforcement algorithms.[5] The policy incentives for using AI-based systems would seem to be straight forward: preventing crime and reducing costs.[6] There is no evidence in the literature that investments in these systems are justified on the basis of a (perceived) need to support high-tech industry.

1.1.2 Selected AI-based systems used for predictive policing

Numerous AI-based predictive policing systems are or have been in use in the United States. This report will concentrate on three of them: PredPol, HunchLab, and CivicScape.

PredPol

The PredPol tool was the result of a research collaboration conducted by the University of California, Los Angeles (UCLA) and the Los Angeles Police Department (LAPD).[7] The project received funding from the Federal Bureau of Justice Assistance.[8] Launched as a

---

[3] New Orleans, La, Code of Ordinances, § 147-1 (Ord No 28559, 1 January 2021).
[4] Oakland, Cal, Municipal Code § 9.64.010(11) (Ord No 13635, 9 February 2021).
[5] Ashley S Deeks, 'Predicting Enemies' (2018) 104 Va L Rev 1529, 1534 fn 16.
[6] See, eg, Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1109, 1124-25.
[7] Mark Puente, 'Crime Forecast Software Is Losing Its Luster' *Los Angeles Times* (Los Angeles, 8 July 2019) A1; PredPol, 'Overview' (*PredPol*) <www.predpol.com/about/> accessed 16 May 2023.
[8] Tim Lau, 'Predictive Policing Explained' (*Brennan Center,* 1 April 2020) <www.brennancenter.org/our-work/research-reports/predictive-policing-explained> accessed 5 May 2023.

company around 2012,[9] PredPol's proprietary algorithm is based on a 'near-repeat' machine-learning model that proceeds on the basis of the assumption that if a crime occurs at a given location, the immediate surroundings are at increased risk for future criminal activity.[10] The place-focused (geospatial) model, first developed by anthropologist Jeffrey Brantingham and mathematician George Mohler,[11] is an extrapolation of an algorithm used to predict the distribution of earthquake aftershocks.[12] It processes three years' worth of data, whereby more recent data are weighted more heavily. The algorithm generates 500 by 500 square foot predictive boxes on maps that indicate areas where particular crimes are most likely to occur.[13] PredPol uses historical event datasets to train the algorithm. It uses only three data points, crime type, crime location, and crime time[14]; it does not use any personally identifiable information.[15] The only inputs in PredPol's system are incident records.[16] The name of the company was changed in 2021 to Geolitica, a mashup of 'geographical analytics'. The name change was undertaken because the word 'predictive' was seen to be inadequate and because the software was 'predictionless'.[17]

HunchLab

HunchLab, also a geospatial predictive policing software, was developed with federal grant support by Philadelphia-based start-up company Azavea and criminal justice professors Jerry H Ratcliffe and Ralph B Taylor (Temple University) and Joel Caplan and

---

[9] Ali Winston and Ingrid Burrington, 'A Pioneer in Predictive Policing Is Starting a Troubling New Project' (*The Verge*, 26 April 2018) <www.theverge.com/2018/4/26/17285058/predictive-policing-predpol-pentagon-ai-racial-bias> accessed 16 May 2023.

[10] Aziz Z Huq, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043, 1070. For a definition of 'near repeat effect', see Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1109, 1128.

[11] Claudia Aradau and Tobias Blanke, 'Decision' in Claudia Aradau and Tobias Blanke (eds), *Algorithmic Reason: The New Government of Self and Other* (OUP, 2022) 42-43.

[12] Aziz Z Huq, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043, 1068-70 (crediting anthropologist Jeffrey Brantingham and mathematician Andrea Bertozzi).

[13] Sarah Brayne, Alex Rosenblat, and danah boyd, 'Predictive Policing' (Workshop primer presented at annual conference of Data & Civil Rights: A New Era of Policing and Justice, Washington, DC, 27 October 2015) <www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf> accessed 17 May 2023.

[14] Sarah Brayne, Alex Rosenblat, and danah boyd, 'Predictive Policing' (Workshop primer presented at annual conference of Data & Civil Rights: A New Era of Policing and Justice, Washington, DC, 27 October 2015) <www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf> accessed 17 May 2023.

[15] PredPol, 'How Predictive Policing Works' (*PredPol*) <www.predpol.com/how-predictive-policing-works/> accessed 5 May 2023.

[16] Andrew Guthrie Ferguson, 'Predictive Policing Theory' in Tamara Rice Lave and Eric J Miller (eds), *The Cambridge Handbook of Policing in the United States* (Cambridge Univ Press 2019) 494 (refers to PredPol website).

[17] Claudia Aradau and Tobias Blanke, 'Decision' in Claudia Aradau and Tobias Blanke (eds), *Algorithmic Reason: The New Government of Self and Other* (OUP, 2022) 51.

Les Kennedy (Rutgers University).[18] The first version of HunchLab was on the market in 2008.[19] The software applies machine learning algorithms and includes weather patterns.[20] More specifically, HunchLab's

> 'ensemble machine learning' algorithm uses 'temporal cycles' (day of week, seasonality); 'weather'; 'risk terrain modeling' (locations of bars, bus stops, etc.); 'socioeconomic indicators'; historic crime levels; and near-repeat patterns as a means of predicting individual crime expectations across the jurisdiction.[21]

HunchLab was sold to ShotSpotter in 2018 and renamed ShotSpotter Missions.[22]

CivicScape

CivicScape was launched as a startup in spring 2017 by Brett Goldstein, a former officer of the Chicago Police Department, following a technology transfer out of the University of Chicago.[23] By summer 2017, nine cities were either using the software or in the process of implementing it.[24] The cities of Camden and Linden, New Jersey, and Dearborn, Michigan, for example, were among the early users.[25] CivicScape, which is funded by Ekistic Ventures, applies predictive analytics to policing, using artificial intelligence and neural networks.[26] It uses crime-pattern data, federal weather information, and 311 call

---

18 Azavea, *HunchLab: Under the Hood* (Azavea 2014), available at <www.brennancenter.org/sites/default/files/NYC_0002305_HunchlabPromoOverview%20-%20Copy.pdf> accessed 11 May 2023.

19 Bilel Benbouzid, 'To Predict and to Manage. Predictive Policing in the United States' (2019) 6 Big Data & Soc'y 1.

20 Gwenola Ricordeau, 'Predictive Policing' in Philip L Frana and Michael J Klein (eds), *Encyclopedia of Artificial Intelligence: The Past, Present, and Future of AI* (ABC-CLIO 2021) 271.

21 Aziz Z Huq, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043, 1069-70. See also Sarah Brayne, Alex Rosenblat, and danah boyd, 'Predictive Policing' (Workshop primer presented at annual conference of Data & Civil Rights: A New Era of Policing and Justice, Washington, DC, 27 October 2015) <www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf> accessed 17 May 2023.

22 Jerry H Ratcliffe and others, 'The Philadelphia Predictive Policing Experiment' (2021) 17 J Experimental Criminology 15.

23 David Raths, 'Growing Up Gov Tech' (January/February 2018) Government Technology Magazine 26, 28 <https://archives.erepublic.com/GT/GT_Mag_Jan_Feb2018.pdf> accessed 5 May 2023.

24 Joshua Brustein, The Ex-Cop at the Center of Controversy over Crime Prediction Tech (*Bloomberg*, 10 July 2017) <www.bloomberg.com/news/features/2017-07-10/the-ex-cop-at-the-center-of-controversy-over-crime-prediction-tech?leadSource=uverify%20wall> accessed 12 May 2023.

25 David Raths, 'Growing Up Gov Tech' (January/February 2018) Government Technology Magazine 26, 28 <https://archives.erepublic.com/GT/GT_Mag_Jan_Feb2018.pdf> accessed 5 May 2023.

26 David Raths, 'Growing Up Gov Tech' (January/February 2018) Government Technology Magazine 26, 28, 29 <https://archives.erepublic.com/GT/GT_Mag_Jan_Feb2018.pdf> accessed 5 May 2023.

records[27] to calculate risk scores'[28] It does not consider data concerning arrests for marijuana possession, given that research has shown these arrests exhibit blatant racial disparities; while CivicScape has invited discussion about the types of data used, police departments have prevented the company from sharing some of the data used to produce predictions.[29] In 2018, CivicScape was the only predictive policing firm that had published its code online, a move that 'earned praise from civil rights advocates concerned about the secrecy of algorithms that can send people to jail'.[30]

1.1.3 *Areas where selected AI-based systems are or have been used*

AI-based systems are or have been used by numerous municipal police departments in the United States. PredPol, for example, was used in 2018 by more than 60 police departments around the country, most of them mid-size agencies of 100 to 200 officers.[31] One study found that the areas targeted by PredPol were those most heavily populated by people of color and the poor:

> Analyzing entire jurisdictions, we observed that the proportion of Black and Latino residents was higher in the most-targeted block groups and lower in the least-targeted block groups (about 10% of which had zero predictions) compared to the overall jurisdiction. We also observed the opposite trend for the White population: The least-targeted block groups contained a higher proportion of White residents than the jurisdiction overall, and the most-targeted block groups contained a lower proportion.[32]

1.1.4 *Criminal activities at the focus of AI-based systems*

The theory and initial experiments underlying PredPol technology focused on a limited number of property-based crime (such as burglary and auto-related crime); the

---

[27] 311 is a non-emergency phone number that people can call for information about municipal services or to make complaints or report problems such as graffiti or road damage. Colin Wood, 'What Is 311?' (*Government Technology*, 2 August 2016) <www.govtech.com/dc/what-is-311.html> accessed 17 May 2023.

[28] Liz Coates, 'Case Builds for Greater Transparency in How Police Use Data to Predict Crime' (*Forbes*, 5 June 2017) <www.forbes.com/sites/delltechnologies/2017/06/05/case-builds-for-greater-transparency-in-how-police-use-data-to-predict-crime/> accessed 5 May 2023.

[29] Joshua Brustein, 'The Ex-Cop at the Center of Controversy over Crime Prediction Tech' (*Bloomberg*, 10 July 2017) <www.bloomberg.com/news/features/2017-07-10/the-ex-cop-at-the-center-of-controversy-over-crime-prediction-tech?leadSource=uverify%20wall> accessed 12 May 2023.

[30] Ben Miller, 'Team Ekistic Ventures' (*Government Technology*, 26 March 2018) <www.govtech.com/top-25/team-ekistic-ventures.html> accessed 17 May 2023.

[31] Randy Rieland, 'Artificial Intelligence Is Now Used to Predict Crime. But Is It Biased?' (*Smithsonian Magazine*, 5 March 2018) <www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/> accessed 5 May 2023.

[32] Dhruv Mehrotra and others 'How We Determined Predictive Policing Software Disproportionately Targeted Low-Income, Black, and Latino Neighborhoods' (*Gizmodo*, 2 December 2021) <https://gizmodo.com/how-we-determined-predictive-policing-software-dispropo-1848139456> accessed 4 May 2023.

prediction of violent crimes or individual criminals 'did not inform the early studies'.[33] In a subsequent step, PredPol adapted its focus to include gun violence and gang shootings.[34]

The focus of HunchLab as implemented in St. Louis County is more on serious felonies and less on crimes such as drug possession.[35] Andrew Guthrie Ferguson says that HunchLab focuses

> on direct patrol responses, it weights crimes that respond better to direct patrols more heavily. So, for example, a gun crime might have a higher severity weight (because of the risk to the community), and an aggravated assault charge might have a low police efficacy weight (because those impulsive, violent crimes are less deterred by police patrol). Police officers utilizing HunchLab are provided information about where to patrol (based on the forecast) and then also provided suggested tactics to improve efficiency in those particular areas.[36]

CivicScape focuses on identifying and forecasting property and violent crimes.[37]

1.1.5 Concrete results produced by AI-based systems

CivicScape[38] as well as PredPol and HunchLab are place-based predictive policing technologies: they help police to predict where and when crimes might occur.[39] With this kind of information, police administrators can restructure patrol routes and develop crime suppression strategies.[40]

The results of this kind of restructuring are not necessarily positive, however. In Santa Cruz, for example, Police Chief Andy Mills imposed a moratorium on the use of PredPol in 2017 because in his opinion the system had done more harm than good.

---

33 Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1132, 1148 (2017).

34 Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1134-35 (2017).

35 Jeffrey Benzing, 'The Promise and Dangers of "Predictive" Policing' (*PublicSource*, 15 February 2016) <www.publicsource.org/the-promise-and-dangers-of-predictive-policing> accessed 5 May 2023.

36 Andrew Guthrie Ferguson, 'Predictive Policing Theory' in Tamara Rice Lave and Eric J Miller (eds), *The Cambridge Handbook of Policing in the United States* (Cambridge Univ Press 2019) 497.

37 Liz Coates, 'Case Builds for Greater Transparency in How Police Use Data to Predict Crime' (*Forbes,* 5 June 2017) <www.forbes.com/sites/delltechnologies/2017/06/05/case-builds-for-greater-transparency-in-how-police-use-data-to-predict-crime/> accessed 5 May 2023.

38 Liz Coates, 'Case Builds for Greater Transparency in How Police Use Data to Predict Crime' (*Forbes,* 5 June 2017) <www.forbes.com/sites/delltechnologies/2017/06/05/case-builds-for-greater-transparency-in-how-police-use-data-to-predict-crime/> accessed 5 May 2023; David Raths, 'Growing Up Gov Tech' (January/February 2018) Government Technology Magazine 26, 28 <https://archives.erepublic.com/GT/GT_Mag_Jan_Feb2018.pdf> accessed 5 May 2023.

39 Andrew Guthrie Ferguson, 'Predictive Policing Theory' in Tamara Rice Lave and Eric J Miller (eds), *The Cambridge Handbook of Policing in the United States* (Cambridge Univ Press 2019) 491.

40 Andrew Guthrie Ferguson, 'Predictive Policing Theory' in Tamara Rice Lave and Eric J Miller (eds), *The Cambridge Handbook of Policing in the United States* (Cambridge Univ Press 2019) 491.

> All it did, he said, was inform the [Santa Cruz Police Department] where "to do purely enforcement," leading the department to over-police certain neighbourhoods around the city without building productive relationships with the community that might have helped tackle the root cause of criminal activity.[41]

In 2020, Santa Cruz became one of the first cities in the United States to ban the use of predictive policing.[42]

1.1.6 Public perceptions and media presentation of AI-based systems for predictive policing

Initial reporting about predictive policing was generally quite positive. Newspaper headlines touted the technology as the future of policing, with one headline in the Los Angeles Times from 2010 reading 'Stopping Crime before It Starts,'[43] and Time Magazine dubbing it one of the 50 best inventions of the year in 2011.[44] According to an article in the ABA Journal in 2013, adopting predictive policing software 'often results in a public relations boost for police departments'.[45]

In recent years, however, media coverage has become more critical, and in July 2019, the Los Angeles Times reported that PredPol, 'the widely hailed tool … developed by a UCLA professor in conjunction with the LAPD' had 'come under fire in the last 18 months' and that numerous departments across the country were terminating their use of the software.[46]

Some of the greatest proponents of predictive policing have been high profile police chiefs. In 2019, for example, Chief Michel Moore, LAPD, still supported the technology, saying that the LAPD needed 'location-based strategies to target crime and keep residents safe'.[47] And in the same year the New York Times reported that Charlie Beck, a former police chief of Los Angeles who had just been named interim police chief in

---

[41] Greg Noone, 'The Case Against Predictive Policing' (*Techmonitor,* 5 July 2021) <https://techmonitor.ai/technology/ai-and-automation/case-against-predictive-policing> accessed 4 May 2023.

[42] See text accompanying n 71.

[43] Joel Rubin, 'Stopping Crime before It Starts' *Los Angeles Times* (Los Angeles, 21 August 2010) A1. For additional references, see Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1109, 1131.

[44] Lev Grossman and others, 'The 50 Best Inventions of the Year' (*Time Magazine,* 28 November 2011) <http://content.time.com/time/magazine/article/0,9171,2099708,00.html> accessed 5 May 2023.

[45] Leslie A Gordon, Predictive Policing May Help Bag Burglars – But It May Also Be a Constitutional Problem (*ABA Journal*, 1 September 2013) <www.abajournal.com/magazine/article/predictive_policing_may_help_bag_burglars--but_it_may_also_be_a_constitutio> accessed 5 May 2023.

[46] Mark Puente, 'Crime Forecast Software Is Losing Its Luster' *Los Angeles Times* (Los Angeles, 8 July 2019) A1.

[47] Mark Puente and Cindy Chang, 'LAPD Will Adjust Data-Driven Predictor' *Los Angeles Times* (Los Angeles, 16 October 2019) B3.

Chicago, was known for his support of technology to fight crime, including predictive policing.[48] In 2018, Rod Rosenstein, deputy attorney general of the United States, declared the Trump administration's support for predictive policing.[49]

In contrast, law professors and law students, it seems, tend to view predictive policing critically: numerous articles critical of the technology have been published in law journals in the last ten years. As far as academia as a whole is concerned, in April 2019, 40 graduate students and 28 faculty members from the University of California Los Angeles (UCLA) signed an open letter to the Los Angeles Police Commission discrediting the PredPol program. The letter stated that there was no

> universal agreement or acceptance of the empirical merit and the ethics of [the] research at UCLA in anthropology as a discipline or in other disciplines. On the contrary, many anthropologists and other scholars … believe it represents some of the most troubling legacies of the discipline of anthropology and of social science more generally.[50]

In a subsequent action later that same year, over 400 academics – including some 140 faculty members from universities in the United States and abroad – signed a similar letter to the Police Commission.[51] Another letter, this one signed by some 1,400 mathematicians (holders of a PhD or doctoral students in the field), was sent in June 2020 to the trade journal Notices of the American Mathematical Society.[52] The signees called for their colleagues to stop collaborating with police because of the disparities in how law enforcement agencies treat people of various races and ethnicities: 'Given the structural racism and brutality in US policing, we do not believe that mathematicians should be collaborating with police departments in this matter. It is simply too easy to

---

48 Mitch Smith and Julie Bosman, 'Chicago Police Superintendent to Retire' (*New York Times*, 7 November 2019) <www.nytimes.com/2019/11/07/us/chicago-police-eddie-johnson-retires.html> accessed 5 May 2023.

49 Rod J Rosenstein, 'Remarks at the Project Safe Neighborhoods National Conference' (*US Dep't of Justice*, 5 December 2018) <www.justice.gov/opa/speech/deputy-attorney-general-rod-j-rosenstein-delivers-remarks-project-safe-neighborhoods> accessed 5 May 2023.

50 Mark Puente, 'Second LAPD Data Program Draws Scrutiny' Los Angeles Times (Los Angeles, 10 April 2019) A1. For text of letter and list of signees, see Stop LAPD Spying Coalition, 'On Tuesday April 2nd, 2019 Twenty-Eight Professors and Forty Graduate Students' (*Medium*, 4 April 2019) <https://stoplapdspying.medium.com/on-tuesday-april-2nd-2019-twenty-eight-professors-and-forty-graduate-students-of-university-of-8ed7da1a8655> accessed 8 May 2023.

51 Mark Puente and Cindy Chang, 'LAPD Will Adjust Data-Driven Predictor' *Los Angeles Times* (Los Angeles, 16 October 2019) B3. For text of letter and list of signees, see Stop LAPD Spying Coalition, 'Over 450 Academics Reject PredPol' (*Medium*, 9 October 2019) <https://stoplapdspying.medium.com/over-450-academics-reject-predpol-790e1d1b0d50> accessed 8 May 2023.

52 Davide Castelvecchi, 'Mathematicians Urge Colleagues to Boycott Police Work in Wake of Killings' (*Nature*, 19 June 2020) <www.nature.com/articles/d41586-020-01874-9> accessed 17 May 2023.

create a "scientific" veneer for racism.'[53] Criticism is aimed at predictive policing in general and PredPol (co-founded by a mathematician), in particular.

Numerous NGOs, including the American Civil Liberties Union, the Brennan Center for Justice, and the Electronic Frontier Foundation have also pointed out the dangers posed by predictive policing to racial justice and equity.[54] And New York University's AI Now Institute, which studies the social impacts of artificial intelligence, published a paper in May 2019 in which it concluded that predictive policing is potentially discriminatory if it relies on 'dirty data' from policing practices that disproportionately affect minority populations.[55]

1.1.7 Evaluation of the reliability of AI-based systems used for predictive policing

Despite the purported benefits of predictive-policing systems, evidence regarding their accuracy, reliability, and overall utility is, at best, mixed.[56]

Typically, defense teams and the general public do not have access to the source code that defines predictive policing software such as PredPol, HunchLab, and CivicScape. They do not have information on how the software was constructed nor do they have information concerning its reliability. Indeed, vendors often prohibit independent, third-party review of their systems, and they even deny requests to allow expert witnesses to review the details of the system under protective order.[57]

1.1.8 Evaluation of the impartiality of the AI-based systems

Studies that evaluate the impartiality of AI-based systems are rare. One study of predictive policing, conducted by interested parties and published in 2018, found 'no significant differences in the proportion of arrests by racial-ethnic group between control and treatment conditions'. This study concluded that 'predictive policing did not result in biased arrests'.[58]

---

[53] Letter to AMS Notices <https://docs.google.com/forms/d/e/1FAIpQLSfdmQGrgdCBCexTrpne7KXUzpbiI9LeEtd0Am-qRFimpwuv1A/viewform> accessed 17 May 2023.

[54] Am Civ Liberties Union, 'Statement of Concern about Predictive Policing by ACLU and 16 Civil Rights Privacy, Racial Justice, and Technology Organizations' (*American Civil Liberties Union,* 31 August 2016) <www.aclu.org/other/statement-concern-about-predictive-policing-aclu-and-16-civil-rights-privacy-racial-justice> accessed 5 May 2023.

[55] Rashida Richardson, Jason M Schultz, and Kate Crawford, 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice' (2019) 94 NYU L Rev Online 15.

[56] Am Law Inst, Principles of the Law: Policing, Tentative Draft No 3 (ALI 2021) § 2.06.

[57] Stephanie J Lacambra, Jeanna Matthews, and Kit Walsh, 'Opening the Black Box' (2018) 42 Champion 28, 29.

[58] P Jeffrey Brantingham, Matthew Valasik, and George O Mohler, 'Does Predictive Policing Lead to Biased Arrests? Results from a Randomized Controlled Trial' (2018) 5:1 Statistics and Public Policy 1, 6, DOI: 10.1080/2330443X.2018.1438940.

