Law Society of Ontario | Barreau de l'Ontario

TAB 4

AI for Litigators

Litigating Algorithms in Criminal Law: A Defence Perspective

Gerald Chan
Stockwoods LLP

April 16, 2021

cpd

*This paper was originally presented at the LSO Special Lectures 2019: Innovation, Technology and the Practice of Law (November 21, 2019)

Litigating Algorithms in Criminal Law: A Defence Perspective

By Gerald Chan[1]

I. Introduction

It is often said that lawyers went to law school because their science grades were not good enough to get them into medical school. If that is true of life science, then it may be doubly true of computer science. Ask the average lawyer about artificial intelligence and machine learning, and you will likely elicit some combination of fascination and bewilderment. However, we can no longer afford the latter. Having penetrated every other aspect of our lives, it was inevitable that computer algorithms would eventually make their way into the criminal justice system. That reality has arrived. If criminal defence lawyers are to continue their job of protecting our liberty, they need to be prepared to live in this brave new world.

This paper will provide a general overview of three ways in which algorithms can be used in the criminal justice system: (1) to direct policing; (2) to assess the risk of an individual; and (3) to establish the identity of the perpetrator. The examples given are not exhaustive. Instead, the goal is a modest one: to use a few examples to survey the relevance and reliability of computer algorithms in the criminal justice process, from the beginning of a police investigation to the end of a trial.

II. Uses of Algorithms

A. Using Algorithms to Direct Policing

The first way that algorithms can be used in the criminal justice system is to assist police in deciding how to allocate its resources. Like every aspect of the justice system, policing is subject to budgetary constraints. Each police force only has so many officers, and they can only be in so many places at one time. Thus, a number police forces have begun to use predictive policing systems to direct their resources to what are considered riskier neighbourhoods and communities.[2]

---

[1] Partner, Stockwoods LLP in Toronto. The author is hugely indebted to Richard Mahal (Harold G. Fox scholar) for his invaluable assistance with the research for and preparation of this paper. The author also wishes to thank Uttra Gautam (recent University of Ottawa law graduate) for her research assistance.

[2] For a general overview of the subject, see Walter Perry, Brian McInnis, Carter Price, Susan Smith, John Hollywood, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (2013, RAND

In the United States, police forces use algorithms that rely on variables like records of past offences, weather, and social media to create predictive maps showing crime hotspots.[3] Canadian police forces have started down this path as well (they term it "forecasting").[4] In 2015, Saskatoon Police created a "Predictive Analytics Lab".[5] Vancouver Police has also been reported to be employing predictive analytics to guide their law enforcement efforts.[6]

So how do predictive policing algorithms work? One approach is to "train" computers with large sets of historical data to identify variables and patterns in the data that correlate to particular outcomes.[7] The algorithm's predictive accuracy is then tested against new sets of data.[8] When the algorithm is deemed to be sufficiently accurate in its prediction, it is used in live cases.[9] The algorithm can continue to evolve with the addition of new data from the