In contrast, an experiment conducted by the Human Rights Data Analysis Group,[59] using PredPol software fed with drug arrest history from Oakland, California, found the algorithm only reinforced police bias in reporting when making its predictions. Results of the experiment, which were published in 2016,[60] can be summarized as follows:

> Due to inequalities in which neighborhoods were being policed in the first place, there was an inequality in the data being produced and input into the training set. Locations that were heavily patrolled by police, like lower income communities of color, were over-represented in the police report data that was being used to train the algorithm. As a result, the algorithm learned not about patterns in actual crime, but about patterns in how police record crime, and used these patterns to predict and deploy patrols to the same areas overrepresented in the police data – reinforcing and hiding biased police practices by using a supposedly 'impartial' software program.[61]

Finally, a 2021 study that claimed to be the first independent effort to evaluate actual PredPol crime predictions analyzed predictions in 38 cities and countries in the United States. The study found that PredPol's algorithm 'disproportionately targeted vulnerable populations, including low-income communities and residents of public housing' and that its predictions 'disproportionately targeted neighborhoods with proportionately more Black and Latino residents'.[62]

1.1.9 Evaluation of the effectiveness of using AI-based systems for policing

There is very little research on whether predictive policing technology actually works. Most of the few existing studies are either written or paid for by the very companies that developed the technology at issue.[63]

Andrew Guthrie Ferguson pointed out in 2017 that the PredPol company had begun analyzing the performance of its product but that most other commercial products made no scientific claims as to the effectiveness of their technology. One reason for this absence of data and peer-reviewed publications, he explained, was that researchers 'require time and funding to conduct experiments, and policing urban areas with real criminals and

---

[59] The group defines itself as a 'non-profit, non-partisan organization that applies rigorous science to the analysis of human rights violations around the world'. Human Rights Data Analysis Group, 'About HRDAG' (*HRDAG*) <https://hrdag.org/knowledge-base/> accessed 23 May 2023.
[60] Kristian Lum and William Isaac, 'To Predict and Serve?' (2016) 13 Significance 14, 18.
[61] Stephanie J Lacambra, Jeanna Matthews, and Kit Walsh, 'Opening the Black Box' (2018) 42 Champion 28, 33.
[62] Dhruv Mehrotra and others 'How We Determined Predictive Policing Software Disproportionately Targeted Low-Income, Black, and Latino Neighborhoods' (*Gizmodo,* 2 December 2021) <https://gizmodo.com/how-we-determined-predictive-policing-software-dispropo-1848139456> accessed 4 May 2023.
[63] Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364, 378-79.

real victims provides an imperfect testing environment'.[64] He also referred to the difficulty of drawing causal conclusions, given the multitude of variables that play a role in why crime occurs or why crime rates drop across jurisdictions and over time.[65]

As of 2018, there seems to have been only one study that found a statistically significant decline in reported crime.[66] According to this study, which was conducted cooperatively by academics and police officials from the authorities (including the LAPD) using the system, crime volume decreased by an average 7.4% in the case of police patrols using PredPol forecasts, whereas there was no significant effect on crime volume associated with patrols based on analyst predictions.[67] Doubts concerning this result have been raised, however, as LAPD crime statistics show that crime in areas where PredPol was not in use decreased by as much as 16% during the same time period.[68]

Finally, in 2019, the Office of the Inspector General of the Los Angeles Police Commission was unable to determine whether PredPol, the LAPD's predictive-policing program, helped reduce crime.[69]

1.1.10 Public authorities that have decided not to use AI-based systems for predictive policing in the future

In the last several years, numerous public authorities have decided not to use AI-based systems for predictive policing. The following list is not exhaustive.[70]

The Santa Cruz Police Department, which began a pilot project on predictive policing in 2011, placed a moratorium on the practice in 2017 and banned it by city ordinance in 2020.[71] According to the ordinance, the propensity of predictive policing technology to

---

[64] Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1109, 1162.
[65] Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1109, 1162.
[66] William S Isaac, 'Hope, Hype, and Fear: The Promise and Potential Pitfalls of Artificial Intelligence in Criminal Justice' (2018) 15 Ohio St J Crim L 543, 556. See also Aziz Z Huq, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043.
[67] GO Mohler and others, 'Randomized Controlled Field Trials of Predictive Policing' (2015) 110 J Am Stat Ass'n 1399, 1400.
[68] William S Isaac, 'Hope, Hype, and Fear: The Promise and Potential Pitfalls of Artificial Intelligence in Criminal Justice' (2018) 15 Ohio St J Crim L 543, 556.
[69] Mark Puente and Cindy Chang, 'LAPD Will Adjust Data-Driven Predictor' *Los Angeles Times* (Los Angeles, 16 October 2019). For the publication of the Office of the Inspector General, an independent entity established through a voter-approved amendment to the City Charter, see Mark P Smith, Review of Selected Los Angeles Police Department Data-Driven Policing Strategies (*Office of the Inspector General*, 12 March 2019) <www.oig.lacity.org/_files/ugd/b2dd23_21f6fe20f1b84c179abf440d4c049219.pdf> accessed 23 May 2023.
[70] Richmond, California, for example, stopped using PredPol in 2016. Emily Thomas, 'Why Oakland Police Turned Down Predictive Policing' (*Vice*, 28 December 2016, <www.vice.com/en/article/ezp8zp/minority-retort-why-oakland-police-turned-down-predictive-policing> accessed 5 May 2023.
[71] Kristi Sturgill, 'Santa Cruz Becomes the First U.S. City to Ban Predictive Policing' (Los Angeles Times, 26 June 2020) <www.latimes.com/california/story/2020-06-26/santa-cruz-becomes-first-u-s-city-to-ban-

endanger civil rights and civil liberties outweighs its purported benefits, and the technology appears to have the propensity to exacerbate racial injustice.

In April 2020, the LAPD announced that it would stop using PredPol.[72] Public documents revealed at the time detailed how the technology 'reinforced decisions to patrol certain people and neighborhoods over others, leading to the over-policing of Black and brown communities'.[73] According to then police chief Michel R Moore, however, the decision to stop using the software was not because of concerns raised by activists but because of financial constraints due to COVID-19.[74]

In June 2020, the City of Pittsburgh suspended CrimeScan, its predictive policing program, due to concerns about racial bias.[75]

In December 2020, the New Orleans City Council passed an ordinance banning predictive policing.[76]

The Oakland City Council voted to ban the use of predictive policing technology in January 2021.[77] According to the city ordinance, predictive policing technology 'uses arrest data that can encode patterns of racist policing behavior and as a result, [is] more likely to predict a high potential for crime in minority neighborhoods or among minority people'. The ordinance also refers to studies that have shown that the technology perpetuates systemic racism and leads to disparate arrest rates.[78]

According to the National Association of Criminal Defense Lawyers, writing in September 2021, four cities had terminated their contracts with Geolitica (the successor company of PredPol) 'because "the minimal benefit did not justify continuing costs"' (Milpitas Police Department), 'because "it wasn't telling anything [the department] didn't know"' (Rio Rancho Police Department), 'because the "results were mixed"'

---

predictive-policing> accessed 24 May 2023. Santa Cruz, Cal, City Ord No 2020-17 § 9.85.030 (Ord No 2020-17, 11 August 2020).

72 Shivaun Field, 'Predictive Policing: What Is It and Should It Be Used in 2020?' (*UCIPT,* 20 April 2020) <https://predictiontechnology.ucla.edu/predictive-policing-what-is-it-and-should-it-be-used-in-2020/> accessed 5 May 2023.

73 Johana Bhuiyan, 'LAPD Ended Predictive Policing Programs Amid Public Outcry. A New Effort Shares Many of Their Flaws' (*The Guardian,* 8 November 2021) <www.theguardian.com/us-news/2021/nov/07/lapd-predictive-policing-surveillance-reform> accessed 5 May 2023.

74 Caroline Haskins, 'The Los Angeles Police Department Says It Is Dumping a Controversial Predictive Policing Tool' (*BuzzFeed News,* 22 April 2020) <www.buzzfeednews.com/article/carolinehaskins1/los-angeles-police-department-dumping-predpol-predictive> accessed 5 May 2023.

75 Ashley Murray and Kate Giammarise, 'City Halts Predictive Policing Program Algorithm' *Pittsburgh Post-Gazette* (Pittsburgh, 24 June 2020) A1. Pittsburg, Pa, City Ord § 116.15(c)(1) (Ord No 31-2020, § 1, eff 24 September 2020).

76 New Orleans, La, Code of Ordinances, § 147-2(b)(4) (Ord No 28559, 1 January 2021).

77 Nathan Sheard, 'Oakland's Progressive Fight to Protect Residents from Government Surveillance' (*EFF,* 20 January 2021) <www.eff.org/deeplinks/2021/01/oaklands-progressive-fight-protect-residents-government-surveillance> accessed 24 May 2023.

78 Oakland, Cal, Ord No 13635, 9 February 2021 (amending Oakland Municipal Code ch 9.64).

(Mountain View Police Department), or 'because [the department] "didn't find it effective" … and "it didn't help [the department] solve crime"' (Palo Alto Police Department).[79]

1.2 Normative Framework

1.2.1 National legal rules governing AI-based systems for predictive policing

There are no national legal rules that govern AI-based systems for predictive policing in the United States. Moreover, despite mounting concerns about the risks to society posed by AI tools in general and generative AI in particular,[80] as of May 2023, there seem to be no relevant bills pending in Congress. Commentators have, however, called for Congress (as well as state, county, and local legislatures) to pass legislation that would prevent predictive policing programs from being used, inadvertently, to increase arrests unnecessarily. They also call for legislation that would require police officers using such programs to be taught that persons encountered at 'hot spot' locations are not automatically subject to search and/or seizure.[81] Given the political climate, it is considered unlikely that Congress would pass legislation regulating the use of big data analytics such as predictive policing.[82]

1.2.2 Soft law sources addressing predictive policing

In 2022, the American Law Institute (ALI)[83] completed a project called *Principles of the Law, Policing*[84] whose goal is to provide guidance and suggest best practices to courts, legislatures, and police. Tentative Draft No. 3, which contains – among other things – a section devoted to police use of algorithms and profiles, was approved at the annual

[79] Wendy Lee, Jumana Musa, and Michael Pinard, 'Garbage In, Gospel Out: How Data-Driven Policing Technologies Entrench Historic Racism and "Tech-Wash" Bias in the Criminal Legal System' (*Nat'l Ass'n Crim Def Law,* September 2021) 72 <www.nacdl.org/datadrivenpolicing> accessed 10 May 2023.

[80] Cat Zakrzewski, Cristiano Lima, and Will Oremus, 'CEO behind ChatGPT Warns Congress AI Could Cause "Harm to the World"' (*Washington Post,* 16 May 2023) <www.washingtonpost.com/technology/2023/05/16/sam-altman-open-ai-congress-hearing/> accessed 19 May 2023.

[81] See, eg, I India Thusi and Robert L Carter, 'Transforming the System' (*The Opportunity Agenda,* 2016) <http://transformingthesystem.org/pdfs/Transforming-The-System-CJReport.pdf> accessed 5 May 2023. See also Richard Edwards, 'Criminal Justice Reform Topic: Affirmative' (*Nat'l Fed'n St High School Ass'ns*) <www.nfhs.org/media/4016198/criminal-justice-topic_affirmative-cases.pdf> accessed 10 May 2023 (calling for a ban on predictive policing techniques).

[82] Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364, 419 fn 315.

[83] The ALI, founded in 1923, consists of judges, lawyers, and academics. Its mission is 'to promote the clarification and simplification of the law and its better adaptation to social needs, to secure the better administration of justice, and to encourage and carry on scholarly and scientific legal work'. It served as a model for the European Law Institute, founded in 2011. Am Law Inst, *American Law Institute Annual Report 2021-2022* (ALI 2022) 3.

[84] News, 'Principles of the Law, Policing Is Approved' (*American Law Institute,* 18 May 2022) <www.ali.org/news/articles/principles-law-policing-approved> accessed 18 May 2023.

meeting of the ALI in 2021. According to the Draft, an agency 'should not rely on an algorithm or profile to direct police resources to a particular location, to identify potential targets for further investigation or surveillance, or to assess the risk of harm that individuals may pose to others' unless the following five criteria have been met:[85]

- the algorithm or profile is sufficiently accurate and transparent;

- officials using the algorithm or profile receive adequate training;

- the use of protected characteristics, such as race or ethnicity, satisfies established criteria;

- inputs that may reflect prior discriminatory enforcement practices are avoided; and

- the agency regularly examines the algorithm and underlying data to ensure accuracy and avoid bias.

1.2.3 Non-criminal case law regarding AI-based systems used for predictive policing

In the United States, only a few non-criminal cases have been decided that deal with AI-based systems used for predictive policing. One of them followed a Freedom of Information Law (FOIL)[86] request submitted in June 2016 by the Brennan Center for Justice at New York University School of Law[87] to the New York City Police Department (NYPD). The request sought disclosure of numerous records regarding the Department's use of predictive policing technology. In December 2016, after the NYPD failed to respond adequately to the FOIL request, the Center sued the Department, invoking the public's significant interest in the transparency of predictive policing systems. The NYPD responded that it had to respect nondisclosure agreements it had entered into and that disclosure of the test results of the predictive policing products of vendors that had bid unsuccessfully for the project would discourage potential vendors from contracting with the NYPD in the future, thereby limiting the pool of technology available to the Department. Additionally, the NYPD argued that information about tests of commercial products could reveal a vendor's trade secrets, while information about inputs and algorithms in use in the Department would potentially allow criminals to replicate the results and predict police movements.[88]

---

[85] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06.

[86] The Freedom of Information Law, art 6 (ss 84-90) of the New York Public Officers Law, provides the public rights of access to certain records maintained by government agencies.

[87] Brennan Center for Justice, 'About Us' (*Brennan Center*) <www.brennancenter.org/about> accessed 10 May 2023.

[88] For information and documents associated with the litigation, see Brennan Center for Justice, 'Court Case Tracker: Brennan Center for Justice v. New York Police Department (*Brennan Center,* 6 August 2021) <www.brennancenter.org/our-work/court-cases/brennan-center-justice-v-new-york-police-department> accessed 10 May 2023.

In a decision handed down in December 2017,[89] the trial court in New York held that nondisclosure agreements with vendors could not, without more, insulate the NYPD from a FOIL request. It directed the NYPD to disclose (redacted) email correspondence with the vendors, output data from its predictive policing system from its inception until six months before the date of the decision, and, finally (for in camera review), the summary of results of the 45-day trial of the three unsuccessful vendors' products. The request for disclosure of input data was denied.

The Brennan Center sees the result of the litigation as 'a victory for members of the public who deserve to know how their police department is allocating its resources, and for civil liberties advocates committed to understanding the implications of predictive policing for individual freedoms'.[90]

The response of legal commentators to this decision has been muted, but generally positive.[91]

1.2.4 *Criminal case law regarding AI-based systems used for predictive policing*

As of 2023, there have not been many criminal cases involving predictive policing decided in the United States. One case, however, decided by a federal appeals court in 2020, offers a multi-faceted discussion of the technology, with concurring and dissenting judges arguing among themselves.[92] The case involved a suspicionless search of the defendant, Billy Curry, Jr, conducted by the police in September 2017, which turned up a revolver on his person. The police officers, who were responding to reports of gunfire, were specifically assigned to monitor the area where the incident occurred as a result of predictive policing strategies that had been adopted by the Richmond Police Department.[93] In an 8-6 decision, the US Court of Appeals for the 4th Circuit, sitting en banc, agreed with the district court's decision to grant Curry's motion to suppress evidence of his revolver based on the unreasonable search that led to its discovery.

The concurring and dissenting opinions address the use of predictive policing technology and debate the various constitutional and public policy concerns such

---

[89] Brennan Ctr for Just at NY Univ Sch of Law v NYC Police Dep't, 2017 NY Misc LEXIS 5138 (NY Sup Ct, 22 December 2017).

[90] Brennan Center for Justice, 'Court Case Tracker: Brennan Center for Justice v. New York Police Department (*Brennan Center*, 6 August 2021) <www.brennancenter.org/our-work/court-cases/brennan-center-justice-v-new-york-police-department> accessed 10 May 2023.

[91] See, eg, Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364, 378; Céline Castets-Renard, 'Accountability of Algorithms in the GDPR and Beyond: A European Legal Framework on Automated Decision-Making' (2019) 30 Fordham Intell Prop Media & Ent LJ 91, 104 (lauding the decision as 'a first step toward more transparency').

[92] *United States v Curry*, 965 F3d 313 (4th Cir 2020)(en banc).

[93] For a description of the preventive technology used by the Richmond (Virginia) police, see Jennifer Bachner, *Predictive Policing: Preventing Crime with Data and Analytics* (IBM Center for the Business of Government 2013) 29-30.

technology raises. The concurring judges argued that the use of predictive policing technology could lessen the Fourth Amendment protections[94] of people who live in high crime areas and could contribute to the perpetuation of racial bias and profiling in the criminal justice system. The dissenting judges argued that predictive policing technology enhances community safety and ensures that less affluent, high-crime areas are not abandoned by law enforcement.[95]

The response of the legal community to this decision has been neutral to positive.

1.2.5 Normative instruments addressing the reliability, impartiality, and effective-ness of AI-based systems used for predictive policing

There are no hard national laws in the United States that expressly address the reliability, impartiality, and effectiveness of AI-based systems used for predictive policing. Tentative Draft No. 3 of the ALI Principles of the Law, Policing, encourages agencies that use predictive policing tools 'to take steps to verify that the algorithm … has been evaluated in some manner and has been shown to perform with acceptable accuracy'.[96] It also calls on agencies to 'insist that third-party developers disclose the inputs used to develop an algorithm and provide the agency with enough information to assess on an ongoing basis the efficacy of the predictive tool and any disparate impact it may have'. Furthermore, if an algorithm is used as part of the justification for a search or seizure or to inform a later determination in the course of the criminal process, the Draft states that 'the same information about the reliability of the algorithm, as well as the inputs used, should be made available to the defendant and to the court'.[97] The Draft also points out that evidence regarding the accuracy, reliability, and overall utility of predictive-policing systems is, at best, mixed and that the technology has the potential to exacerbate old problems and generate entirely new ones, particularly racial bias. It also questions both the accuracy and reliability of predictive policing algorithms.[98]

Regulation of AI-based systems for predictive policing and the companies that produce them

As of 2023, there is no obligation in the United States for AI-based systems to be certified or labelled before they can be used for predictive policing. In addition, there is no special

---

[94] The Fourth Amendment to the US Constitution establishes the right of the people to be secure in their persons against unreasonable searches and seizures. US Const amend IV. The protection against unreasonable seizures includes brief investigatory stops. See, eg, *Terry v Ohio*, 392 US 1 (1968).

[95] For a summary of the *Curry* decision, see Margo McGehee, 'Predictive Policing Technology: Fourth Amendment and Public Policy Concerns' (*U Cin L Rev Blog*, 17 February 2021) <https://uclawreview.org/2021/02/17/predictive-policing-technology-fourth-amendment-and-public-policy-concerns/> accessed 4 May 2023.

[96] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Comment c.

[97] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Comment c.

[98] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Reporters' Note b.

mechanism in place for holding companies that produce such systems accountable for the results the systems provide.

1.2.6 Rules governing the monitoring and adjustment of AI-based systems for predictive policing

There are no national laws in the United States that require authorities that use AI-based systems for predictive policing to monitor and adjust these systems at regular intervals. There are, however, a number of municipalities that have ordinances in place requiring periodic audits, such as, for example, Cambridge, Massachusetts. The Cambridge ordinance requires evaluation of the effectiveness of surveillance technologies, including predictive policing software, both before adoption and on an annual basis once such a technology is in use.[99]

Tentative Draft No. 3 of the ALI Principles of the Law, Policing, urges agencies to undertake routine audits of both the algorithm and the underlying data and to make all changes necessary to ensure accuracy and reliability:

> In particular, agencies should take steps to verify that any databases on which algorithms or profiles rely are up to date. And they should consider avoiding using inputs such as gang affiliations that have been shown to have a high error rate.[100]

After such an audit is conducted, the Tentative Draft encourages agencies to make the changes indicated by the audit results. According to the Draft, a number of agencies that had conducted audits of algorithm-based programs in recent years had found that the programs were less effective than expected.[101]

1.2.7 Transparency regarding the technological functioning of AI-based systems

Transparency with regard to the technological functions of AI-based systems used for predictive policing in the United States is not guaranteed. Indeed, companies that produce such systems may seek to avoid releasing relevant information by citing contractual nondisclosure agreements and by claiming the technology used is shielded by trade secret protection.[102] According to Tentative Draft No. 3 of the ALI Principles of the Law, Policing, some companies that sell predictive-policing software to public agencies consider their algorithms to be proprietary; the algorithm may be a black box 'even to the developers themselves'.[103] The developers may know what data the

[99] Cambridge, Mass, Municipal Code ch 2.128.060 (2018).

[100] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Comment c.

[101] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Reporters' Note c.

[102] Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364, 377. See discussion of the Brennan Center FOIL request, text accompanying nn 86ff.

[103] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Comment c.

algorithm has access to, but they may not know how the algorithm uses the data to calculate its predictions.[104] This lack of transparency may prevent agencies from understanding the systems they use.[105]

1.2.8 Transparency regarding use by organizations of AI-based systems for predictive policing

Organizations that use AI-based systems for predictive policing in the United States are not required to guarantee transparency about their practices. Indeed, even the very fact that such technology is in use is not always disclosed. While some municipal ordinances, expressly require city council approval before surveillance technology (including predictive policing software) can be acquired,[106] and while it could be expected that public budgeting discussions would, in any case, lead to disclosure,[107] this is not always the case. In New Orleans, for example, it was possible to keep the use of the technology secret because the company provided it free of charge, a strategy that enabled the Palantir initiative to be budgeted as a 'philanthropic venture'. Thus, there was no need for a public vetting of the program, and it 'flew under the radar'.[108] Furthermore, authorities may avoid releasing relevant information about predictive policing systems by citing contractual nondisclosure agreements and developer claims of trade secrets.[109]

1.2.9 Accountability of organizations that use AI-based systems

There are no special national rules for holding organizations that use AI-based systems for predictive policing accountable if the system malfunctions. Instead, generally applicable rules, such as Fourth Amendment guarantees regarding search and seizure, can be used to suppress evidence found in the course of a search conducted by officers in an area identified by predictive policing technology as a so-called 'hot spot'.[110]

Tentative Draft No. 3 of the ALI Principles of the Law, Policing, includes recommendations designed to enable agencies that use predictive policing tools to take responsibility for the tools they use and for the inputs and inferences on which these tools rely:

> A threshold requirement for using algorithms and profiles is that public officials – including both high-level agency officials who make the decision to acquire a

[104] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Comment b.
[105] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Reporters' Note b.
[106] Eg Cambridge, Mass, Municipal Code ch 2.128.030 (2018).
[107] Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364, 376.
[108] Nicole Lindsey, 'Predictive Policing Raises Important Privacy and Human Rights Concerns' (CPO Magazine, 16 March 2018) <www.cpomagazine.com/2018/03/16/predictive-policing-raises-important-privacy-and-human-rights-concerns> accessed 5 May 2023.
[109] Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364, 377.
[110] *United States v Curry*, 965 F3d 313 (4th Cir 2020)(en banc). See text accompanying nn 92ff.

particular tool and the agency officials who are authorized to use it – have at least a basic understanding of the inputs that the algorithm or profile uses and the limits of any inferences that may be drawn. This enables agency heads to take responsibility both for the tools they are using and for the various inputs and inferences on which they rely. This means that agencies generally should avoid the use of algorithms that are developed using machine-learning techniques that render the algorithms entirely opaque, even to developers themselves. This is particularly important if an algorithm is to be used as a basis for any sort of adverse action against a member of the public. Under these circumstances, it is essential that the government be able to explain the factors on which the algorithm relies and the reason why they are relevant to the determination at issue.[111]

1.2.10 Substantive obligations imposed on police authorities that use AI-based systems

There are no substantive obligations imposed nationally on the police authorities that use AI-based systems for predictive policing. Indeed, legal and policy debates associated with predictive policing have not kept up with the technological developments. For the most part, local police administrators who choose to enter into contracts with predictive policing companies are subject to very little public oversight.[112] Tentative Draft No. 3 of the ALI Principles of the Law, Policing, sets out a number of recommendations for police agencies.[113]

1.3 General principles of law

1.3.1 Protection of the right to equality – or the right to non-discrimination – with respect to AI-based systems used for predictive policing

Many legal scholars who have published in this area are critical of AI-based predictive policing systems. They argue that machine learning algorithms learn and reproduce the – often biased – data they are trained on[114] and that predictive policing is capable of perpetuating racial bias.[115] They also express skepticism about the capacity of even race-blind algorithms (i.e., those that are provided no race-specific data) to reduce racial biases in policing.[116]

---

[111] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Comment c.

[112] Andrew Guthrie Ferguson, 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1109, 1143.

[113] Am Law Inst, *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021) § 2.06, Comment c.

[114] Karl Manheim and Lyric Kaplan, 'Artificial Intelligence: Risks to Privacy and Democracy' (2019) 21 Yale JL & Tech 106, 109 ('[P]redictive policing and AI sentencing in criminal cases can reinforce discriminatory societal practices, but in a way that pretends to be objective.').

[115] Renata M O'Donnell, 'Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause' (2019) 94 NYU L Rev 544.

[116] Molly Griffard, 'A Bias-Free Predictive Policing Tool: An Evaluation of the NYPD's Patternizr' (2019) 47 Fordham Urb LJ 43, 52.

The discussion about the threats posed by the use of AI-based systems for predictive policing has even taken on constitutional overtones. In 2016, for example, a coalition of 17 civil rights, privacy, racial justice, and technology organizations issued a statement of concern that drew attention to the racial bias inherent in predictive policing and expressed misgivings regarding the constitutionality of the practice:

> Predictive policing systems threaten to undermine the constitutional rights of individuals … [and] must not be allowed to erode rights of due process and equal protection. Systems that manufacture unexplained 'threat' assessments have no valid place in constitutional policing.[117]

Passages from this document were cited with approval in a concurring opinion by Circuit Judge Thacker in the *Curry* case.[118] The right of equal protection referred to in the coalition's statement of concern is guaranteed by the Fourteenth Amendment to the United States Constitution.[119]

A 2019 law review publication claims to be the first piece to argue that machine learning-based predictive policing algorithms are 'a facial, race-based violation of the Equal Protection Clause'.[120] According to Renata M O'Donnell, the law-student author, machine learning confers upon an algorithm the capacity to 'learn, mimic, and refine patterns that exist in the real world' and, in the context of policing, 'allows an algorithm to associate race and criminality, and thereby discriminate via race-based facial classifications'. In her view, the Equal Protection Clause is 'the obvious remedy for facial discrimination'.[121] Basing her argument on equal protection jurisprudence, she goes on to say, however, that claimants will 'face significant barriers to success because of the difficulties of attributing private action to state actors and the difficulties of gathering proof of the algorithms' classifications on the basis of race'.[122]

A widely cited law journal article by Aziz Z Huq, also published in 2019, outlines the possible racial effects of algorithmic criminal justice, which the author defines as encompassing machine-learning (including deep-learning) tools likely to be deployed

---

[117] Am Civ Liberties Union, 'Statement of Concern about Predictive Policing by ACLU and 16 Civil Rights Privacy, Racial Justice, and Technology Organizations' (*American Civil Liberties Union*, 31 August 2016) <www.aclu.org/other/statement-concern-about-predictive-policing-aclu-and-16-civil-rights-privacy-racial-justice> accessed 5 May 2023.
[118] *United States v Curry*, 965 F3d 313, 345 (4th Cir 2020) (en banc). See text accompanying nn 92ff.
[119] 'No state shall … deny to any person within its jurisdiction the equal protection of the laws.' US Const amend XIV, § 1.
[120] Renata M O'Donnell, 'Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause' (2019) 94 NYU L Rev 544.
[121] Renata M O'Donnell, 'Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause' (2019) 94 NYU L Rev 544, 578.
[122] Renata M O'Donnell, 'Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause' (2019) 94 NYU L Rev 544, 578. See also, Note, 'Beyond Intent: Establishing Discriminatory Purpose in Algorithmic Risk Assessment' (2021) 134 Harv L Rev 1760.

for prediction purposes.[123] Huq argues that the Constitution's Equal Protection Clause, which 'purports to provide a general norm regulating the state's use of race', focuses on intent and classification and is thus ill-suited to 'the forms and dynamics of algorithmic criminal justice tools'.[124] Instead of relying on the tools provided by the Constitution, he proposes a different metric entirely for considering racial equity concerns in algorithmic design, namely, that criminal justice algorithms be evaluated in terms of their 'long-term, dynamic effects on racial stratification'.[125]

1.3.2 Protection of the right to privacy with regard to AI-based systems used for predictive policing

There is a multifaceted discussion about the conflict between privacy interests and the use of AI-based systems for predictive policing. Claims of privacy are raised both by the police departments that employ such systems and by the individuals affected by them. The police, for example, have been known to refuse to release information about their predictive programs, claiming concern for individual privacy.[126]

With the rights of individuals in mind, commentators have asked whether predictive policing instruments infringe on privacy rights protected by the Fourth Amendment to the United States Constitution.[127] The test for evaluating what 'reasonable' is under the Fourth Amendment was established in 1967 by *Katz v United States*.[128] Using this test, courts can assess whether law enforcement has violated an individual's constitutionally protected reasonable expectation of privacy. According to *Katz*, '[A] man's home is, for most purposes, a place where he expects privacy, but objects, activities, or statements that he exposes to the "plain view" of outsiders are not "protected," because no intention to keep them to himself has been exhibited.'[129] Consequently, a number of commentators have concluded that since individuals do not have a reasonable expectation of privacy outside their homes, the police can observe them without any court interference, and

---

[123] Aziz Z Huq, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043.

[124] Aziz Z Huq, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043, 1083.

[125] Aziz Z Huq, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043.

[126] Dave Collins, 'Should Police Use Computers to Predict Crimes and Criminals?' (*Phys.Org,* 5 July 2018) <https://phys.org/news/2018-07-police-crimes-criminals.html> accessed 5 May 2023. Dan Hunter, Mirko Bagaric, and Nigel Stobbs, 'A Framework for the Efficient and Ethical Use of Artificial Intelligence in the Criminal Justice System' (2020) 47 Fla St U L Rev 749, 765.

[127] The Fourth Amendment prohibits unreasonable searches and seizures. US Const amend IV ('The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.').

[128] 389 US 347 (1967).

[129] *Katz*, 389 US at 361.

predictive policing, which involves observation, does not require reasonable suspicion and as such does not activate Fourth Amendment protections.[130]

This situation has left a number of scholars unsatisfied. Already in 2012, Andrew Guthrie Ferguson, stated that predictive policing may necessitate a 'reconsideration of some of the existing reasonable suspicion doctrine'.[131] And more recently, Margaret Hu has suggested that technological developments may challenge existing views of which expectations of privacy are reasonable, thus calling into question the continued viability of the *Katz* test.[132]

In the meantime, several United States Supreme Court Justices have raised privacy concerns based on the aggregated nature of surveillance inherent in big data analytics. While predictive policing systems have not yet been addressed directly, in their concurring opinions to *United States v Jones*, Justices Alito and Sotomayor have pointed to the need for the Court to begin to take account of the ways in which big data technologies may change reasonable expectations of privacy under the Fourth Amendment.[133] Analysis of other recent Supreme Court decisions has led Sarah Valentine to express cautious hope that 'courts will be hesitant to accept overreaching reliance on algorithmically determined probable cause or reasonable suspicion'.[134]

1.3.3 *Protection of the right to liberty and security of persons against AI-based systems used for predictive policing*

Discussion about possible conflicts between the right to liberty, on the one hand, and the use of AI-based predictive policing systems, on the other, may take place within the framework of the rights guaranteed by the Fourth Amendment.[135] In a law review publication from 2015, for example, the law-student author Kelly K Koss discussed the role of predictive policing technology in the reasonable suspicion calculus undertaken to determine the permissibility of so-called Terry stops in alleged high-crime areas.[136] 'Terry stops' are a form of liberty-infringing, investigative intrusion governed by the reasonableness requirement of the Fourth Amendment.[137] A Terry stop is justified when

---

[130] See, eg, Erik Bakke, 'Predictive Policing: The Argument for Public Transparency' (2018) 74 NYU Ann Surv Am L 131, 143 (2018). See also Lindsey Barrett, 'Reasonably Suspicious Algorithms: Predictive Policing at the United States Border' (2017) 41 NYU Rev L & Soc Change 327, 329.

[131] Andrew Guthrie Ferguson, 'Predictive Policing and Reasonable Suspicion' (2012) 62 Emory LJ 259, 263.

[132] Margaret Hu, 'Cybersurveillance Intrusions and an Evolving Katz Privacy Test' (2018) 55 Am Crim L Rev 127, 129.

[133] *United States v Jones*, 565 US 400 (2012). See also Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364, 409.

[134] Sarah Valentine, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364, 410.

[135] See also text accompanying nn 127ff and nn 146ff.

[136] Kelly K Koss, 'Leveraging Predictive Policing Algorithms to Restore Fourth Amendment Protections in High-Crime Areas in a Post-Wardlow World' (2015) 90 Chi-Kent L Rev 301.

[137] *Terry v Ohio*, 392 US 1 (1968).

'an officer has reasonable suspicion – more than a "hunch" but less than probable cause – that criminal activity is occurring or will imminently occur'.[138] In such a situation, the officer may briefly stop, detain, and question the person in question if

> the suspicion is supported by articulable facts that are particularized to [that person]. Those facts and the rational inferences from them must reasonably justify the intrusion, and must be proportionate to the scope of the intrusion. The calculus of reasonable suspicion is determined in light of the totality of the circumstances.[139]

It would appear that the consensus of scholars writing in this area can be summed up by Andrew Guthrie Ferguson's 2015 statement: '[T]he growth of "big data" has the potential to change the reasonable suspicion calculus because more personal or predictive information about a suspect will make it easier for police to justify stopping a suspect.'[140]

1.3.4 Principle of proportionality and AI-based systems for predictive policing

Proportionality – disproportionality, really – is not infrequently mentioned in the context of AI-based systems for predictive policing. The topic often comes up in connection with feedback loops of crime prediction, that is, in connection with causing neighborhoods disproportionately targeted by law enforcement in the past to be overrepresented in the historical crime data used to train and build predictive crime algorithms.[141] And Kristian Lum and William Isaac, who have shown that predictive policing of drug crimes leads to increasingly disproportionate policing of historically over-policed communities, point to the discriminatory nature of the policy of imposing the disproportionate real costs of over-policing on affected communities.[142]

There is, however, very little discussion of the principle of proportionality in this context. Christopher Slobogin, for one, has argued that dragnet searches should be governed, among other things, by proportionality standards.[143] And he has been credited with the

---

[138] Lindsey Barrett, 'Reasonably Suspicious Algorithms' (2017) 41 NYU Rev L & Soc Change 327, 331 (citation omitted).

[139] Lindsey Barrett, 'Reasonably Suspicious Algorithms' (2017) 41 NYU Rev L & Soc Change 327, 331 (citation omitted).

[140] Andrew Guthrie Ferguson, 'Big Data and Predictive Reasonable Suspicion' (2015) 163 U Pa L Rev 327, 351. See text accompanying nn 146ff.

[141] Wendy Lee, Jumana Musa, and Michael Pinard, 'Garbage In, Gospel Out: How Data-Driven Policing Technologies Entrench Historic Racism and "Tech-Wash" Bias in the Criminal Legal System' (*Nat'l Ass'n Crim Def Law,* September 2021) 7-8 <www.nacdl.org/datadrivenpolicing> accessed 10 May 2023.

[142] Kristian Lum and William Isaac, 'To Predict and Serve?' (2016) 13 Significance 14, 19. See also Elizabeth E Joh, 'Feeding the Machine: Policing, Crime Data, and Algorithms' (2017) 26 Wm & Mary Bill Rts J 287, 301.

[143] Christopher Slobogin, 'Government Dragnets' (2010) 73 L & Contemp Probs. 107, 109.

observation that big data and predictive policing have contributed to a destabilization of the current framework that governs administrative searches.[144]

1.3.5 Procedural legality and AI-based predictive policing

There is a great deal of discussion about the effect of AI-based predictive policing on the requirement that law enforcement authorities base their investigations on suspicion. Some scholars argue that predictive policing systems pose a threat to rights protected by the Fourth Amendment to the US Constitution. The Fourth Amendment, which prohibits 'unreasonable searches and seizures'[145] is at the root of the doctrine requiring 'reasonable suspicion' or 'probable cause' for a police stop. Writing in 2012, Andrew Guthrie Ferguson was not yet sure how predictive policing would affect reasonable suspicion analysis, but of the fact that it would affect the analysis he expressed no doubt.[146] Three years later, he stated that one effect of 'big data' might be to provide the police with information that they could use to justify stopping a suspect.[147] In Ferguson's words, the shift from 'small data' to 'big data' 'simultaneously undermines the protection that reasonable suspicion provides against police stops and potentially transforms reasonable suspicion into a means of justifying those same stops'.[148] Numerous scholars argue that predictive analytics tools such as AI-based predictive policing may indeed make it easier for police to claim that an individual meets the reasonable suspicion standard, ultimately justifying more stops.[149]

---

[144] GS Hans, 'Curing Administrative Search Decay' (2018) 24 BU J Sci & Tech L 1, 20.

[145] See text accompanying nn 127ff and nn 136ff.

[146] Andrew Guthrie Ferguson, 'Predictive Policing and Reasonable Suspicion' (2012) 62 Emory LJ 259, 325.

[147] Andrew Guthrie Ferguson, 'Big Data and Predictive Reasonable Suspicion' (2015) 163 U Pa L Rev 327, 351.

[148] Andrew Guthrie Ferguson, 'Big Data and Predictive Reasonable Suspicion' (2015) 163 U Pa L Rev 327, 331 (2015).

[149] See, eg, Elizabeth E Joh, 'The Undue Influence of Surveillance Technology Companies on Policing' (2017) 92 NYU L Rev Online 19, 42 ('At some point in the near future, courts will have to determine whether an algorithm's determination can form the basis, at least in part, of Fourth Amendment suspicion. If informants and tips can help develop reasonable suspicion, it is likely that courts will accept big data analysis as another source of information for the police as well.'); Tim Lau, 'Predictive Policing Explained' (*Brennan Center,* 1 April 2020) <www.brennancenter.org/our-work/research-reports/predictive-policing-explained> accessed 5 May 2023; Wendy Lee, Jumana Musa, and Michael Pinard, 'Garbage In, Gospel Out: How Data-Driven Policing Technologies Entrench Historic Racism and "Tech-Wash" Bias in the Criminal Legal System' (*Nat'l Ass'n Crim Def Law,* September 2021) 10 <www.nacdl.org/datadrivenpolicing> accessed 10 May 2023 ('Does a person loitering on a corner in an identified "hotspot" translate to reasonable suspicion? What if that person was identified by an algorithm as a gang member or someone likely to be involved in drug dealing or gun violence? Can an algorithm alone ever satisfy the probable cause or reasonable suspicion requirement?')(citation omitted); Kelly Blount, 'Using Artificial Intelligence to Prevent Crime: Implications for Due Process and Criminal Justice' (*AI & Soc'y*, 24 June 2022) <https://doi.org/10.1007/s00146-022-01513-z> accessed 17 May 2023.

References

Am Civ Liberties Union, 'Statement of Concern about Predictive Policing by ACLU and 16 Civil Rights Privacy, Racial Justice, and Technology Organizations' (*American Civil Liberties Union*, 31 August 2016) <www.aclu.org/other/statement-concern-about-predictive-policing-aclu-and-16-civil-rights-privacy-racial-justice> accessed 5 May 2023

Am Law Inst, *American Law Institute Annual Report 2021-2022* (ALI 2022)

– – *Principles of the Law: Policing, Tentative Draft No. 3* (ALI 2021)

Aradau C and Blanke T, 'Decision' in C Aradau and T Blanke (eds), *Algorithmic Reason: The New Government of Self and Other* (OUP, 2022) 42-66

Azavea, *HunchLab: Under the Hood* (Azavea 2014) available at <www.brennancenter.org/sites/default/files/NYC_0002305_HunchlabPromoOverview%20-%20Copy.pdf> accessed 11 May 2023

Bachner J, *Predictive Policing: Preventing Crime with Data and Analytics* (IBM Center for the Business of Government, 2013)

Bakke E, 'Predictive Policing: The Argument for Public Transparency' (2018) 74 NYU Ann Surv Am L 131

Barrett L, 'Reasonably Suspicious Algorithms: Predictive Policing at the United States Border' (2017) 41 NYU Rev L & Soc Change 327

Benbouzid B, 'To Predict and to Manage. Predictive Policing in the United States' (2019) 6 Big Data & Soc'y 1

Benzing J, 'The Promise and Dangers of "Predictive" Policing' (*PublicSource*, 15 February 2016) <www.publicsource.org/the-promise-and-dangers-of-predictive-policing> accessed 5 May 2023

Bhuiyan J, 'LAPD Ended Predictive Policing Programs Amid Public Outcry. A New Effort Shares Many of Their Flaws' (*The Guardian*, 8 November 2021) <www.theguardian.com/us-news/2021/nov/07/lapd-predictive-policing-surveillance-reform> accessed 5 May 2023

Blount K, 'Using Artificial Intelligence to Prevent Crime: Implications for Due Process and Criminal Justice' (*AI & Soc'y*, 24 June 2022) <https://doi.org/10.1007/s00146-022-01513-z> accessed 17 May 2023

Brantingham PJ, Valasik M, and Mohler GO, 'Does Predictive Policing Lead to Biased Arrests? Results from a Randomized Controlled Trial' (2018) 5:1 Statistics and Public Policy 1-6

Brayne S, Rosenblat A, and boyd d, 'Predictive Policing' (Workshop primer presented at annual conference of Data & Civil Rights: A New Era of Policing and Justice, Washington, DC, 27 October 2015) <www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf> accessed 17 May 2023

Brennan Center for Justice, 'About Us' (*Brennan Center*) <www.brennancenter.org/about> accessed 10 May 2023

– – 'Court Case Tracker: Brennan Center for Justice v. New York Police Department' (*Brennan Center,* 6 August 2021) <www.brennancenter.org/our-work/court-cases/brennan-center-justice-v-new-york-police-department> accessed 10 May 2023

Brustein J, 'The Ex-Cop at the Center of Controversy over Crime Prediction Tech' (*Bloomberg,* 10 July 2017) <www.bloomberg.com/news/features/2017-07-10/the-ex-cop-at-the-center-of-controversy-over-crime-prediction-tech?leadSource=uverify%20wall> accessed 12 May 2023

Castelvecchi D, 'Mathematicians Urge Colleagues to Boycott Police Work in Wake of Killings' (*Nature,* 19 June 2020) <www.nature.com/articles/d41586-020-01874-9> accessed 17 May 2023

Castets-Renard C, 'Accountability of Algorithms in the GDPR and Beyond: A European Legal Framework on Automated Decision-Making' (2019) 30 Fordham Intell Prop Media & Ent LJ 91

Coates L, 'Case Builds for Greater Transparency in How Police Use Data to Predict Crime' (*Forbes,* 5 June 2017) <www.forbes.com/sites/delltechnologies/2017/06/05/case-builds-for-greater-transparency-in-how-police-use-data-to-predict-crime/> accessed 5 May 2023

Collins D, 'Should Police Use Computers to Predict Crimes and Criminals?' (*Phys.Org,* 5 July 2018) <https://phys.org/news/2018-07-police-crimes-criminals.html> accessed 5 May 2023

Deeks AS, 'Predicting Enemies' (2018) 104 Va L Rev 1529

Edwards R, 'Criminal Justice Reform Topic: Affirmative' (*Nat'l Fed'n St High School Ass'ns*) <www.nfhs.org/media/4016198/criminal-justice-topic_affirmative-cases.pdf> accessed 10 May 2023

Ferguson AG, 'Predictive Policing and Reasonable Suspicion' (2012) 62 Emory LJ 259

– – 'Big Data and Predictive Reasonable Suspicion' (2015) 163 U Pa L Rev 327

– – 'Policing Predictive Policing' (2017) 94 Wash U L Rev 1109

– – *The Rise of Big Data Policing* (New York University Press 2017)

– – ‘Predictive Policing Theory’ in TR Lave and EJ Miller (eds), *The Cambridge Handbook of Policing in the United States* (Cambridge Univ Press 2019)

Field S, ‘Predictive Policing: What Is It and Should It Be Used in 2020?’ (*UCIPT*, 20 April 2020) <https://predictiontechnology.ucla.edu/predictive-policing-what-is-it-and-should-it-be-used-in-2020/> accessed 5 May 2023

Gordon LA, ‘Predictive Policing May Help Bag Burglars – But It May Also Be a Constitutional Problem’ (*ABA Journal*, 1 September 2013) <www.abajournal.com/magazine/article/predictive_policing_may_help_bag_burglars--but_it_may_also_be_a_constitutio> accessed 5 May 2023

Griffard M, ‘A Bias-Free Predictive Policing Tool: An Evaluation of the NYPD’s Patternizr’ (2019) 47 Fordham Urb LJ 43

Grossman L and others, ‘The 50 Best Inventions of the Year’ (*Time Magazine*, 28 November 2011) <http://content.time.com/time/magazine/article/0,9171,2099708,00.html> accessed 5 May 2023

Hans GS, ‘Curing Administrative Search Decay’ (2018) 24 BU J Sci & Tech L 1

Haskins C, ‘The Los Angeles Police Department Says It Is Dumping a Controversial Predictive Policing Tool’ (BuzzFeed News, 22 April 2020) <www.buzzfeednews.com/article/carolinehaskins1/los-angeles-police-department-dumping-predpol-predictive> accessed 5 May 2023

Hu M, ‘Cybersurveillance Intrusions and an Evolving Katz Privacy Test’ (2018) 55 Am Crim L Rev 127

Human Rights Data Analysis Group, ‘About HRDAG’ (*HRDAG*) <https://hrdag.org/knowledge-base/> accessed 23 May 2023

Hunter D, Bagaric M, and Stobbs N, ‘A Framework for the Efficient and Ethical Use of Artificial Intelligence in the Criminal Justice System’ (2020) 47 Fla St U L Rev 749

Huq AZ, ‘Racial Equity in Algorithmic Criminal Justice’ (2019) 68 Duke LJ 1043, 1043 (2019)

Isaac WS, ‘Hope, Hype, and Fear: The Promise and Potential Pitfalls of Artificial Intelligence in Criminal Justice’ (2018) 15 Ohio St J Crim L 543

Joh EE, ‘Feeding the Machine: Policing, Crime Data, and Algorithms’ (2017) 26 Wm & Mary Bill Rts J 287

– – ‘The Undue Influence of Surveillance Technology Companies on Policing’ (2017) 92 NYU L Rev Online 19

Koss KK, 'Leveraging Predictive Policing Algorithms to Restore Fourth Amendment Protections in High-Crime Areas in a Post-Wardlow World' (2015) 90 Chi-Kent L Rev 301

Lacambra SJ, Matthews J, and Walsh K, Opening the Black Box, (2018) 42 Champion 28

Lau T, 'Predictive Policing Explained' (*Brennan Center,* 1 April 2020) <www.brennancenter.org/our-work/research-reports/predictive-policing-explained> accessed 5 May 2023

Lee W, Musa J, and Pinard M, 'Garbage In, Gospel Out: How Data-Driven Policing Technologies Entrench Historic Racism and "Tech-Wash" Bias in the Criminal Legal System' (*Nat'l Ass'n Crim Def Law,* September 2021) <www.nacdl.org/datadrivenpolicing> accessed 10 May 2023

Lindsey N, 'Predictive Policing Raises Important Privacy and Human Rights Concerns' (*CPO Magazine,* 16 March 2018) <www.cpomagazine.com/2018/03/16/predictive-policing-raises-important-privacy-and-human-rights-concerns> accessed 5 May 2023

Lum K and Isaac W, 'To Predict and Serve?' (2016) 13 Significance 14

Manheim K and Kaplan L, 'Artificial Intelligence: Risks to Privacy and Democracy' (2019) 21 Yale JL & Tech 106

McGehee M, 'Predictive Policing Technology: Fourth Amendment and Public Policy Concerns' (*U Cin L Rev Blog,* 17 February 2021) <https://uclawreview.org/2021/02/17/predictive-policing-technology-fourth-amendment-and-public-policy-concerns/> accessed 4 May 2023

Mehrotra D and others, 'How We Determined Predictive Policing Software Disproportionately Targeted Low-Income, Black, and Latino Neighborhoods' (*Gizmodo,* 2 December 2021) <https://gizmodo.com/how-we-determined-predictive-policing-software-dispropo-1848139456> accessed 4 May 2023

Miller B, 'Team Ekistic Ventures' (*Government Technology,* 26 March 2018) <www.govtech.com/top-25/team-ekistic-ventures.html> accessed 17 May 2023

Mohler GO and others, 'Randomized Controlled Field Trials of Predictive Policing' (2015) 110 J Am Stat Ass'n 1399

Murray A and Giammarise K, 'City Halts Predictive Policing Program Algorithm' *Pittsburgh Post-Gazette* (Pittsburgh, 24 June 2020) A1

News, 'Principles of the Law, Policing Is Approved' (*American Law Institute,* 18 May 2022) <www.ali.org/news/articles/principles-law-policing-approved> accessed 18 May 2023

Noone G, 'The Case against Predictive Policing' (*Techmonitor,* 5 July 2021) <https://techmonitor.ai/technology/ai-and-automation/case-against-predictive-policing> accessed 4 May 2023

Note, 'Beyond Intent: Establishing Discriminatory Purpose in Algorithmic Risk Assessment' (2021)134 Harv L Rev 1760

O'Donnell RM, 'Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause' (2019) 94 NYU L Rev 544

PredPol, 'How Predictive Policing Works' (*PredPol*) <www.predpol.com/how-predictive-policing-works/> accessed 5 May 2023

– – 'Overview' (*PredPol*) <www.predpol.com/about/> accessed 16 May 2023

Puente M, 'Second LAPD Data Program Draws Scrutiny' *Los Angeles Times* (Los Angeles, 10 April 2019) A1

– – 'Crime Forecast Software Is Losing Its Luster' *Los Angeles Times* (Los Angeles, 8 July 2019) A1.

– – and Chang C, 'LAPD Will Adjust Data-Driven Predictor' *Los Angeles Times* (Los Angeles, 16 October 2019) B3

Ratcliffe JH and others, 'The Philadelphia Predictive Policing Experiment' (2021) 17 J Experimental Criminology 15

Raths D, 'Growing Up Gov Tech' (January/February 2018) Government Technology Magazine 26 <https://archives.erepublic.com/GT/GT_Mag_Jan_Feb2018.pdf> accessed 5 May 2023

Richardson R, Schultz JM, and Crawford K, 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice' (2019) 94 NYU L Rev Online 15

Ricordeau G, 'Predictive Policing' in PL Frana and MJ Klein (eds), *Encyclopedia of Artificial Intelligence: The Past, Present, and Future of AI* (ABC-CLIO 2021)

Rieland R, 'Artificial Intelligence Is Now Used to Predict Crime. But Is It Biased?' (*Smithsonian Magazine,* 5 March 2018) <www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/> accessed 5 May 2023

Rosenstein RJ, 'Remarks at the Project Safe Neighborhoods National Conference' (*US Dep't of Justice,* 5 December 2018) <www.justice.gov/opa/speech/deputy-attorney-general-rod-j-rosenstein-delivers-remarks-project-safe-neighborhoods> accessed 5 May 2023

Rubin J, 'Stopping Crime before It Starts' *Los Angeles Times* (Los Angeles, 21 August 2010) A1.

Sheard N, 'Oakland's Progressive Fight to Protect Residents from Government Surveillance' (*EFF,* 20 January 2021) <www.eff.org/deeplinks/2021/01/oaklands-progressive-fight-protect-residents-government-surveillance> accessed 24 May 2023

Slobogin C, 'Government Dragnets' (2010) 73 L & Contemp Probs 107

Smith MP, Review of Selected Los Angeles Police Department Data-Driven Policing Strategies (*Office of the Inspector General,* 12 March 2019) <www.oig.lacity.org/_files/ugd/b2dd23_21f6fe20f1b84c179abf440d4c049219.pdf> accessed 23 May 2023

Smith M and Bosman J, 'Chicago Police Superintendent to Retire' (*New York Times,* 7 November 2019) <www.nytimes.com/2019/11/07/us/chicago-police-eddie-johnson-retires.html> accessed 5 May 2023

Stop LAPD Spying Coalition, 'On Tuesday April 2nd, 2019 Twenty-Eight Professors and Forty Graduate Students' (*Medium,* 4 April 2019) <https://stoplapdspying.medium.com/on-tuesday-april-2nd-2019-twenty-eight-professors-and-forty-graduate-students-of-university-of-8ed7da1a8655> accessed 8 May 2023

– – 'Over 450 Academics Reject PredPol' (*Medium,* 9 October 2019) <https://stoplapdspying.medium.com/over-450-academics-reject-predpol-790e1d1b0d50> accessed 8 May 2023

Sturgill K, 'Santa Cruz Becomes the First U.S. City to Ban Predictive Policing' (*Los Angeles Times,* 26 June 2020) <www.latimes.com/california/story/2020-06-26/santa-cruz-becomes-first-u-s-city-to-ban-predictive-policing> accessed 24 May 2023

Thomas E, 'Why Oakland Police Turned Down Predictive Policing' (*Vice,* 28 December 2016) <www.vice.com/en/article/ezp8zp/minority-retort-why-oakland-police-turned-down-predictive-policing> accessed 5 May 2023

Thusi II and Carter RL, 'Transforming the System' (The Opportunity Agenda, 2016) <http://transformingthesystem.org/pdfs/Transforming-The-System-CJReport.pdf> accessed 5 May 2023

Valentine S, 'Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control' (2019) 46 Fordham Urb LJ 364

Winston A and Burrington I, 'A Pioneer in Predictive Policing Is Starting a Troubling New Project' (*The Verge,* 26 April 2018) <www.theverge.com/2018/4/26/17285058/predictive-policing-predpol-pentagon-ai-racial-bias> accessed 16 May 2023

Wood C, 'What Is 311?' (*Government Technology,* 2 August 2016) <www.govtech.com/dc/what-is-311.html> accessed 17 May 2023

Zakrzewski C, Lima C, and Oremus W, 'CEO behind ChatGPT Warns Congress AI Could Cause "Harm to the World"' (*Washington Post,* 16 May 2023) <www.washingtonpost.com/technology/2023/05/16/sam-altman-open-ai-congress-hearing/> accessed 19 May 2023

2 Predictive Justice

2.1 National practices

2.1.1 Definition of 'predictive justice'

There is no single official legal definition of 'predictive justice' in the United States; nevertheless, the term has been in use for decades. One working definition, articulated in early 2022 in an editorial on the use of artificial intelligence (AI) in the administration of justice, viewed it as a process involving the use of machine learning algorithms 'that perform a probabilistic analysis of any given particular legal dispute using case law precedents'.[150] Another aspect of predictive justice, discussed in a piece published in 2018, involves machine learning systems that employ risk-assessment algorithms to estimate the likelihood of recidivism.[151]

Writing in 2008, a prolific law professor referred to lectures he delivered at the University of Cincinnati in 1973 as 'an occasion to lay out a general theory of predictive justice ... to articulate a theory of preventive actions based on predictive decisions'.[152] The focus of the professor's theory was preventive confinement.[153] And, already decades earlier, there was so much literature on the topic of preventive justice in 1958 that a detailed description and illustration of 'the predictive devices developed for sentencing to various types of imprisonment, for placement on probation, for release on parole, and for predicting the postparole conduct of former prisoners over a considerable span of time' would require 'too extensive a discussion' for a single article.[154]

2.1.2 Selected AI-based systems used for predictive justice

According to the Partnership Report on Artificial Intelligence published in 2019,[155] criminal justice risk assessment tools are basic forms of AI, even though they are usually

150 Raffaele Giarda, 'International: Artificial Intelligence in the Administration of Justice' (*LegalBytes*, January 2022) <https://bakerxchange.com/rv/ff008a110bc355ed8ab8ecd14f4f8822ba8d30ae/p=0> accessed 27 March 2023.

151 See, eg, Slava Polonski, 'Mitigating Algorithmic Bias in Predictive Justice: 4 Design Principles for AI Fairness' (24 November 2018) Towards Data Science <https://towardsdatascience.com/mitigating-algorithmic-bias-in-predictive-justice-ux-design-principles-for-ai-fairness-machine-learning-d2227ce28099> accessed 28 March 2023.

152 Alan M Dershowitz, 'Visibility, Accountability and Discourse as Essential to Democracy: The Underlying Theme of Alan Dershowitz's Writing and Teaching' (2008) 71 Alb L Rev 731.

153 Dershowitz (n 152) 745 ('although preventive confinement has always been and will always be practiced, no jurisprudence of preventive intervention has ever emerged. … . No philosopher, legal writer, or political theorist has ever, to this writer's knowledge, attempted to construct a systematic theory of when it is appropriate for the state to confine preventively.').

154 Sheldon Glueck, 'Predictive Devices and the Individualization of Justice' (1958) 23 Law & Contemp Probs 461, 471.

155 Partnership on AI is a 'non-profit partnership of academic, civil society, industry, and media organizations creating solutions so that AI advances positive outcomes for people and society'. Partnership on AI, 'About Us' <https://partnershiponai.org/about/> accessed 29 March 2023.

much simpler than the deep neural networks used in many modern AI systems. While some of them use heuristic frameworks to produce their scores, 'most use simple machine learning methods to train predictive models from input datasets.'[156] Arguably, there are a number of AI-based systems being used for predictive justice in the various jurisdictions of the United States. Three such systems will be introduced here: the commercially available Correctional Offender Management Profiling for Alternative Sanctions (COMPAS); the federal Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN); and the bespoke tool developed by the Pennsylvania Board of Probation and Parole.

COMPAS, a well-known commercially available instrument that seems to date to at least the late 1990s,[157] is widely used in the United States.[158] In Wisconsin, for example, it was implemented in 2012.[159] A proprietary algorithm sold by the private company currently known as equivant,[160] it is referred to as a fourth generation tool.[161] Fourth generation tools 'use machine learning in their modeling' and, in contrast to third generation tools, 'can output an explicit forecast, rather than a score'; when such a forecast is generated, 'it can be difficult to understand precisely what led to the system's determination'.[162] In

---

[156] Partnership on AI, 'Report on Algorithmic Risk Assessment Tools in the US Criminal Justice System' 7 (*Partnership on AI*, 23 April 2019) <https://partnershiponai.org/paper/report-on-machine-learning-in-risk-assessment-tools-in-the-u-s-criminal-justice-system> accessed 28 March 2023.

[157] JC Oleson, 'Risk in Sentencing: Constitutionally Suspect Variables and Evidence-Based Sentencing' (2011) 64 SMU L Rev 1329, 1349 fn125 (2011). See Mapping Pretrial Injustice, 'Common Pretrial Risk Assessments' <https://pretrialrisk.com/the-basics/common-prai> accessed 28 March 2023; Alexandra 'Mac' Taylor, 'AI Prediction Tools Claim to Alleviate and Overcrowded American Justice System … But Should They Be Used?' Stanford Politics (13 September 2020) <https://stanfordpolitics.org/2020/09/13/ai-prediction-tools-claim-to-alleviate-an-overcrowded-american-justice-system-but-should-they-be-used> accessed 28 March 2023; Tim Brennan, William Dieterich, and Beate Ehret, 'Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System' (2009) 36 Crim Just & Behavior 21.

[158] EPIC, 'AI in the Criminal Justice System' <https://epic.org/issues/ai/ai-in-the-criminal-justice-system> accessed 28 March 2023.

[159] Andrew Lee Park, 'Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing' (UCLA Law Review, 19 February 2019) <www.uclalawreview.org/injustice-ex-machina-predictive-algorithms-in-criminal-sentencing> accessed 28 March 2023.

[160] Three corporations, Northpointe, CourtView Justice Solutions, and Constellation Justice Systems, consolidated into a single branded entity called equivant in January 2017. Anne L. Washington, 'How to Argue with an Algorithm: Lessons from the COMPAS-ProPublica Debate' (2018) 17 Colo Tech LJ 131, 133 fn5.

[161] See, eg, Susan Turner and others, 'Development of the California Static Risk Assessment (CSRA): Recidivism Risk Prediction in the California Department of Corrections and Rehabilitation' 2 (September 2013) UC Irvine Center for Evidence-Based Correction Working Paper <https://bpb-us-e2.wpmucdn.com/sites.uci.edu/dist/0/1149/files/2013/12/Development-of-the-CSRA-Recidivism-Risk-Prediction-in-the-CDCR.pdf> accessed 28 March 2023.

[162] Michael E Donohue, 'A Replacement for Justitia's Scales: Machine Learning's Role in Sentencing' (2019) 32 Harv JL & Tech 657, 661. See also Orna Rabinovich-Einy and Ethan Katsh, 'Artificial Intelligence and the Future of Dispute Resolution – The Age of AI-DR' in Mohamed Abdel Wahab, Daniel Rainey, and Ethan Katsh (eds), *Online Dispute Resolution: Theory and Practice* (2nd edn, Eleven International Publishing 2021) pp. 471-488 (referring to COMPAS as an AI-based predictive algorithm).

the jurisdictions where COMPAS has been applied or adapted, judges may draw on the algorithm's output when making sentencing decisions.[163]

The Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN) was developed and implemented by the Federal Bureau of Prisons in accordance with legislation known as the First Step Act of 2018.[164] PATTERN, which was initially released in July 2019,[165] takes an AI-like approach.[166] It should be noted, however, that the question of whether machine learning was used to develop PATTERN – a question that was raised in Congressional testimony – was not immediately answered by the Department of Justice.[167] Staff of the Federal Bureau of Prisons use PATTERN to score inmates in their custody.[168]

In 2013, members of the Pennsylvania Board of Probation and Parole began using machine learning forecasts to help inform discrete parole release decisions. Funding to develop the Board's state-of-the-art risk assessment tools was provided by the National Institute of Justice.[169]

At this point it should be emphasized that the question of which tools in use in the various criminal justice systems of the United States in fact rely on machine-learning algorithms does not lead to uniform, straight-forward answers. According to an article published in 2022, 'a number of states now rely on algorithmic and Artificial Intelligence ("AI") systems to fine tune the assessment of future dangerousness.'[170] In contrast, the following was claimed in a 2021 article:

---

163 Ellora Thadaney Israni, 'When an Algorithm Helps Send You to Prison' The NY Times (New York, 26 October 2017) <https://www.nytimes.com/2017/10/26/opinion/algorithm-compas-sentencing-bias.html> accessed 27 March 2023.

164 Public Law No 115-391, 132 Stat 5195 (21 December 2018). See Michael Santos, 'PATTERN Risk and Needs Assessment Under First Step Act' (Prison Professors) <https://prisonprofessors.com/pattern-first-step-act> accessed 28 March 2023. See also National Institute of Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool* (No 303859, December 2021) <www.ojp.gov/pdffiles1/nij/303859.pdf> accessed 28 March 2023.

165 DOJ, 'Department of Justice Announces Enhancements to the Risk Assessment System and Updates on First Step Act Implementation' (15 January 2020 <www.justice.gov/opa/pr/department-justice-announces-enhancements-risk-assessment-system-and-updates-first-step-act> accessed 28 March 2023.

166 Harold J Krent and Robert Rucker, 'The First Step Act - Constitutionalizing Prison Release Policies' (2022) 74 Rutgers UL Rev 631, 643.

167 Amy B Cyphert, 'Reprogramming Recidivism: The First Step Act and Algorithmic Prediction of Risk' (2020) 51 Seton Hall L Rev 331, 360. '"The DOJ Report provides so few details on weighting, it is unclear what type(s) of models were used (such as regressions) and/or whether any type of machine learning (supervised or unsupervised) was employed."' Id. at 360 fn176.

168 National Institute of Justice (n 164).

169 Richard Berk, 'An Impact Assessment of Machine Learning Risk Forecasts on Parole Board Decisions and Recidivism' (2017) 13 J Experimental Criminology 193, 195. See also Aziz Z Huq, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043, 1076.

170 Krent and Rucker (n 166) 633 (footnotes omitted).

> Algorithmic tools have taken root in some court systems at least as aids to human decision-making in criminal cases with respect to questions of bail, sentencing, and parole. But so far, virtually none of these tools appear to rely on machine-learning algorithms. … As best we can determine, only one jurisdiction (Pennsylvania) has implemented any risk assessment tool in criminal justice that is based on machine learning. … . Despite somewhat frequent claims to the contrary in the popular media, all other algorithmic tools used by courts appear to be based on standard indices or conventional logistic regression models – not machine-learning algorithms.[171]

This article refers specifically to COMPAS as a non-learning algorithmic tool.[172]

2.1.3 Description and role in the decision-making process of AI-based systems in use in the United States

According to some authors, AI-based systems have been in use in the criminal justice systems of the United States since at least the early years of the 21st century. Of these, risk assessment tools – some of which may incorporate machine learning – are used in a variety of contexts, including pretrial risk assessment (pretrial detainment/bail), sentencing, parole and probation, and prison rehabilitation programs.[173] According to a 2019 law review article, AI, 'void of all human interaction, has been used to inform probation, sentencing, and parole decisions on the state level, and probation on the federal level'.[174]

In the context of sentencing, states tend to make the use of risk assessment tools advisory, rather than presumptive or mandatory.[175] In the Loomis decision of 2016, for example, the Wisconsin Supreme Court held that a sentencing court may consider a COMPAS risk assessment at sentencing but that COMPAS scores are but one of many factors that may be considered and weighed: risk scores alone may not be used to determine whether an offender is incarcerated or to determine the severity of an offender's sentence, and they may not be the determinative factor in deciding whether an offender can be supervised

---

171 Cary Coglianese and Lavi M Ben Dor, 'AI in Adjudication and Administration' (2021) 86 Brook L Rev 791, 801-803 (footnotes omitted).
172 Coglianese and Ben Dor (n 171) 803.
173 See Danielle Kehl, Priscilla Guo, and Samuel Kessler, 'Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing' (2017) <https://dash.harvard.edu/handle/1/33746041> accessed 28 March 2023. See also Mirko Bagaric and others, 'The Solution to the Pervasive Bias and Discrimination in the Criminal Justice System: Transparent and Fair Artificial Intelligence' (2022) 59 Am Crim L Rev 95, 130 (discussing impact of AI in the use of bail).
174 Rachel DiBenedetto, 'Reducing Recidivism or Misclassifying Offenders: How Implementing Risk and Needs Assessment in the Federal Prison System Will Perpetuate Racial Bias' (2019) 27 JL & Pol'y 414, 417 (footnotes omitted).
175 Brandon Garrett and John Monahan, 'Assessing Risk: The Use of Risk Assessment in Sentencing' (2019) 103 Judicature 42, 43.

safely and effectively in the community.[176] As a result of the limitations placed on the use of COMPAS, the discretion of the decision-maker continues to play an important role in the sentencing process; furthermore, there is very little information available about how judges actually use risk assessments in practice.[177]

In the federal prison system, in contrast, eligibility for early release is determined by PATTERN alone. No discretion on the part of prison authorities is involved:[178] 'Unlike COMPAS, PATTERN is not just one factor that is weighed in deciding who is eligible for benefits like early release, it is THE factor.'[179]

Outside the field of risk assessment tools, AI does not yet seem to have advanced to the point where it is relied upon to 'produce judicial decisions', but it has been used in other ways, such as predicting a Supreme Court ruling on a particular issue.[180] Also, by analyzing massive amounts of data, software developed in recent years can assist in the exercise of legal judgment, work that was traditionally thought to be immune to automation.[181] For instance, thanks to its 'machine-learning, artificial intelligence and natural language processing technologies', Ravel Law, acquired by LexisNexis in 2017,[182] 'provides strategic insight into an array of factors that affect a judge's decision-making'.[183]

In conclusion, in 2019, 'the more fantastic ideas such as using AI to objectively decide cases by analyzing facts and applying law' were still 'figments of creative imaginations'.[184] And as recently as 2021, authors who knew of 'no machine-learning tool that has been adopted in any court in the United States to make an ultimate, fully automated determination on a legal or factual question',[185] made the following statement:

---

176 *State v Loomis*, 881 NW2d 749, 769 (Wis 2016).

177 Garrett and Monahan (n 175), 43.

178 Krent and Rucker (n 166) 634.

179 Cyphert (n 167) 342 (emphasis in original).

180 See, eg, Matthew Hutson, 'Artificial Intelligence Prevails at Predicting Supreme Court Decisions' (Science, 2 May 2017) <www.sciencemag.org/news/2017/05/artificial-intelligence-prevails-predicting-supreme-court-decisions> accessed 28 March 2023. See Taylor B. Schaefer, 'The Ethical Implications of Artificial Intelligence in the Law' 55 Gonz L Rev 221, 225.

181 Dana Remus and Frank Levy, 'Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law' (2017) 30 Geo J Legal Ethics 501, 524 (2017).

182 LexisNexis, 'LexisNexis Announces Acquisition of Ravel Law' (8 June 2017) <www.lexisnexis.com/community/pressroom/b/news/posts/lexisnexis-announces-acquisition-of-ravel-law> accessed 28 March 2023.

183 PRWeb, 'Ravel Law Announces Unprecedented Judge Analytics Offering' (16 April 2015) <www.prweb.com/releases/2015/04/prweb12656883.htm> accessed 28 March 2023.

184 See also Richard C Kraus, 'Artificial Intelligence Invades Appellate Practice: The Here, The Near, and The Oh My Dear' (2019 Winter Edition) Appellate Issues <www.americanbar.org/groups/judicial/publications/appellate_issues/2019/winter/artificial-intelligence-invades-appellate-practice-the-here-the-near-and-the-oh-my-dear> accessed 28 March 2023.

185 Coglianese and Ben Dor (n 171) 798 (footnote omitted).

> Although it is still early in courts' assessment of judicial use of algorithmic tools, it seems noteworthy that, in all the cases decided to date that have actually wrestled with the issues, courts appear to have taken pains to emphasize that such tools only serve as one of multiple factors that a judge takes into account in reaching a decision.[186]

On the other hand, when John Roberts, Chief Justice of the US Supreme Court, was asked in April 2017 whether 'smart machines, driven with artificial intelligences, will assist with courtroom fact finding or, more controversially even, judicial decision making', he replied, 'It's a day that's here, and it's putting a significant strain on how the judiciary goes about doing things.'[187]

2.1.4 How AI-based systems used for predictive justice in the United States work

According to some scholars, COMPAS uses machine learning.[188] But because COMPAS is proprietary software, it is difficult to say much about how it functions. Indeed, 'there is almost no transparency about its inner workings'.[189] The COMPAS tool 'is organized around an algorithm that uses the answers to some 137 questions about a criminal suspect to rank them on a scale of 1 to 10 … with higher scores indicating a greater risk of recidivism'.[190] It considers variables from five main areas (criminal involvement, relationships and lifestyles, personality and attitudes, family, and social exclusion) and uses a combination of static and dynamic factors to assess the risk of recidivism. The algorithm is 'largely considered to be a black box: though its basic input information is available, the weighting of these inputs within the algorithm are proprietary, and thus not available to the public'.[191]

PATTERN takes an AI-like approach,[192] where AI is defined as 'the ability of a machine to perceive and respond to its environment independently and perform tasks that would typically require human intelligence and decision-making processes, but without direct human intervention'.[193] Although PATTERN does not utilize a fully autonomous AI or machine-learning algorithm, 'its algorithm nonetheless provides the foundation for

---

[186] Coglianese and Ben Dor (n 171) 811.

[187] Adam Liptak, 'Sent to Prison by a Software Program's Secret Algorithms' The NY Times (New York, 1 May 2017), A22.

[188] Donohue (n 162) 661. But Coglianese and Ben Dor (n 171) 803 (COMPAS is a 'non-learning algorithmic tool adopted by several state court systems for pretrial decisions'); Jeff Ward, 'Black Box Artificial Intelligence and the Rule of Law' 84 Law & Contemp Prob i, ii (2021)(referring to COMPAS as a simple, statistically-based algorithm).

[189] Kehl, Guo, and Kessler (n 173) 9, 11.

[190] Huq, 'Racial Equity' (n 169) 1047.

[191] Taylor (n 157). See also Kehl, Guo, and Kessler (n 173) 11.

[192] Krent and Rucker (n 166) 643.

[193] Christopher Rigano, 'Using Artificial Intelligence to Address Criminal Justice Needs' (January 2019) NIJ Journal 280 <https://www.nij.gov/journals/280/Pages/using-artificial-intelligence-to-address-criminal-justice-needs.aspx> accessed 28 March 2023

greater application as a more AI-like tool, including for example, automatic updating independent of human intervention.'[194] PATTERN (as updated following publication of the July 2019 Risk and Needs Assessment Report) incorporates fifteen factors: eleven dynamic and four static.[195]

As far as the bespoke Pennsylvania Board of Probation and Parole risk assessment tools are concerned, training data were provided by the state's Department of Corrections, and several machine learning procedures were applied. Random forests were determined to be the most effective.[196] The Pennsylvania tool was trained using data provided by the Department of Corrections. The data included information concerning the inmate's capacity for violence, sex offender status, conduct in prison, arrest and conviction history, gender, age, and intelligence as well as information from the inmate's Level of Service Inventory-Revised interview.[197]

2.1.5 Use of AI-based systems at various stages of the criminal process

Judicial authorities in numerous jurisdictions throughout the United States are required to use risk assessment tools at one or more stages of the criminal process. While the nature of the tools in question is not always clear, the Partnership Report states generally that 'criminal justice risk assessment tools are basic forms of AI', even if they are 'usually much simpler than the deep neural networks used in many modern artificial intelligence systems'.[198]

As far as the pretrial stage is concerned, information on the use throughout the country of risk assessment tools is provided by a website run since 2020[199] by the organizations 'Media Alliance Project' and 'MediaJustice'. According to the website ('Mapping Pretrial Injustice'), risk assessment tools are required by court order in Jefferson County,

---

[194] Krent and Rucker (n 166) 643 fn85.

[195] Dynamic factors: 1. Conviction(s) for any type of infraction during current incarceration period; 2. Conviction(s) for serious and violent infractions during current incarceration period; 3. Infraction-free (any type) during current incarceration period; 4. Infraction-free (serious and violent) during current incarceration period; 5. Number of programs completed (any); 6. Work programming; 7. Drug treatment while incarcerated; 8. Non-compliance with financial responsibility; 9. History of violence; 10. History of escapes; 11. Education score. Static factors: 1. Age at time of assessment; 2. Violent offense of conviction; 3. Sex offender status; 4. Criminal history score. See DOJ, 'The First Step Act of 2018: Risk and Needs Assessment System – UPDATE' 10–11 (footnotes omitted) (January 2020) <www.bop.gov/inmates/fsa/docs/the-first-step-act-of-2018-risk-and-needs-assessment-system-updated.pdf> accessed 28 March 2023.

[196] Berk (n 169) 195.

[197] See Berk (n 169) 195, 213-214.

[198] Partnership on AI, 'Report' (n 156) 7 ('Some [of the tools] use heuristic frameworks to produce their scores, though most use simple machine learning methods to train predictive models from input datasets.').

[199] Ethan Corey, 'New Data Suggests Risk Assessment Tools have Little Impact on Pretrial Incarceration' (*The Appeal*, 7 February 2020) <https://theappeal.org/new-data-suggests-risk-assessment-tools-have-little-impact-on-pretrial-incarceration> accessed 28 March 2023.

Alabama; they are required by legislation in Connecticut, Delaware, Hawaii, Kentucky, New Jersey, Rhode Island, Vermont, Virginia, and West Virginia; they are required by the state supreme court in Indiana and Nevada; and they are required by judicial council in Minnesota.[200]

Furthermore, the website reports that COMPAS is in use in at least 11 counties.[201] In addition to common national pretrial tools (such as COMPAS),[202] 11 states have developed their own tools and 49 counties across 22 states use locally-developed or otherwise county-specific tools.[203]

As far as sentencing is concerned, by 2017, 'numerous states', including Kentucky and Oklahoma, required sentencing judges to consider the results of 'evidence-based tools'.[204] In Kentucky, judges are required to consider the results of a defendant's risk and needs assessment;[205] in Oklahoma, the judge is required to review the defendant's risk and needs assessment if the defendant is a felony offender being considered for a community punishment pursuant to the Oklahoma Community Sentencing Act.[206] In other states, the use of risk assessment tools in the context of sentencing is advisory, rather than presumptive or mandatory.[207] In any case, given the lack of available information, it is difficult to determine what it means, in practice, for a sentencing courts to 'use' such a tool. Finally, the role of discretion in a decision-maker's determination should not be underestimated.[208]

2.1.6 Incentives for using AI-based systems

Rapid growth in the use of increasingly sophisticated risk assessment tools in criminal justice systems across the United States is due in part to reform efforts undertaken in order to reduce the country's extremely high incarceration rates.[209] Other potential

---

200 Mapping Pretrial Injustice, 'State Laws on Rats' <https://pretrialrisk.com/national-landscape/state-laws-on-rats> accessed 28 March 2023.

201 Mapping Pretrial Injustice, 'How Many Jurisdictions Use Each Tool?' <https://pretrialrisk.com/national-landscape/how-many-jurisdictions-use-each-tool> accessed 28 March 2023. Note that COMPAS has separate tools for use at sentencing and during the pretrial stage, Public Safety Risk Assessment Clearinghouse, 'Correctional Offender Management Profiling for Alternative Sanctions (COMPAS)' <https://bja.ojp.gov/sites/g/files/xyckuh186/files/media/document/compas.pdf> accessed 28 March 2023.

202 Mapping Pretrial Injustice, 'Common Pretrial Risk Assessments' <https://pretrialrisk.com/the-basics/common-prai> accessed 28 March 2023.

203 'How Many Jurisdictions' (n 201).

204 John Lightbourne, 'Damned Lies & Criminal Sentencing Using Evidence-Based Tools' (2017) 15 Duke L & Tech Rev 327, 332.

205 Ky Rev Stat § 532.007(3) (2022).

206 22 Okl St § 988.18 (2022). The Oklahoma Community Sentencing Act is codified at 22 Okl St §§ 988.1–988.24.

207 Garrett and Monahan (n 175) 43.

208 Garrett and Monahan (n 175) 43.

209 Partnership on AI, 'Report' (n 156) 7.

advantages of harnessing these tools include decreasing the disparities caused by the cash-bail system[210] and providing outcomes that are fairer and less punitive than those produced by human decision-makers when they act with complete discretion.[211] One of the aims of the federal First Step Act of 2018, which led to the development of PATTERN, was to lower federal prison numbers by providing for the early release of non-violent offenders.[212]

2.1.7 Alternative dispute resolution based on AI calculations

Online dispute resolution, which is in the process of taking the place of alternative dispute resolution, has 'gained significant traction in the United States'.[213] While AI has begun showing up in this context, mostly in the form of AI-based predictions, its use has been limited, at least where decision-making is involved.[214]

2.1.8 Public, media, and scholarly responses to AI-based systems for predictive justice

AI-based systems for predictive justice have received a great deal of negative press in recent years. The media, NGOs, and legal scholars tend to emphasize the negative aspects of the technology, particularly the risks of bias that accrue to the detriment of poorer communities and communities of color, groups already suffering from structural racism and human-emanating bias in the criminal justice context. The public perception reflects this. Fewer scholars, it seems, focus on the advantages that AI-based systems have to offer.

2.1.9 Reliability and impartiality of AI-based systems for predictive justice in use in the United States

A 2013 study of 19 criminal risk and need assessment tools in use in the United States found that validity, in most cases, had been examined in 'one or two studies' and that these investigations were frequently completed by the very people who had developed the instrument.[215] Another study, conducted between September 2019 and July 2020 by the Electronic Privacy Information Center (EPIC),[216] consisted of a survey of state usage

---

[210] Jessica Brand and Jessica Pishko, 'Bail Reform: Explained' (*The Appeal*, 14 June 2018) <https://theappeal.org/bail-reform-explained-4abb73dd2e8a> accessed 28 March 2023.

[211] Partnership on AI, 'Report' (n 156) 7.

[212] Bagaric and others (n 173) 123.

[213] Amy J Schmitz and Janet Martinez, 'ODR and Innovation in the United States' in Mohamed Abdel Wahab, Daniel Rainey, and Ethan Katsh (eds), *Online Dispute Resolution: Theory and Practice* (2nd edn, Eleven International Publishing 2021) pp. 611-638.

[214] Rabinovich-Einy and Katsh (n 162).

[215] Sarah L Desmarais and Jay P Singh, Risk Assessment Instruments Validated and Implemented in Correctional Settings in the United States 2 (*Council of State Governments Justice Center*, 27 March 2013) <https://csgjusticecenter.org/wp-content/uploads/2020/02/Risk-Assessment-Instruments-Validated-and-Implemented-in-Correctional-Settings-in-the-United-States.pdf> accessed 28 March 2023.

[216] EPIC is an NGO that was established in 1994 to 'protect privacy, freedom of expression, and democratic values in the information age'. EPIC, 'About Us' < https://epic.org/about> accessed 29 March 2023.

(pre-trial as well as other contexts) of risk assessment tools. A table summarizing the results of the survey indicates which of the numerous tools in use had been subject to a validation study. The table does not, however, indicate who carried out the study, nor does it identify tools that are AI-based.[217] As far as the performance of algorithmic risk tools with minorities is concerned, the few studies to date show 'some evidence that minorities are more likely to be ranked at higher risk, though this result is not consistent across studies and for all tools'.[218]

COMPAS has been evaluated by numerous entities, both independent and internal.[219] It has been the subject of research that questions accuracy, utility, and fairness.[220] In 2021, available validation studies of COMPAS were 'typically performed by employees, consultants, or research funding recipients' of the tool's owners.[221] In a summary published in 2010 of research findings from multiple studies,[222] the Northpointe Research and Development Department came to the overall conclusion that COMPAS was reliable and had both good predictive and construct validity.[223] The authors acknowledged that:

> much of the evidence for the reliability and validity of the COMPAS is found in the results of research studies conducted by Northpointe. We know that critics may discount this research. However, most of our in-house research is conducted for state agencies, and that [sic] competent research divisions within those agencies closely scrutinize our methods and results. Such state-sponsored studies are, thus, often subjected to a far more thorough vetting than that provided by the

217 EPIC, 'Liberty at Risk: Pre-Trial Risk Assessment Tools in the US' (September 2020) <https://epic.org/wp-content/uploads/2022/02/Liberty-At-Risk-Report-FALL-2020-UPDATE.pdf> accessed 28 March 2023.

218 Melissa Hamilton, 'Algorithmic Risk Assessment: A Progressive Policy in Pre-trial Release' (2021) 57 Idaho L Rev 615, 630 (citing Whitney Threadcraft-Walker and others, 'Gender, Race/Ethnicity and Prediction: Risk in Behavioral Assessment' (2018) 54 J Crim Just 12 (note that the studies do not indicate whether the tools were AI-based).

219 See, eg, Northpointe Research and Development Department, 'COMPAS Scales and Risk Models – Validity and Reliability: A Summary of Results from Internal and Independent Studies' (*Elec Priv Info Ctr*, 20 July 2010) <https://epic.org/algorithmic-transparency/crim-justice/EPIC-16-06-23-WI-FOIA-201600805-COMPASSummaryResults.pdf> accessed 28 March 2023.

220 Mirko Bagaric, Dan Hunter, and Nigel Stobbs, 'Erasing the Bias against Using Artificial Intelligence to Predict Future Criminality: Algorithms Are Color Blind and Never Tire' (2020) 88 U Cin L Rev 1037, 1044.

221 Melissa Hamilton, 'Algorithmic Risk Assessment: A Progressive Policy in Pretrial Release' (2021) 57 Idaho L Rev 615, 628 (2021).

222 Northpointe Research and Development Department, 'COMPAS Scales and Risk Models – Validity and Reliability: A Summary of Results from Internal and Independent Studies' (*Elec Priv Info Ctr*, 20 July 2010) <https://epic.org/algorithmic-transparency/crim-justice/EPIC-16-06-23-WI-FOIA-201600805-COMPASSummaryResults.pdf> accessed 28 March 2023.

223 'Construct validity' is relevant with regard to COMPAS' 'needs' scales; 'predictive validity' is relevant with regard to its 'risk scales'. Northpointe Research and Development Department, 'COMPAS Scales and Risk Models – Validity and Reliability: A Summary of Results from Internal and Independent Studies' 2-4, 7 (*Elec Priv Info Ctr*, 20 July 2010) <https://epic.org/algorithmic-transparency/crim-justice/EPIC-16-06-23-WI-FOIA-201600805-COMPASSummaryResults.pdf> accessed 28 March 2023.

editors of peer-reviewed journals often resulting from the fact that such agencies have direct access to the same data, can scrutinize such data and often can replicate and test our findings.[224]

By law, PATTERN is subject to annual review and validation by the Attorney General.[225] In August 2020, the National Institute of Justice contracted with two investigators to serve as consultants and to conduct the annual review and revalidation of PATTERN.[226]

The impact of machine learning forecasts used to help the Pennsylvania Board of Probation and Parole make parole release decisions was evaluated in a paper published in 2017.[227]

2.1.10 Findings of studies mentioned in question 2.1.9

COMPAS

One of the evaluations of COMPAS, carried out by ProPublica[228] and published in 2016,[229] garnered an enormous amount of attention in both the popular media and the scholarly literature.[230] According to the ProPublica study, the risk scores calculated by COMPAS

---

[224] Northpointe Research and Development Department, 'COMPAS Scales and Risk Models – Validity and Reliability: A Summary of Results from Internal and Independent Studies' 2 (*Elec Priv Info Ctr*, 20 July 2010) <https://epic.org/algorithmic-transparency/crim-justice/EPIC-16-06-23-WI-FOIA-201600805-COMPASSummaryResults.pdf> accessed 28 March 2023. A 2021 review of the literature on the two popular risk tools that were the focus of prior empirical studies, including COMPAS, revealed that 'the available validation studies were typically performed by employees, consultants, or research funding recipients of the tools' owners.' Melissa Hamilton, 'Evaluating Algorithmic Risk Assessment' (2021) 24 New Crim L Rev 156, 181.

[225] 18 USC § 3631(b)(4). See National Institute of Justice (n 164).

[226] National Institute of Justice (n 164).

[227] Berk (n 169).

[228] ProPublica describes itself as 'an independent, nonprofit newsroom that produces investigative journalism with moral force'. With a team of more than 100 journalists, ProPublica 'covers a range of topics including government and politics, business, criminal justice, the environment, education, health care, immigration, and technology'. See ProPublica, 'About Us' <www.propublica.org/about> accessed 28 March 2023.

[229] Julia Angwin and others, 'Machine Bias' (*ProPublica*, 23 May 2016) <www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 28 March 2023; Jeff Larson and others, 'How We Analyzed the COMPAS Recidivism Algorithm' (*ProPublica*, 23 May 2016) <www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm> accessed 28 March 2023.

[230] For scholarly literature, see, eg, Mirko Bagaric and Gabrielle Wolf, 'Sentencing by Computer: Enhancing Sentencing Transparency and Predictability and (Possibly) Bridging the Gap between Sentencing Knowledge and Practice' (2018) 25 Geo Mason L Rev 653; Huq, 'Racial Equity' (n 169); Sandra G Mayson, 'Dangerous Defendants' (2017) 127 Yale LJ 490. For popular media, see, eg, Julia Angwin, 'Make Algorithms Accountable' The NY Times (New York, 1 August 2016) A17; Sam Corbett-Davies and others, 'A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased against Blacks. It's Actually Not That Clear' The Washington Post (Washington, D.C., 17 October 2016)

were ‘remarkably unreliable in forecasting violent crime: Only 20 percent of the people predicted to commit violent crimes actually went on to do so’; furthermore, when a full range of crimes, including misdemeanors, was taken into account, ‘the algorithm was somewhat more accurate than a coin flip.’[231] Also, as far as false positives were concerned, ProPublica found that ‘the formula was particularly likely to falsely flag black defendants as future criminals, wrongly labeling them this way at almost twice the rate as white defendants’ and that, with regard to false negatives, ‘white defendants were mislabeled as low risk more often than black defendants.’[232]

The ProPublica study was not, however, without its detractors. For example, in September 2016, just a few months after the ProPublica study was published, an article strongly critical of its findings appeared in Federal Probation. Its authors stated:

> We think ProPublica’s report was based on faulty statistics and data analysis, and that the report failed to show that the COMPAS itself is racially biased, let alone that other risk instruments are biased. Not only do ProPublica’s results contradict several comprehensive existing studies concluding that actuarial risk can be predicted free of racial and/or gender bias, a correct analysis of the underlying data (which we provide below) sharply undermines ProPublica’s approach.[233]

And in October 2017, in response to the ProPublica-Northpointe contretemps, a group of computer science researchers wrote the following in the Washington Post:

> Algorithms have the potential to dramatically improve the efficiency and equity of consequential decisions, but their use also prompts complex ethical and scientific questions. … The problems we discuss apply equally to human decision-makers, and humans are additionally biased in ways that machines are not. We must continue to investigate and debate these issues as algorithms play an increasingly prominent role in the criminal justice system.[234]

PATTERN

As reported in the 2021 publication ‘Review and Revalidation of the First Step Act Risk Assessment Tool’ (Report), discrepancies with some of the measures used to create PATTERN version 1.2 were identified. The staff of the Bureau of Prison’s Office of Research and Evaluation together with the National Institute of Justice’s review and

<www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas> accessed 27 March 2023; Thadaney Israni (n 163).

231 Angwin and others (n 229).

232 Angwin and others (n 229).

233 Anthony W Flores, Kristin Bechtel, and Christopher T Lowenkamp, ‘False Positives, False Negatives, and False Analyses: A Rejoinder to Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And It’s Biased Against Blacks’ (2016) 80 Fed Probation 38.

234 Corbett-Davies and others (n 230).

revalidation expert consultants collaborated to correct the discrepancies.[235] Updated data were used to created PATTERN version 1.3. Among other things, the Report reviewed and analyzed the predictive validity and the racial and ethnic neutrality of PATTERN version 1.3: According to the Report, the results suggested that PATTERN 1.3 displayed a high level of predictive accuracy. And, with respect to racial and ethnic neutrality, the review of the risk and needs assessment system (as mandated by the First Step Act) must include 'an evaluation of the rates of recidivism among similarly classified prisoners to identify any unwarranted disparities, including disparities among similarly classified prisoners of different demographic groups, in such rates'.[236] The Report contains results of evaluations of PATTERN using a number of approaches that 'reflect the current scientific standards for assessing instrument neutrality'.[237] Racial and ethnic neutrality was examined in a number of different ways, including through differential prediction analyses, which assess a key question: 'Do racial and ethnic subgroups have different probabilities of recidivism controlling for PATTERN score?' According to the Report, PATTERN shows relatively high predictive accuracy across the five racial/ethnic (White, Black, Hispanic, Native American, Asian) groups.

> The predictive value … and differential prediction results …, however, are mixed and complex. The differential prediction analyses reveal statistically significant results in 28 of 48 tests (analyses of main effects). These include the overprediction of Black, Hispanic, and Asian males and females on some of the general recidivism tools and the underprediction of Black males and females and Native American males, relative to white individuals, on some of the violent recidivism tools. The magnitudes of differential prediction include:
>
> • 6 to 7 percent relative overprediction for Black females on the general recidivism tool
>
> • 12 to 15 percent relative underprediction of Native American males and females on the general recidivism tools
>
> • 5 to 8 percent relative overprediction of Asian males on the general and violent recidivism tools
>
> … . Statistically significant results do not necessarily invalidate a tool, particularly with large sample sizes. However, due to the importance of the FSA mandate to examine the risk and needs assessment system for racial and ethnic neutrality, these results will be a central focus of subsequent review and revalidation efforts.
>
> The NIJ consultants will also continue to investigate potential solutions for the differential prediction issues identified during this review, including testing

[235] National Institute of Justice (n 164).

[236] 18 USC § 3631(b)(4)(E).

[237] National Institute of Justice (n 164).

emerging debiasing techniques and engaging with stakeholders to explore the most promising and supportable approaches.[238]

According to an evaluation of PATTERN carried out by a legal scholar and published in 2020, the tool 'will have a disproportionate impact on Black inmates'[239] although it includes 'certain best practices in recidivism prediction, and its developers have made a good faith effort to engage advocates and scholars about the tool's development'.[240]

Pennsylvania Board of Probation and Parole's machine-learning protocol

The performance evaluation of the Pennsylvania Board of Probation and Parole's machine-learning protocol, published in 2017, showed that the machine learning forecasts 'apparently had no effect on the overall parole release rate but did appear to alter the mix of inmates released'.[241] The forecasts appeared to lead to reductions in rearrests for both nonviolent and violent crime.[242]

2.1.11 Neutrality compared: AI-based systems used for predictive justice versus humans

The question of whether AI-based systems for predictive justice provide more neutrality in the criminal justice system than humans has been and continues to be hotly debated. One side of the debate emphasizes the advantages of AI-based systems. This approach points out that the current process for making sentencing decisions, a process dominated by judges, has been shown to be heavily biased against disadvantaged groups, and it refers to research findings showing that under this process 'groups such as African Americans and unattractive people receive disproportionately heavier sentences than other people.' Scholars in this camp emphasize the fact that algorithms, unlike humans, 'have no subconscious thinking paths' and 'do exactly what they are programmed to do'.[243]

On the other side of the debate are exponents of Melvin Kranzberg's first law of technology.[244] The following quote, made in the context of AI risk assessments, stems from an AI sceptic: '[T]hese algorithms are neither good nor bad, but they are certainly

[238] National Institute of Justice (n 164).
[239] Cyphert (n 167) 331.
[240] Cyphert (n 167) 381.
[241] Berk (n 169).
[242] Huq, 'Racial Equity' (n 169) 1076. See Berk (n 169) 212-213.
[243] Mirko Bagaric, Dan Hunter, and Nigel Stobbs, 'Erasing the Bias against Using Artificial Intelligence to Predict Future Criminality: Algorithms Are Color Blind and Never Tire' (2020) 88 U Cin L Rev 1037, 1039, 1065, available at: https://scholarship.law.uc.edu/uclr/vol88/iss4/3.
[244] 'Technology is neither good nor bad; nor is it neutral.' Melvin Kranzberg, 'Technology and History: "Kranzberg's Laws"' (1986) 27 Technology and Culture 544, 545.

not neutral. To accept AI in our courts without a plan is to defer to machines in a way that should make any advocate of judicial or prosecutorial discretion uncomfortable.'[245]

2.1.12 Consistency compared: AI-based systems versus humans

Opinions regarding the consistency of AI-based systems compared to that of humans appear to be mixed and findings limited. In one law review article, published in 2021, the authors wrote in support of the consistency of machine-learning tools: 'If machine-learning tools are used as substitutes for – or even just as complements to – human decision-making, they could potentially reduce inconsistencies and other foibles that permeate human judgment.'[246] In contrast, the authors of a 2017 law review article were more critical of risk assessment tools and software, including those that incorporate machine learning. In their eyes, while such tools 'have the potential to improve sentencing accuracy in the criminal justice system and reduce the risk of human error and bias, they also have the potential to reinforce or exacerbate existing biases and to undermine certain basic tenets of fairness that are central to our justice system'.[247]

A qualitative study published in 2020 examined attorney attitudes – specifically, prosecutors and defense attorneys – towards risk assessment in sentencing and plea bargaining.[248] The findings of the study can be summarized as follows:

> Prosecutors, for example, favored the use of risk assessment tools for sentencing, arguing they "were likely a more consistent and fair way than relying on intuition or personal experience." On the other hand, defense attorneys "were consistently opposed to using future recidivism risk as a factor in sentencing," as tools measuring future recidivism were based on "group means" rather than individual ones.[249]

2.1.13 Effect of AI-based systems on responses to crime

There is not a great deal of information available to show whether AI-based systems lead to harsher or more lenient responses to crimes or other violations of the law. In a 2017 assessment of the machine learning risk forecasts used by the Pennsylvania Board of Probation, it was found that the forecasts did not seem to have an effect on the overall parole release rate but did seem to alter the mix of inmates released.[250] The conclusion

---

[245] Jason Tashea, 'Courts Are Using AI to Sentence Criminals. That Must Stop Now' 17 April 2017, Wired <www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now> accessed 28 March 2023.
[246] Coglianese and Ben Dor (n 171) 828 (footnote omitted).
[247] Kehl, Guo, and Kessler (n 173) 36.
[248] Anne Metz and others, 'Valid or Voodoo: A Qualitative Study of Attorney Attitudes Towards Risk Assessment in Sentencing and Plea Bargaining' 8 (10 March 2020) Virginia Public Law and Legal Theory Research Paper No 2020-25, Duke Law School Public Law & Legal Theory Series No 2020-15 <https://papers.ssrn.com/sol3/papers.cfm?abstractid=3552018> accessed 28 March 2023.
[249] Bagaric and others (n 173), 122 (footnotes omitted).
[250] Berk (n 169) 193.

was that 'risk assessments based on machine learning forecasts can improve parole release decisions, especially when distinctions are made between re-arrests for violent and nonviolent crime.'[251]

2.1.14 Future of AI-based systems for predictive justice purposes

Despite controversy surrounding the use of (AI-based) systems for predictive justice, numerous jurisdictions in the United States continue to use COMPAS, and PATTERN remains in use on the federal level. Examples of public authorities that have terminated their use of AI-based systems for predictive justice purposes are not readily apparent.

2.2 Normative framework

2.2.1 National legal rules governing the use of AI-based systems for predictive justice

As of 2023, there were no national legal rules specifically governing the use of AI-based systems for predictive justice in the United States.

> Given the federalist structure of the United States, the development and implementation of AI technology in the public sector … is not determined by any central institution. … Decisions about digital technologies used by courts throughout the United States are … made by a plethora of institutions and actors. … Any one of these numerous … entities could in principle have its own policy with respect to … the use of algorithms to support decision-making.[252]

In April 2021, however, a bill, the 'Justice in Forensic Algorithms Act of 2021,'[253] was introduced in the US House of Representatives. Had it been enacted into law before the end of the 117th Congress (2021-2022), the Act would have established a federal framework to govern the use of computational forensic software. The bill defined computational forensic software as 'software that relies on an automated or semiautomated computational process, including one derived from machine learning, statistics, or other data processing or artificial intelligence techniques, to process, analyze, or interpret evidence.' The framework would have contained various elements, including the following:

> – requirements for the establishment of testing standards and a testing program for computational forensic software,
>
> – requirements for the use of computational forensic software by federal law enforcement agencies and related entities (e.g., crime labs),

[251] 'An Impact Assessment of Machine Learning Risk Forecasts on Parole Board Decisions and Recidivism' (2017) 13 J Experimental Criminology 193, 193.
[252] Coglianese and Ben Dor (n 171) 793.
[253] HR 2438 (117th Congress, 2021-2022).

– a ban on the use of trade secret evidentiary privilege to prevent federal criminal defendants from accessing evidence collected using computational forensic software or information about the software (e.g., source code), and

– limits on the admissibility of evidence collected using computational forensic software.[254]

Arguments for adoption of this federal legislation included the advantages of interpretable, not black box, technologies: if interpretable information is accessible to judges, prosecution, and defense counsel, they can understand the results produced by the technologies and can, in turn, explain them to jurors and other stakeholders in the criminal justice system.[255] Most importantly, defendants and defense counsel who are able to understand how forensic technologies reach their conclusions could have contested them in a meaningful way if such findings were used as evidence against them.[256] As of March 2023, the bill had not been reintroduced in the 118th Congress (2023-2024).

Legislative activity in this area also takes place at the state and local levels. Idaho, for example, enacted legislation in 2019 that specifically addresses the transparency, accountability, and explainability of pretrial risk assessment tools. The law requires all information used to build or validate such tools to be open to public inspection; entitles parties to criminal cases in which the court has considered or an expert witness has relied upon such a tool to review all calculations and data used to calculate the defendant's risk score; and prohibits builders and users of pretrial risk assessment tools from asserting trade secret or other intellectual property protections to quash discovery of relevant information in criminal and civil cases.[257]

---

[254] Quoted from the Bill Summary authored by CRS (Congressional Research Service). Both the Bill Summary and the text of the bill are available at <www.congress.gov/bill/117th-congress/house-bill/2438?r=3&s=6> accessed 31 March 2023.

[255] See Duke Government Relations, 'The Need for Transparency and Interpretability at the Intersection of AI and Criminal Justice' (22 November 2021) <https://governmentrelations.duke.edu/2021/11/22/the-need-for-transparency-and-interpretability-at-the-intersection-of-ai-and-criminal-justice> accessed 28 March 2023.

[256] Julie Pattison-Gordon, 'Justice-Focused Algorithms Need to Show Their Work, Experts Say' (12 May 2022) government technology <www.govtech.com/computing/justice-focused-algorithms-need-to-show-their-work-experts-say> accessed 28 March 2023. See also Brookings Institution, 'Forensic Algorithms: The Future of Technology in the US Legal System' (*Brookings*, 12 May 2022) <www.brookings.edu/events/forensic-algorithms-the-future-of-technology-in-the-us-legal-system> accessed 28 March 2023; Doug Austin, 'The Justice in Forensic Algorithms Act: Legal Technology Trends' (*eDiscovery Today*, 26 May 2022), https://ediscoverytoday.com/2022/05/26/the-justice-in-forensic-algorithms-act-legal-technology-trends> accessed 28 March 2023.

[257] Idaho Code § 19-1910 (2022). For additional examples of proposed and enacted legislation at the state and local levels (validation study requirements, transparency in how prosecutors use risk assessments, AI task forces and commissions, etc.), see 'Liberty at Risk' (n 217).

2.2.2 *Normative instruments produced by the executive authorities of your country deal with AI-based systems for predictive justice*

In February 2019, President Donald Trump promulgated Executive Order 13859, entitled 'Maintaining American Leadership in Artificial Intelligence.' The order[258] required the Office of Management and Budget to issue a memorandum to agencies urging them to 'consider ways to reduce barriers to the use of AI technologies in order to promote their innovative application while protecting civil liberties, privacy, American values, and United States economic and national security'.[259] In November 2019, in response to the order, the memorandum 'Guidance for Regulation of Artificial Intelligence Applications' was issued for the heads of executive departments and agencies.[260] While not specifically focused on AI-based systems for predictive justice, the memorandum encouraged agencies to coordinate with each other 'to ensure consistency and predictability of AI-related policies that advance American innovation and adoption of AI' and reminded them of the need appropriately to protect 'privacy, civil liberties, national security, and American values" and to allow 'sector- and application-specific approaches'.[261]

2.2.3 *Soft law sources concerning predictive justice*

In October 2022, the White House Office of Science and Technology Policy published a document entitled 'Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People'. The document, a white paper intended to 'support the development of policies and practices that protect civil rights and promote democratic values in the building, deployment, and governance of automated systems', is non-binding and does not constitute US government policy.[262] As stated in the title, it is a blueprint rather than an actual AI bill of rights. Its five principles state that people should be protected from automated systems that are unsafe or ineffective; they should be protected from algorithmic discrimination; they should enjoy data privacy; they should be notified when an automated system is being used, and explanations of outcomes should be provided; and they should, where appropriate, be able to opt out from automated systems in favor of a human alternative. While not specifically targeting

[258] An executive order is a declaration by the president that has the force of law. Executive orders do not require any action by Congress to take effect, and they cannot be overturned by Congress. See Legal Information Institute, 'Executive Order' (*Legal Information Institute*) <www.law.cornell.edu/wex/executive_order> accessed 28 March 2023.

[259] Exec Order No 13859 of 11 February 2019, Maintaining American Leadership in Artificial Intelligence, 84 Fed Reg 3967 (14 February 2019).

[260] Russell T Vought, Memorandum for the Heads of Executive Departments and Agencies, *Guidance for Regulation of Artificial Intelligence Applications* (17 November 2020) <https://trumpwhitehouse.archives.gov/wp-content/uploads/2020/11/M-21-06.pdf> accessed 28 March 2023.

[261] Vought (n 260).

[262] The White House, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People, 2 (October 2022) <www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf> accessed 29 March 2023.

predictive justice, the blueprint calls for 'enhanced protections and restrictions for data and interferences related to sensitive domains', including criminal justice; furthermore, automated systems intended for use within sensitive domains such as criminal justice should be 'tailored to the purpose, provide meaningful access for oversight, include training for any people interacting with the system, and incorporate human consideration for adverse or high-risk decisions'.[263]

Another soft-law source, the Model Penal Code,[264] prominently endorsed the consideration of risk in the sentencing process in its 2017 revision (MPC-S).[265] Once risks and needs processes developed by the sentencing commission[266] – including, presumably, those based on AI – prove to be sufficiently reliable, they may be incorporated into the sentencing guidelines:

> MPC-S § 6B.09. Evidence-Based Sentencing; Offender Treatment Needs and Risk of Reoffending.
>
> (1) The sentencing commission shall develop instruments or processes to assess the needs of offenders for rehabilitative treatment, and to assist the courts in judging the amenability of individual offenders to specific rehabilitative programs. When these instruments or processes prove sufficiently reliable, the commission may incorporate them into the sentencing guidelines.
>
> (2) The commission shall develop actuarial instruments or processes, supported by current and ongoing recidivism research, that will estimate the relative risks that individual offenders pose to public safety through their future criminal conduct. When these instruments or processes prove sufficiently reliable, the commission may incorporate them into the sentencing guidelines.
>
> (3) The commission shall develop actuarial instruments or processes to identify offenders who present an unusually low risk to public safety, but who are subject to a 32 presumptive or mandatory sentence of imprisonment under the laws or guidelines of the state. When accurate identifications of this kind are reasonably feasible, for cases in which the offender is projected to be an unusually low-risk

[263] White House (n 262) 6-7.

[264] The Model Penal Code, first promulgated in 1962, is a model code assembled by the American Legal Institute. Am Law Inst, '*Model Penal Code: Official Draft and Explanatory Notes*' (The American Law Institute, 1985). The part on sentencing has been revised in light of the changes in sentencing philosophy and practice that have taken place in the last half century. The project was approved in 2017; publication of the official text is expected in 2023. Am Law Inst, 'Model Penal Code: Sentencing' <www.ali.org/projects/show/sentencing> accessed 29 March 2023.

[265] Am Law Inst, Model Penal Code: Sentencing, Proposed Final Draft § 6B.09(3) (Proposed Final Draft, 2017) <https://robinainstitute.umn.edu/publications/model-penal-code-sentencing-proposed-final-draft-approved-may-2017> accessed 28 March 2023.

[266] The MPC-S recommends that all American jurisdictions establish a permanent sentencing commission as an essential agency of the criminal justice system. See MPC-S § 6A.

offender, the sentencing court shall have discretion to impose a community sanction rather than a prison term, or a shorter prison term than indicated in statute or guidelines. The sentencing guidelines shall provide that such decisions are not departures from the sentencing guidelines.

2.2.4 *Case law that addresses AI-based systems used for predictive justice: Criminal courts*

The courts have only just begun to grapple with the legal implications of the use of algorithmic risk assessment tools in sentencing. Several such cases have been litigated in recent years in the United States. While the focus has not been on the use of AI, the holdings would seem to be applicable in the context of AI-based systems as well. The most prominent of these cases is *State v Loomis*. After a short introduction to the *Loomis* case, three additional cases (*Malenchik v State, State v Rogers*, and *State v Walls*) will be introduced. The section will end with a summary of the discussion.

State v Loomis

Defendant Loomis pleaded guilty in Wisconsin state court to charges relating to his involvement in a drive-by shooting.[267] He challenged the state's use of the risk assessment portion of the COMPAS report at sentencing. In determining his sentence, the court relied in part on the fact that Loomis had been 'identified, through the COMPAS assessment, as an individual who is at high risk to the community'. Loomis argued that the use of the COMPAS risk assessment violated his right to due process for three reasons: first, it violated his right to be sentenced on the basis of accurate information; second, it violated his right to an individualized sentence; and third, it improperly used gendered assessments in sentencing.

In its 2016 holding, the Wisconsin Supreme Court rejected all of Loomis's due process challenges: First, the variables used by the COMPAS algorithms were publicly available and the outcome of the risk assessment was based entirely on Loomis's answers to the questions or on publicly available information. Due process was satisfied because Loomis had 'the opportunity to verify that the questions and answers listed on the COMPAS report were accurate'. Second, while the COMPAS assessment did involve group data, the assessment was only one of multiple factors considered by the sentencing court so that Loomis received an individualized sentence.[268] Third, COMPAS's use of gender in calculating risk scores did not violate any due process rights since its use simply accounted for differences in recidivism rates between men and women; also, there was no proof that the court actually relied on gender as a factor in sentencing.

---

[267] *State v Loomis*, 881 NW2d 749 (Wis 2016), cert denied, *Loomis v Wisconsin*, 137 S Ct 2290 (June 26, 2017). For a detailed summary of the *Loomis* case, see Recent Cases, 'Criminal Law - Sentencing Guidelines - Wisconsin Supreme Court Requires Warning before Use of Algorithmic Risk Assessments in Sentencing - *State v Loomis*, 881 NW2d 749 (Wis 2016)' (2017) 130 Harv L Rev 1530.
[268] *State v Loomis*, 881 NW2d 749, 765 (Wis 2016).

Despite denying Loomis's claims, the court expressly recognized the finding that risk assessment tools may not perform as well for non-whites as for whites. It also pointed out that the accuracy of such tools, without constant re-norming, is short-lived. As a result, it established the requirement that all Presentence Investigation Reports containing a COMPAS risk assessment inform the sentencing court of the following cautions regarding the risk assessment's accuracy:

> - The proprietary nature of COMPAS has been invoked to prevent disclosure of information relating to how factors are weighed or how risk scores are determined.
>
> - Because COMPAS risk assessment scores are based on group data, they are able to identify groups of high-risk offenders –not a particular high-risk individual.
>
> - Some studies of COMPAS risk assessment scores have raised questions about whether they disproportionately classify minority offenders as having a higher risk of recidivism
>
> - A COMPAS risk assessment compares defendants to a national sample, but no cross-validation study for a Wisconsin population has yet been completed. Risk assessment tools must be constantly monitored and re-normed for accuracy due to changing populations and subpopulations.
>
> - COMPAS was not developed for use at sentencing, but was intended for use by the Department of Corrections in making determinations regarding treatment, supervision, and parole.[269]

In sum, the *Loomis* opinion 'essentially implemented a mandatory disclaimer on the practice of using a COMPAS risk assessment at sentencing'.[270] It also stressed that risk scores may not be used as the sole determinative factor in sentencing. Emphasis of this point was expressed in a concurring opinion: 'consideration of COMPAS is permissible; reliance on COMPAS for the sentence imposed is not permissible.'[271] Loomis appealed the decision to the United States Supreme Court, but the Court declined to hear the case.[272]

Malenchik v State

Defendant Malenchik challenged the trial court's use of the results of two risk assessment tests at sentencing, both of which indicated that Malenchik was at high risk of

[269] *State v Loomis*, 881 NW2d 749, 769-770 (Wis 2016).
[270] Chris Miller, 'The Prospects of Constitutional Challenges to COMPAS Risk Assessment (26 April 2021)' <www.americanbar.org/groups/litigation/committees/privacy-data-security/articles/2021/constitutional-challenges-compas-risk-assessment> accessed 28 March 2023.
[271] *State v Loomis*, 881 NW2d at 774 (Roggensack, CJ, concurring).
[272] *Loomis v Wisconsin*, 137 S Ct 2290 (26 June 2017).

recidivism.[273] In its 2010 holding that a trial court can properly 'supplement and enhance' its evaluation of the evidence before it at sentencing by considering assessment tool scores, the Indiana Supreme Court stressed that the sentence imposed by the trial court was not based solely on the risk assessments but that other factors had also been considered (eg, the defendant's prior criminal history, unwillingness to change his behavior, and refusal to accept responsibility for his actions); furthermore, it pointed out that the trial court had not relied on either test as an independent aggravating factor.

State v Rogers

This case addressed the question of whether a defendant was entitled to reconsideration of a sentence if the sentence was imposed without the use of a risk assessment instrument.[274] While West Virginia's highest court in 2015 denied the motion on procedural grounds, a concurring opinion sought to clarify the role of risk and needs assessments in relation to sentencing: accordingly, a risk and needs assessment 'is merely a tool that may be used by [trial court] judges during sentencing' and '[trial court] judges are not required to consider or use any of the information contained in [such an] assessment.'[275]

State v Walls

Defendant Walls' sentence reflected the risk assessment report that deemed him a high-risk, high-needs probation candidate, but the sentencing court refused to make the report available to defense counsel.[276] Walls challenged the sentence, arguing that he had a statutory and constitutional right to review and verify the question, answers, and scoring decisions contained in the report. In 2017, the Kansas Court of Appeals found in his favor: Depriving him of the report 'necessarily denied him the opportunity to challenge the accuracy of the information upon which the court was required to rely in determining the conditions of his probation'. Since a defendant has a right to an 'effective opportunity to rebut the allegations likely to affect the sentence', the sentencing court's decision to deny him access to the output of the risk assessment tool on which it had relied in setting his sentence violated Wall's right to procedural due process.

Summary

Many legal scholars point out with approval the fact that courts 'appear to have taken pains to emphasize that [algorithmic assessment] tools only serve as one of multiple factors that a judge takes into account in reaching a decision'.[277] One law professor, for example, has stated that 'it would be a dark future if computer algorithms ever replaced

---

[273] *Malenchik v State*, 928 NE2d 564 (Ind 2010).
[274] *State v Rogers*, No 14-0373, 2015 W Va LEXIS 3, 2015 WL 869323 (W Va, 9 January, 2015) (memorandum decision).
[275] ibid.
[276] *State v Walls*, 2017 Kan App Unpub LEXIS 487; 396 P3d 1261 (Kan App 2017).
[277] Coglianese and Ben Dor (n 171) 811. See also Bagaric and others (n 173) 134.

a judge's sentencing decision' and that she 'can't imagine that a risk tool alone could produce just verdicts'. In her opinion, 'the judicial function can't be outsourced to a math problem.'[278] In contrast, another law professor sees this differently. In his view, the results of well-constructed risk assessment instruments are superior to lay judgments and should be given presumptive effect. 'Unfortunately', he writes, 'that rarely occurs'; instead, judges see the results of risk assessment instruments as mere tools and themselves as the definitive answer. This scholar views critically the holdings of judicial decisions (such as *Loomis* and *Malenchik*) according to which 'the results of a [risk assessment instrument] are but one factor to consider and should not be dispositive.'[279] He argues that:

> Judges and parole boards are clearly the ultimate decision-makers about offender risk. But they should be aware that evaluator, judicial, and parole board adjustments to [a risk assessment instrument] usually do not improve on the actuarial assessment. In fact, consistent with the studies comparing actuarial and clinical judgment, several studies find that professional "overrides" of [a risk assessment instrument's] risk estimate, whether by judges, probation officers, or other correctional professionals, decrease accuracy in predicting offending.[280]

While underscoring the superiority of risk assessment systems, the author also advocates for transparency, specifically, for risk algorithms to be made available for evaluation so as to enable defendants to engage in meaningful challenges to the results of risk assessments. Transparency is still suboptimal, both with regard to COMPAS – '[T]he company that produces the COMPAS refuses to reveal its algorithm or the weights assigned to risk factors, claiming trade secret protection.' – and the purportedly publicly developed PATTERN – 'Congress required that the PATTERN be made public, but did not require that the validation procedure that led to development of the instrument nor the data underlying it be disclosed.'[281] And he points out that 'the integration of sophisticated machine learning into [risk assessment instrument] construction could make matters worse.'[282]

2.2.5 Case law that addresses AI-based systems used for predictive justice: Civil courts

Various aspects of risk assessment tools used for predictive justice have claimed the attention of civil courts. Two such cases will be mentioned here. The first, *Henderson v Stensberg*, raised equal protection claims involving the use of COMPAS in a parole

---

[278] Joe Forward, The Loomis Case: The Use of Proprietary Algorithms at Sentencing, The InsideTrack (July 19, 2017), www.wisbar.org/newspublications/insidetrack/pages/article.aspx?Volume=9&ArticleID=25730 (quoting Prof. Cecelia Klingele).

[279] Christopher Slobogin, 'Preventive Justice: How Algorithms, Parole Boards, and Limiting Retributivism Could End Mass Incarceration' (2021) 56 Wake Forest L Rev 97, 138.

[280] Slobogin (n 279) 138.

[281] Slobogin (n 279) 164.

[282] Slobogin (n 279) 164.

decision. The second, *Rodgers v Christie*, raised products liability claims regarding a risk assessment tool. Neither claim was successful.

Henderson v Stensberg

Plaintiff Henderson, incarcerated in Wisconsin, was denied parole in 2015.[283] He argued that prison officials discriminated against him and other Black prisoners by using COMPAS to assess their suitability for parole. Among other things, he brought Fourteenth Amendment equal protection claims in federal court against those prison officials as well as against the company that developed COMPAS. The judge granted summary judgment to the defendants. He expressly stated that Henderson's equal protection claims were not foreclosed and acknowledged that 'there is growing concern that risk-assessment algorithms unfairly disadvantage Black offenders.'[284] But he granted summary judgment because 'Henderson's recidivism score was the lowest possible'; he could not 'show that his COMPAS recidivism score was the reason he was denied parole'; and he thus 'failed to adduce admissible evidence that he was harmed by his COMPAS assessment or that he was denied parole for a discriminatory reason'.[285]

Rodgers v Christie

In 2017, Christian Rodgers[286] was murdered, allegedly by a man who, days before, had been granted pretrial release by a state court due to a decision informed in part by the court's use of a risk estimation tool (the Public Safety Assessment, PSA).[287] The victim's mother brought products liability claims in federal court against the foundation responsible for developing the PSA, alleging that the tool was designed in a defective manner. The Third Circuit, ruling in 2020, affirmed dismissal of the case, however, holding that the PSA is not a 'product' pursuant to the New Jersey Products Liability Act.[288]

Summary

As of yet, there is not much relevant legal scholarship assessing these rulings: commentators have yet to weigh in on the *Henderson* opinion of March 2021. As for the *Rodgers* decision, it was described by a law professor in 2021 as the only case to date involving accusations that the PSA tool harmed a third party not involved in a criminal

---

[283] *Henderson v Stensberg*, 2021 US Dist LEXIS 58010 (WD Wis, 26 March 2021).
[284] *Henderson v Stensberg*, 2021 US Dist LEXIS 58010 at *2.
[285] *Henderson v Stensberg*, 2021 US Dist LEXIS 58010 at *2, *18.
[286] *Rodgers v Christie*, 795 F App'x 878 (3d Cir 2020) (note that the disposition of this case is not an opinion of the full Court and does not constitute biding precedent).
[287] While Coglianese and Ben Dor refer to the Public Safety Assessment as a non-learning algorithmic tool (see Coglianese and Ben Dor (n 171) 803), the outcome of the case is nevertheless of interest in the context of this study.
[288] *Rodgers v Christie*, 795 F App'x 878 (3d Cir 2020) (note that the disposition of this case is not an opinion of the full Court and does not constitute biding precedent).

matter. The scholar argued that both the *Loomis* and the *Henderson* decisions show that the lack of transparency of algorithms restricts the ability of plaintiffs to be heard and to prepare plausible causes of action.[289]

2.2.6 *Laws governing reliability, impartiality, equality, and adaptability of AI-based predictive justice*

In the United States, reliability, impartiality, equality, and adaptability are not specifically addressed by federal legislation for the context of AI-based predictive justice (risk assessment). These or related issues have, however, been addressed indirectly in a recent executive order, in a bill introduced in a state legislature (Massachusetts), and in model legislation (EPIC).

In May 2022, President Joe Biden issued an executive order entitled 'Advancing Effective Accountable Policing and Criminal Justice Practices to Enhance Public Trust and Public Safety'.[290] The order directs the National Academy of Sciences to conduct and publish a study of – among other things – predictive algorithms, with a particular focus on the use of such algorithms by federal law enforcement agencies. The study must assess concerns in the areas of privacy, civil rights, civil liberties, accuracy, or disparate impact that arise in association with the use of such algorithms. Subsequently, the study will be used to make any necessary changes to Federal law enforcement practices.[291]

In February 2022, an 'Act Establishing a Commission on Automated Decision-Making by Government in the Commonwealth' was introduced in the Massachusetts legislature.[292] Virtually the same bill, Bill S.33, was introduced in early 2023.[293] The act would establish a commission to study and make recommendations related to the use in Massachusetts of automated decision systems that may affect human welfare, including the legal rights and privileges of individuals. It describes the responsibilities and composition of the commission and lists the reporting requirements with which it would have to comply.

In particular, the commission would undertake a survey of

> (b) the training specific Massachusetts offices provide to individuals using automated decision systems, and the procedures for enforcing the principles, policies, and guidelines regarding their use;

---

[289] Sonia M Gipson Rankin, 'Technological Tethereds: Potential Impact of Untrustworthy Artificial Intelligence in Criminal Justice Risk Assessment Instruments' (2021) 78 Wash & Lee L Rev 647, 705–706.
[290] Exec Order No 14074 of 25 May 2022, Advancing Effective, Accountable Policing and Criminal Justice Practices to Enhance Public Trust and Public Safety, 87 Fed Reg 32945 (31 May 2022).
[291] See The White House, 'Fact Sheet' (*The White House*, 25 May 2022) <www.whitehouse.gov/briefing-room/statements-releases/2022/05/25/fact-sheet-president-biden-to-sign-historic-executive-order-to-advance-effective-accountable-policing-and-strengthen-public-safety> accessed 29 March 2023. See also Exec Order No 14074 of 25 May 2022, s 13 (d).
[292] Bill S.2688, <https://malegislature.gov/Bills/192/S2688> accessed 29 March 2023.
[293] Bill S.33, <https://malegislature.gov/Bills/193/S33> accessed 29 March 2023.

(c) the manner by which Massachusetts offices validate and test the automated decision systems they use, and the manner by which they evaluate those systems on an ongoing basis …;

(d) matters related to the transparency, explicability, auditability, and accountability of automated decision systems in use in Massachusetts offices …;

(e) the manner and extent to which Massachusetts offices make the automated decision systems they use available to external review …; and

(f) procedures and policies in place to protect the due process rights of individuals directly affected by Massachusetts offices' use of automated decision systems, including but not limited to public disclosure and transparency procedures.[294]

The commission would also consult with experts in the fields of machine learning, algorithmic bias, algorithmic auditing, and civil and human rights[295] and would examine research related to the use of automated decision systems that directly or indirectly result in disparate outcomes for individuals or communities based on an identified group characteristic.[296]

Sometime before 2020, EPIC developed a model law for state AI commissions.[297] In its 'findings' section, the model law proposes that the enacting legislature find that the state has begun to deploy AI and other automated decision systems in numerous areas, including in the area of criminal law; that there is an inherent risk of bias and inaccuracy in the use of these technologies; that there is limited public knowledge about the systems; and that existing regulation of automated decision systems is insufficient.[298] The model law calls for the creation of an 18-person commission to carry out a two-phased study.[299] Phase one would involve the reviewing and cataloguing of how algorithms or other automated decision systems are being used by the state, including:

the identity of the developer and pertinent contract terms between the state and the developer; any state bodies or subdivisions using automated decision systems; the inputs used; the source of the inputs used; the purposes for which such systems are used; the validation policies, the logic of the automated decision system; the data maintenance and deletion policies; and the potential harms that could arise from the use of the system and how those risks are currently addressed.[300]

294 Bill S.33 Sec. 11. (b)(i)(b)-(f).
295 Bill S.33 Sec. 11. (b)(ii).
296 Bill S.33 Sec. 11. (b)(iii).
297 EPIC, 'Model State Artificial Intelligence Commission Law' <https://archive.epic.org/EPIC-Model-State-AI-Commission-Bill.pdf> accessed 28 March 2023.
298 Model State (n 297) s 2.
299 Model State (n 297) s 4(b).
300 Model State (n 297) s 4(g)(1).

In a second phase, the commission would propose recommendations regarding, among other things, minimum technological standards for all automated decision systems; uniform data security provisions; procedures by which individuals affected by a decision made by an automated decision system used by the state could seek information concerning that decision; procedures by which individuals could seek human review of automated decisions made about them; procedures to ensure that automated decision system do not reflect unfair bias or make impermissible discriminatory decisions; procedures to ensure that such systems are adequately evaluated; procedures to ensure the accuracy, reliability, and validity of decisions made by such systems; and procedures to establish data provenance.[301]

2.2.7 *Restrictions on marketing AI-based systems for predictive justice*

There is no federal law governing the marketing of AI-based systems for predictive justice nor is there a federal law that imposes technological requirements on producers of AI-based systems for predictive justice. There is no federal law that requires producers of AI-based systems for predictive justice to consult criminal justice professionals regarding the design of the software, and there is no federal law that requires producers of AI-based systems for predictive justice to regularly monitor and update the software. Finally, there is no federal law governing the certification or labelling of AI-based systems for predictive justice.

In *State v Loomis*, the Wisconsin Supreme Court recognized in 2016 that the accuracy of risk assessment tools, without constant re-norming, is short-lived. In its holding, the court required all presentence investigation reports submitted to sentencing judges that contain a COMPAS risk assessment to include a 'written advisement'. The purpose of the advisement was to inform the sentencing court (among other things) that risk assessment tools 'must be constantly monitored and re-normed for accuracy due to changing populations and subpopulations'.[302] It should be noted that the holding is binding only in Wisconsin and that it does not require the monitoring and re-norming of risk assessment tools; it requires only that the advisement regarding the accuracy of risk assessment tools be included in presentence investigation reports that contain a COMPAS risk assessment.

2.2.8 *Training of professionals who rely on AI-based systems*

Of the many jurisdictions that use risk assessment instruments (some of which are – or may soon be – based on AI), very few train judges, lawyers, and correctional officials in their use.[303] As far as the use of risk assessment instruments by sentencing judges is concerned, their discretion continues to play an important role, and very little

[301] Model State (n 297) s 4(g)(2).
[302] *State v Loomis*, 881 NW2d 749, 769 (Wis 2016).
[303] Slobogin (n 279) 168.

information is available about how judges actually use these risk assessments in practice.[304]

2.2.9 Transparency and the technological functioning of AI-based systems

There is no federal law guaranteeing the transparency of the technological functioning of AI-based systems for predictive justice. Generally speaking, companies are allowed to claim their technology is a trade secret and can refuse to be transparent about how their product works. In *State v Loomis*, for example, the defendant requested access to information concerning the inner workings of the COMPAS tool, but the Wisconsin Supreme Court denied the request: the court permitted the proprietary nature of the COMPAS tool – as asserted by its developer, Northpointe, Inc. – to prevent disclosure of information about how factors are weighed or how risk scores are determined.[305]

There is, however, a statute in the state of Idaho,[306] enacted in 2019, that addresses the transparency, accountability, and explainability of pretrial risk assessment tools. Pursuant to the statute, all pretrial risk assessment tools must be transparent; information used to build or validate such tools must be open to public inspection; parties to criminal cases in which such a tool has been relied upon are entitled to review calculations and data used to calculate the defendant's risk score; and builders and users of such tools may not assert trade secret or other intellectual property protections in order to quash discovery of information used in the development or validation of such tools.[307]

In contrast to the transparency required by statute in Idaho at the pretrial stage, Massachusetts does not require such transparency at the parole stage. According to an

---

304 Garrett and Monahan (n 175) 43. See discussion at 1.3. above.

305 *State v Loomis*, 881 NW2d at 761, 769 (Wis 2016).

306 Idaho Code § 19-1910 (2022).

307 Idaho Code § 19-1910 (2022). Pretrial Risk Assessment Tools.

'(1) All pretrial risk assessment tools shall be transparent, and:

(a) All documents, data, records, and information used by the builder to build or validate the pretrial risk assessment tool and ongoing documents, data, records, and written policies outlining the usage and validation of the pretrial risk assessment tool shall be open to public inspection, auditing, and testing;

(b) A party to a criminal case wherein a court has considered, or an expert witness has relied upon, a pretrial risk assessment tool shall be entitled to review all calculations and data used to calculate the defendant's own risk score; and

(c) No builder or user of a pretrial risk assessment tool may assert trade secret or other intellectual property protections in order to quash discovery of the materials described in paragraph (a) of this subsection in a criminal or civil case.

(2) For purposes of this section, "pretrial risk assessment tool" means a pretrial process that creates or scores particular factors in order to estimate a person's level of risk to fail to appear in court, risk to commit a new crime, or risk posed to the community in order to make recommendations as to bail or conditions of release based on such risk, whether made on an individualized basis or based on a grid or schedule.'

amicus brief filed by EPIC[308] in the case of *Jose Rodriguez v Massachusetts Parole Board*,[309] parole applicants in Massachusetts are given only a redacted version of the report provided by the predictive analytical tool in use in that state (LS/CMI[310]). They are not given information about the sources of data that went into their assessment nor are they given information about the logic of the tool or about the role the report played in their parole decisions. Furthermore, they and the public are unable to access even blank scoresheets, scoring guides, training manuals or validation studies. EPIC argued that Massachusetts' lack of transparency concerning use of its predictive analytical tool prevents the public from fully understanding the tool's accuracy and potential for bias and prevents inmates from understanding how the tool decided their recidivism risk and whether those decisions were accurate. While EPIC stated that the predictive analytical tool in use in Massachusetts 'seems' to fall into the category of 'checklist-type tools' rather than the more advanced category of tools that use machine learning, the argumentation in favor of transparency is equally if not more applicable to AI-based tools.

2.2.10 Transparency and the use of AI-based systems for predictive justice

There are no federal rules specifically governing the right of affected individuals to be informed about the use of AI-based systems for predictive justice. Other general rules concerning the right to be informed may, however, apply. In New York State, for example, inmates whose application for parole are denied have a statutory right to be informed in writing of the factors and reasons for such denial, and 'such reasons shall be given in detail and not in conclusory terms.'[311] This right will not always suffice, however, to enable inmates to challenge all the factors that have contributed to their parole denials. Take, for example, inmate Glenn Rodriguez, who was granted parole in 2017 but whose previous parole application was denied on the basis of a COMPAS 'high risk' ranking:

> When inmate Glenn Rodriguez was denied parole, he had a statutory right to be informed in writing of the "factors and reasons" for the denial." Rodriguez filed a grievance showing that there was an error in one of the inputs used to generate his risk assessment score. The tool relies on manual inputs from surveys filled out by a human evaluator. In Rodriguez's case, the evaluator had checked "yes"

---

[308] Benjamin Winters, 'Brief of Amicus Curiae,' *Rodriguez v Massachusetts Parole Board*, SJC-13197 (*Elec Priv Info Ctr*, 14 February 2022) <https://epic.org/documents/rodriguez-v-massachusetts-parole-board> accessed 28 March 2023.

[309] SJC-13197, available at <www.ma-appellatecourts.org/docket/SJC-13197> accessed 30 March 2023. In September 2022, the Massachusetts Supreme Judicial Court affirmed the judgment of the superior court in favor of the parole board, holding that the superior court was correct in affirming the board's decision to deny Rodriguez release on parole. Slip opinion available at <https://cases.justia.com/massachusetts/supreme-court/2022-sjc-13197.pdf?ts=1662552216> accessed 30 march 2023.

[310] Level of Service/Case Management Inventory.

[311] NY Exec Law § 259-i(2)(a) (McKinney 2018).

> where he should have checked "no" in one survey response. Rodriguez knew that when another inmate had received a reassessment to correct the same error, that person's final risk score dropped significantly. But Rodriguez could not prove that the error had any significant effect in his own case because the weights of the input variables are alleged trade secrets. Ultimately, he was unable to convince anyone to correct the mistake and had to return to the parole board six months later with the same erroneous score.[312]

While the problems Glenn Rodriguez encountered involved interactions between human error and trade secrets, the combination of AI and trade secrets, it would seem, would pose inmates with even more challenging situations. (See also discussion of Jose Rodriguez at 2.9. above.)

2.3 General principles of law

2.3.1 Right to equality (right to non-discrimination) with regard to AI-based systems used for predictive justice

In the context of AI-based systems used for predictive justice, there is lively discussion in the United States about equality and non-discrimination – issues of interest from many perspectives, including that of constitutional law. This discussion is taking place both in the popular media[313] and in the academic community.[314] Authors (primarily law faculty members) of a law review article published in 2022, for example, who point to the wide acceptance of the principle of equality before the law as a fundamental tenet of justice,[315] recognize that the use of algorithms in the criminal justice system has both positive and negative effects on the realization of this principle. Citing an opinion piece written by two professors for the Boston Globe in 2019,[316] these authors opine that 'eliminating algorithms – or reducing their use – would cause "more, not less, discrimination.' And, citing an article published in a nonprofit news organization in 2019, they state that 'the bias present in algorithms presents less of an obstacle than human bias because it can

---

[312] Rebecca Wexler, 'Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System' (2018) 70 Stan L Rev 1343, 1371 (footnotes omitted). See also Rebecca Wexler, 'Code of Silence: How Private Companies Hide Flaws in the Software That Governments Use to Decide Who Goes to Prison and Who Gets Out' (June/July/August 2017) Washington Monthly.

[313] See, eg, Thadaney Israni (n 163); Liptak (n 187); Jens Ludwig and Cass R. Sunstein, 'Discrimination in the Age of Algorithms' The Boston Globe (Boston, 24 September 2019) A8.

[314] See, eg, Itay Ravid and Amit Haim, 'Progressive Algorithms' (2022) 12 UC Irvine L Rev 527; Note, 'Beyond Intent: Establishing Discriminatory Purpose in Algorithmic Risk Assessment' (2021) 134 Harv L Rev 1760; Vincent M. Southerland, 'The Intersection of Race and Algorithmic Tools in the Criminal Legal System' (2021) 80 Md L Rev 487; Aziz Z. Huq, 'Constitutional Rights in the Machine-Learning State' (2020) 105 Cornell L Rev 1875; Huq, 'Racial Equity' (n 169); Michael Simon and others, 'Lola v Skadden and the Automation of the Legal Profession' (2018) 20 Yale JL & Tech 234.

[315] Bagaric and others (n 173) 100; Michael Simon and others, 'Lola v Skadden and the Automation of the Legal Profession' (2018) 20 Yale JL & Tech 234.

[316] Bagaric and others (n 173) 98, citing Ludwig and Sunstein (n 313).

more easily "be observed, studied, and corrected in ways that human bias cannot."'[317] On the other hand, they recognize that the use of algorithms trained with '"past biased data" are likely to recreate the same biases in their decision-making processes, further exacerbating discrimination and unfairness'.[318] In the reform section of their article, they propose that predictive systems 'be developed carefully with a focus on preventing the operation of factors that lead to indirect discrimination' in order to 'minimize the potential for race and other immutable factors to influence the outcomes of risk assessment algorithms'.[319]

In a 2021 law review article, the author, a law professor, defended the use of risk assessment instruments against wide-ranging attacks on accuracy and fairness grounds. His conclusion against claims of egalitarian injustice was that 'with a few caveats, such instruments are not violative of equal protection if they provide relevant and probative results.'[320]

Finally, specifically in the machine-learning context, another law professor reexamined questions of intent and classification – issues at the heart of the constitutional jurisprudence of the Equal Protection Clause and federal antidiscrimination statutes. In a 2020 law review article, he suggested that 'the equality concerns commonly raised by algorithmic systems in practice are better conceptualized in terms of their impact on pernicious social stratification.'[321] And in another 2020 journal article, the co-authors, both law professors, saw the use of machine learning and other forms of AI in the adjudication of criminal proceedings as a 'context in which questions of equity and fairness receive heightened attention'.[322]

2.3.2 AI-based systems and judicial independence

Discussion in the United States on the effects of AI-based systems on judicial independence is not widespread. There are no means or methods designed specifically to guarantee judicial independence in the context of AI usage.

2.3.3 Right of access to a human judge

There is discussion in the United States about whether and under what circumstances there should be a right of access to a human judge.[323] For example, a 2020 law review

---

[317] Bagaric and others (n 173) 98 (citing Matt Henry, 'Risk Assessment: Explained' (*The Appeal*, 25 March 2019) <https://theappeal.org/risk-assessment-explained> accessed 28 March 2023.
[318] Bagaric and others (n 173) 133.
[319] Bagaric and others (n 173) 135.
[320] Slobogin (n 279) 143.
[321] Huq, 'Constitutional Rights' (n 314) 1917.
[322] Carla L Reyes and Jeff Ward, 'Digging into Algorithms: Legal Ethics and Legal Access' (2020) 21 Nev LJ 325, 333.
[323] See, eg, Michael Simon and others, 'Lola v Skadden and the Automation of the Legal Profession' (2018) 20 Yale JL & Tech 234 (citing Joseph Weizenbaum, Computer Power and Human Reason: From Judgment

article, the author, a law professor, pointed to the holding of the Wisconsin Supreme Court in Loomis[324] when he wrote that 'American law is … making tentative moves toward a … right to a human decision.'[325] (Loomis held that a risk score generated by an algorithm cannot, as a matter of due process, 'be considered as the determinative factor in deciding whether the offender can be supervised safely and effectively in the community'.[326]) According to the author, '[The Loomis] decision precludes full automation of bail determinations. There must be a human judge in the loop.'[327] He also argued that there is no reason why it should not be possible to invoke the Sixth Amendment's right to a jury trial[328] to preclude the use of at least some forms of algorithmically generated inputs in criminal sentencing: 'Indeed, it would seem to follow a fortiori that a right precluding a jury's substitution with a judge would also block its displacement by a mere machine.'[329]

The author's own position in this discussion is to favor 'a right to a well-calibrated machine decision'[330] rather than a right to a decision taken by a human judge, in part because 'machines have the capacity to classify and predict with fewer errors than humans.'[331]

> Algorithmic technologies used by machine decisions are still in their infancy. Now, they can be flawed in many ways. It seems too early, however, to assume that human decisions will be globally superior to machine decisions such that a right to the former is warranted. Sometimes the opposite might be true. We should, therefore, at least consider the possibility that under certain circumstances a right to a well-calibrated machine decision might be the better option.[332]

2.3.4 *The presumption of innocence and the use of AI-based systems to establish the probability that a person is dangerous or is likely to reoffend*

There is discussion in the United States about protecting the presumption of innocence when AI-based risk assessment tools are used to determine whether a person is dangerous or is likely to recidivate. In a 2020 law review article, for example, the author explains the role of the constitutional presumption of innocence in the pretrial phase,

---

to Calculation (WH Freeman and Co, 1976), for the argument that computers should not be permitted to make judicial decisions).

324 *State v Loomis*, 881 NW2d 749 (Wis 2016). See discussion at 2.4. above.

325 Aziz Z Huq, 'A Right to a Human Decision' (2020) 106 Va L Rev 611, 617.

326 *State v Loomis*, 881 NW2d 749, 760 (Wis 2016).

327 Huq, 'A Right to a Human Decision' (n 325) 617.

328 US Const amend VI ('In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the state and district wherein the crime shall have been committed ... .').

329 Huq, 'A Right to a Human Decision' (n 325) 617.

330 Huq, 'A Right to a Human Decision' (n 325) 619, 686.

331 Huq, 'A Right to a Human Decision' (n 325) 650.

332 Huq, 'A Right to a Human Decision' (n 325) 687-688.

when risk assessment instruments are used to help decide whether a detainee should be incarcerated or released. She examines the implications of using machine learning to develop the instruments used in this phase of the criminal justice system and asks whether AI-based tools represent a threat to the system. After comparing seven such instruments, two of which (COMPAS and the Kleinberg et al. tool) have a machine-learning component, she concludes that there are 'more similarities than differences' between tools using machine learning and those using regression analysis and that 'adding a machine learning aspect to risk assessment tools will not worsen the outcome, and in many cases may improve it.'[333]

In contrast, in a 2021 journal article on the cascading effect of algorithmic bias in risk assessments, the author came to the following conclusion: 'To the extent the U.S. justice system is predicated on the presumption of innocence, the use of algorithmic tools to predict the probability of future crime in deciding the length of one's sentence is a contradiction.'[334]

2.3.5 Fair trial rights and AI-based systems used for predictive justice

There is widespread discussion in the United States about due process/fair trial issues in the context of AI-based systems used for predictive justice. A 2021 note in the Harvard Law Review, for example, cited case law rejecting a defendant's due process claim against the use of the COMPAS risk assessment systems[335] and argued that with algorithmic systems 'concerns of bias … can infringe upon the individual liberty interest in a fair trial.'[336]

As far as the right to contest decisions made by AI is concerned, the authors of a comprehensive law review article published in 2021 showed that the issue of what rights to an appeal, if any, people should have when they are subjected to decision-making by artificial intelligence is unclear. This lack of clarity exists even though 'the right to challenge decisions with significant effects is a core principle of the rule of law.'[337] Their article reviews suggestions made by numerous legal experts over the past several years and concludes that 'while earlier scholars called for some kind of due process, the recent trend has been to favor systemic governance over the companies or government entities that build and use Al over establishing individual rights such as a right to contest.'[338]

---

[333] Doaa Abu Elyounes, 'Bail Or Jail? Judicial versus Algorithmic Decision-Making in the Pretrial System' (2020) 21 Colum Sci & Tech L Rev 376, 445.
[334] Tim O'Brien, 'Compounding Injustice: The Cascading Effect of Algorithmic Bias in Risk Assessments' (2021) 13 Geo JL & Mod Critical Race Persp 39, 80-81.
[335] Note, 'Beyond Intent' (n 314) 1762 (citing *People v Younglove*, No 341901, 2019 WL 846117, at *3 (Mich Ct App, 21 February 2019).
[336] Note, 'Beyond Intent' (n 314) 1771.
[337] Margot E Kaminski and Jennifer M Urban, 'The Right to Contest AI' (2021) 121 Colum L Rev 1957, 1959-1960.
[338] Kaminski and Urban (n 337) 1984.

2.3.6 *Right to defense against algorithmic calculations*

Robot lawyers are making inroads in the legal profession. Such intelligent machines represent a challenge to the existing liability regime. According to the author of a 2019 law review article, a law professor, human lawyers who fail to deliver competent legal services to their clients are subject to both ethical discipline and malpractice suits.[339] Indeed, 'their responsibility can extend to the actions of third parties' involved in the provision of legal services.[340] When lawyer robots make mistakes, however, the question of who should be held responsible and who should compensate injured clients is still an open one. The author points out that some clients, particularly sophisticated corporate clients, 'are likely to negotiate warranties and other protections into their engagements to shield themselves from any errors resulting from the use of artificial intelligence'.[341] In contrast, to these clients, ordinary individuals, he argues 'are not in a position to negotiate these protections or to assess the quality of the legal services they receive'.[342]

Regarding the hurdles defendants face if they seek meaningfully to challenge algorithmic assessments used to determine their sentences, we turn again to the Loomis case. Here, defendant Loomis challenged the system (COMPAS) that labeled him a high risk for recidivism. The Wisconsin Supreme Court denied his challenge on a number of grounds and, recognizing its protection by trade secret law, did not grant Loomis full access to the COMPAS algorithm. Co-authors of a journal article published in 2020, both law professors, stated the following:

> Despite the [Wisconsin] Supreme Court's decision, concern remains that denying access to the process by which sentencing and other impactful determinations are made represents a due process problem in and of itself. Those engaged in debates at the intersection of law and algorithms employ Loomis' experience and the court's response as a rallying cry for technologies' potential to inject additional inequity into the criminal justice system, rather than less. These concerns amplify as state and federal government institutions adopt technological tools in an increasing number of government-citizen interactions.[343]

2.3.7 *Principles of constitutional law and the use of AI-based systems for predictive justice*

[339] Milan Markovic, 'Rise of the Robot Lawyers' (2019) 61 Ariz L Rev 325, 343.
[340] Markovic (n 339) 343.
[341] Markovic (n 339) 344.
[342] Markovic (n 339) 344.
[343] Reyes and Ward (n 322) 328. See also Justin Snyder, 'RoboCourt: How Artificial Intelligence Can Help Pro Se Litigants and Create a "Fairer" Judiciary' (2022) 10 Ind JL & Soc Equal 200.

There is wide-ranging discussion in the United States about principles of constitutional law affected by the use of AI-based systems for predictive justice.[344] (A number of these principles – due process, equal protection, right to jury trial, etc. – are discussed elsewhere in this report.) In 2020, for example, a group of experts from Harvard (mathematics, economics, and law) published a comprehensive law review article analyzing the constitutional issues presented by the use of risk-assessment technologies – including those based on AI – in the criminal justice system. The issues addressed include the relationship between due process and certain proprietary algorithmic models and the challenges to existing equal protection jurisprudence posed by the discriminatory nature of risk assessment instruments. The article discusses possible ways to challenge the constitutionality of risk assessment technologies in state courts and concludes with suggestions for how to improve the technology and satisfy constitutional standards simultaneously.[345]

2.3.8 Privatization of criminal justice and equality of litigants

There is discussion about the privatization of aspects of criminal justice in the United States.[346] Privatization is particularly problematic in the context of sentencing: 'private developers play a significant part in sentencing determinations without being subject to traditional constitutional accountability mechanisms.'[347]

The question of the equality of litigants in the criminal justice system is also a topic of concern in the United States. The question has been asked whether increased reliance on AI will lead

> to one or more inequitable two-tiered systems. Some fear an eventual system with expensive – but superior – human lawyers and inexpensive – but inferior – AI driven legal assistance. Others fear almost the reverse problem: that AI will be superior to human lawyers but will be expensive and available only to large law firms and their wealthy clients. Still others fear that AI's impact will not overcome the status quo where some can afford legal services while others cannot.[348]

And asymmetries may arise if law enforcement can access data possessed by private companies but investigators for the defense cannot.[349] In such cases, 'law enforcement

---

344 See, eg, Krent and Rucker (n 166) (due process, ex post facto issues); Huq, 'Constitutional Rights' (n 314) (due process, privacy, equality); Andrea Nishi, 'Privatizing Sentencing: A Delegation Framework for Recidivism Risk Assessment' (2019) 119 Colum L Rev 1671 (due process, equal protection).
345 Michael Brenner and others, 'Constitutional Dimensions of Predictive Algorithms in Criminal Justice' (2020) 55 Harv CR-CL L Rev 267.
346 See, eg, Farhang Heydari, 'The Private Role in Public Safety' (2022) 90 Geo Wash L Rev 696.
347 Nishi (n 344) 1688.
348 Drew Simshaw, 'Access to AI Justice: Avoiding an Inequitable Two-Tiered System of Legal Services' (2022) 24 Yale JL & Tech 150, 156 (footnotes omitted).
349 'Privacy Asymmetries: Access to Data in Criminal Defense Investigations' (2021) 68 UCLA L Rev 212, 248

but not defendants will benefit from deploying new algorithmic artificial intelligence and machine learning tools to search and analyze that data.'[350]

References

Abu Elyounes D, 'Bail Or Jail? Judicial versus Algorithmic Decision-Making in the Pretrial System' (2020) 21 Colum Sci & Tech L Rev 376

Am Law Inst, *'Model Penal Code: Official Draft and Explanatory Notes'*, The American Law Institute, 1985

– – 'Model Penal Code: Sentencing'

– – Model Penal Code: Sentencing, Proposed Final Draft, 2017

Angwin J, 'Make Algorithms Accountable' The NY Times (New York, 1 August 2016) A17

Angwin J and others, 'Machine Bias', *ProPublica,* 23 May 2016

Austin D, 'The Justice in Forensic Algorithms Act: Legal Technology Trends', *eDiscovery Today,* 26 May 2022

Bagaric M and others, 'The Solution to the Pervasive Bias and Discrimination in the Criminal Justice System: Transparent and Fair Artificial Intelligence' (2022) 59 Am Crim L Rev 95

Bagaric M, Hunter D, and Stobbs N, 'Erasing the Bias against Using Artificial Intelligence to Predict Future Criminality: Algorithms Are Color Blind and Never Tire' (2020) 88 U Cin L Rev 1037

Bagaric M and Wolf G, 'Sentencing by Computer: Enhancing Sentencing Transparency and Predictability and (Possibly) Bridging the Gap between Sentencing Knowledge and Practice' (2018) 25 Geo Mason L Rev 653

Berk R, 'An Impact Assessment of Machine Learning Risk Forecasts on Parole Board Decisions and Recidivism' (2017) 13 J Experimental Criminology 193

Brand J and Pishko J, 'Bail Reform: Explained', *The Appeal,* 14 June 2018

Brennan T, Dieterich W, and Ehret B, 'Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System' (2009) 36 Crim Just & Behavior 21

[350] Wexler, 'Privacy Asymmetries' (n 349) 248.

Brenner M and others, ‘Constitutional Dimensions of Predictive Algorithms in Criminal Justice’ (2020) 55 Harv CR-CL L Rev 267

Brookings Institution, ‘Forensic Algorithms: The Future of Technology in the US Legal System’, *Brookings,* 12 May 2022

Coglianese C and Ben Dor LM, ‘AI in Adjudication and Administration’ (2021) 86 Brook L Rev 791

Corbett-Davies S and others, ‘A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased against Blacks. It’s Actually Not That Clear’, *The Washington Post,* Washington, D.C., 17 October 2016

Corey E, ‘New Data Suggests Risk Assessment Tools have Little Impact on Pretrial Incarceration’, *The Appeal,* 7 February 2020

Cyphert AB, ‘Reprogramming Recidivism: The First Step Act and Algorithmic Prediction of Risk’ (2020) 51 Seton Hall L Rev 331

Dershowitz AM, ‘Visibility, Accountability and Discourse as Essential to Democracy: The Underlying Theme of Alan Dershowitz’s Writing and Teaching’ (2008) 71 Alb L Rev 731

Desmarais SL and Singh JP, Risk Assessment Instruments Validated and Implemented in Correctional Settings in the United States 2, *Council of State Governments Justice Center,* 27 March 2013

DiBenedetto R, ‘Reducing Recidivism or Misclassifying Offenders: How Implementing Risk and Needs Assessment in the Federal Prison System Will Perpetuate Racial Bias’ (2019) 27 JL & Pol’y 414

DOJ, ‘Department of Justice Announces Enhancements to the Risk Assessment System and Updates on First Step Act Implementation’, 15 January 2020

– – ‘The First Step Act of 2018: Risk and Needs Assessment System – UPDATE’, *Department of Justice,* January 2020

Donohue ME, ‘A Replacement for Justitia’s Scales: Machine Learning’s Role in Sentencing’ (2019) 32 Harv JL & Tech 657

Duke Government Relations, ‘The Need for Transparency and Interpretability at the Intersection of AI and Criminal Justice’, 22 November 2021

EPIC, ‘About Us’ < https://epic.org/about> accessed 29 March 2023

– – ‘AI in the Criminal Justice System’ <https://epic.org/issues/ai/ai-in-the-criminal-justice-system> accessed 28 March 2023

– – 'Liberty at Risk: Pre-Trial Risk Assessment Tools in the US' (September 2020) <https://epic.org/wp-content/uploads/2022/02/Liberty-At-Risk-Report-FALL-2020-UPDATE.pdf> accessed 28 March 2023

– – 'Model State Artificial Intelligence Commission Law' <https://archive.epic.org/EPIC-Model-State-AI-Commission-Bill.pdf> accessed 28 March 2023

Flores AW, Bechtel K, and Lowenkamp CT, 'False Positives, False Negatives, and False Analyses: A Rejoinder to Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks' (2016) 80 Fed Probation 38

Forward J, 'The Loomis Case: The Use of Proprietary Algorithms at Sentencing' InsideTrack, 19 July 19 2017,

Garrett B and Monahan J, 'Assessing Risk: The Use of Risk Assessment in Sentencing' (2019) 103 Judicature 42

Giarda R, 'International: Artificial Intelligence in the Administration of Justice', *LegalBytes,* January 2022

Gipson Rankin SM, 'Technological Tethereds: Potential Impact of Untrustworthy Artificial Intelligence in Criminal Justice Risk Assessment Instruments' (2021) 78 Wash & Lee L Rev 647

Glueck S, 'Predictive Devices and the Individualization of Justice' (1958) 23 Law & Contemp Probs 461

Hamilton M, 'Algorithmic Risk Assessment: A Progressive Policy in Pretrial Release' (2021) 57 Idaho L Rev 615

– – 'Evaluating Algorithmic Risk Assessment' (2021) 24 New Crim L Rev 156

Henry M, 'Risk Assessment: Explained', *The Appeal,* 25 March 2019

Heydari F, 'The Private Role in Public Safety' (2022) 90 Geo Wash L Rev 696

Huq AZ, 'Racial Equity in Algorithmic Criminal Justice' (2019) 68 Duke LJ 1043

– – 'Constitutional Rights in the Machine-Learning State' (2020) 105 Cornell L Rev 1875

– – 'A Right to a Human Decision' (2020) 106 Va L Rev 611

Hutson M, 'Artificial Intelligence Prevails at Predicting Supreme Court Decisions', *Science,* 2 May 2017

Kaminski ME and Urban JM, 'The Right to Contest AI' (2021) 121 Colum L Rev 1957

Kehl D, Guo P, and Kessler S, 'Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing' (2017)

Kranzberg M, ‘Technology and History: “Kranzberg’s Laws”’ (1986) 27 Technology and Culture 544

Kraus RC, ‘Artificial Intelligence Invades Appellate Practice: The Here, The Near, and The Oh My Dear’ (2019 Winter Edition) Appellate Issues

Krent HJ and Rucker R, ‘The First Step Act - Constitutionalizing Prison Release Policies’ (2022) 74 Rutgers UL Rev 631

Larson J and others, ‘How We Analyzed the COMPAS Recidivism Algorithm’, ProPublica, 23 May 2016

Lee Park A, ‘Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing’, UCLA Law Review, 19 February 2019

LexisNexis, ‘LexisNexis Announces Acquisition of Ravel Law’ (8 June 2017)

Lightbourne J, ‘Damned Lies & Criminal Sentencing Using Evidence-Based Tools’ (2017) 15 Duke L & Tech Rev 327

Legal Information Institute, ‘Executive Order’ (*Legal Information Institute*) <www.law.cornell.edu/wex/executive_order> accessed 28 March 2023

Liptak A, ‘Sent to Prison by a Software Program’s Secret Algorithms’ *The NY Times* (New York, 1 May 2017), A22

Ludwig J and Sunstein CR, ‘Discrimination in the Age of Algorithms’ *The Boston Globe* (Boston, 24 September 2019) A8

Mapping Pretrial Injustice, ‘Common Pretrial Risk Assessments’ <https://pretrialrisk.com/the-basics/common-prai> accessed 28 March 2023

– – ‘How Many Jurisdictions Use Each Tool?’ <https://pretrialrisk.com/national-landscape/how-many-jurisdictions-use-each-tool> accessed 28 March 2023

– – ‘State Laws on Rats’ <https://pretrialrisk.com/national-landscape/state-laws-on-rats> accessed 28 March 2023

Markovic M, ‘Rise of the Robot Lawyers’ (2019) 61 Ariz L Rev 325

Mayson SG, ‘Dangerous Defendants’ (2017) 127 Yale LJ 490

Metz A and others, ‘Valid or Voodoo: A Qualitative Study of Attorney Attitudes Towards Risk Assessment in Sentencing and Plea Bargaining’ 8 (10 March 2020) Virginia Public Law and Legal Theory Research Paper No 2020-25, Duke Law School Public Law & Legal Theory Series No 2020-15

Miller C, ‘The Prospects of Constitutional Challenges to COMPAS Risk Assessment (26 April 2021)’, ABA Privacy and Data Security

National Institute of Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool* (No 303859, December 2021)

Nishi A, 'Privatizing Sentencing: A Delegation Framework for Recidivism Risk Assessment' (2019) 119 Colum L Rev 1671

Note, 'Beyond Intent: Establishing Discriminatory Purpose in Algorithmic Risk Assessment' (2021) 134 Harv L Rev 1760

Northpointe Research and Development Department, 'COMPAS Scales and Risk Models – Validity and Reliability: A Summary of Results from Internal and Independent Studies', *Elec Priv Info Ctr,* 20 July 2010

O'Brien T, 'Compounding Injustice: The Cascading Effect of Algorithmic Bias in Risk Assessments' (2021) 13 Geo JL & Mod Critical Race Persp 39

Oleson JC, 'Risk in Sentencing: Constitutionally Suspect Variables and Evidence-Based Sentencing' (2011) 64 SMU L Rev 1329

Partnership on AI, 'Report on Algorithmic Risk Assessment Tools in the US Criminal Justice System', 23 April 2019

– – 'About Us' <https://partnershiponai.org/about/> accessed 29 March 2023

Pattison-Gordon J, 'Justice-Focused Algorithms Need to Show Their Work, Experts Say' (12 May 2022) government technology

Polonski S, 'Mitigating Algorithmic Bias in Predictive Justice: 4 Design Principles for AI Fairness' (24 November 2018) Towards Data Science

ProPublica, 'About Us' <www.propublica.org/about> accessed 28 March 2023

PRWeb, 'Ravel Law Announces Unprecedented Judge Analytics Offering' (16 April 2015) <www.prweb.com/releases/2015/04/prweb12656883.htm> accessed 28 March 2023

Public Safety Risk Assessment Clearinghouse, 'Correctional Offender Management Profiling for Alternative Sanctions (COMPAS)'

Rabinovich-Einy O and Katsh E, 'Artificial Intelligence and the Future of Dispute Resolution – The Age of AI-DR' in MA Wahab, D Rainey, and E Katsh (eds*), Online Dispute Resolution: Theory and Practice* (2nd edn, Eleven International Publishing 2021) pp. 471-488

Ravid I and Haim A, 'Progressive Algorithms' (2022) 12 UC Irvine L Rev 527

Recent Cases, 'Criminal Law - Sentencing Guidelines - Wisconsin Supreme Court Requires Warning before Use of Algorithmic Risk Assessments in Sentencing - State v. Loomis, 881 N.W.2d 749 (Wis. 2016)' (2017) 130 Harv L Rev 1530

Remus D and Levy F, ‘Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law’ (2017) 30 Geo J Legal Ethics 501

Reyes CL and Ward J, ‘Digging into Algorithms: Legal Ethics and Legal Access’ (2020) 21 Nev LJ 325

Rigano C, ‘Using Artificial Intelligence to Address Criminal Justice Needs’ (January 2019) NIJ Journal 280

Santos M, ‘PATTERN Risk and Needs Assessment Under First Step Act’ (*Prison Professors*)

Schaefer TB, ‘The Ethical Implications of Artificial Intelligence in the Law’ 55 Gonz L Rev 221

Schmitz AJ and Martinez J, ‘ODR and Innovation in the United States’ in MA Wahab, D Rainey, and Ethan Katsh (eds), *Online Dispute Resolution: Theory and Practice* (2nd edn, Eleven International Publishing 2021) pp. 611-638

Simon M and others, ‘Lola v Skadden and the Automation of the Legal Profession’ (2018) 20 Yale JL & Tech 234

Simshaw D, ‘Access to AI Justice: Avoiding an Inequitable Two-Tiered System of Legal Services’ (2022) 24 Yale JL & Tech 150

Slobogin C, ‘Preventive Justice: How Algorithms, Parole Boards, and Limiting Retributivism Could End Mass Incarceration’ (2021) 56 Wake Forest L Rev 97

Snyder J, ‘RoboCourt: How Artificial Intelligence Can Help Pro Se Litigants and Create a “Fairer” Judiciary’ (2022) 10 Ind JL & Soc Equal 200

Southerland VM, ‘The Intersection of Race and Algorithmic Tools in the Criminal Legal System’ (2021) 80 Md L Rev 487

Tashea J, ‘Courts Are Using AI to Sentence Criminals. That Must Stop Now’ 17 April 2017, Wired

Taylor AM, ‘AI Prediction Tools Claim to Alleviate and Overcrowded American Justice System … But Should They Be Used?’ Stanford Politics (13 September 2020)

Thadaney Israni E, ‘When an Algorithm Helps Send You to Prison’ *The NY Times* (New York, 26 October 2017)

Threadcraft-Walker W and others, ‘Gender, Race/Ethnicity and Prediction: Risk in Behavioral Assessment’ (2018) 54 J Crim Just 12

Turner S and others, ‘Development of the California Static Risk Assessment (CSRA): Recidivism Risk Prediction in the California Department of Corrections and

Rehabilitation' (September 2013) UC Irvine Center for Evidence-Based Correction Working Paper

Vought RT, Memorandum for the Heads of Executive Departments and Agencies, *Guidance for Regulation of Artificial Intelligence Applications* (17 November 2020)

Ward J, 'Black Box Artificial Intelligence and the Rule of Law' (2021) 84 Law & Contemp Prob

Washington AL, 'How to Argue with an Algorithm: Lessons from the COMPAS-ProPublica Debate' (2018) 17 Colo Tech LJ 131

Weizenbaum J, *Computer Power and Human Reason: From Judgment to Calculation* (WH Freeman and Co, 1976)

Wexler R, 'Code of Silence: How Private Companies Hide Flaws in the Software That Governments Use to Decide Who Goes to Prison and Who Gets Out' (June/July/August 2017) Washington Monthly

– – 'Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System' (2018) 70 Stan L Rev 1343

– – 'Privacy Asymmetries: Access to Data in Criminal Defense Investigations' (2021) 68 UCLA L Rev 212

The White House, 'Fact Sheet' (*The White House,* 25 May 2022)

The White House, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People (October 2022) <www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf> accessed 29 March 2023

Winters B, 'Brief of Amicus Curiae,' Rodriguez v. Massachusetts Parole Board, SJC-13197 (*Elec Priv Info Ctr,* 14 February 2022)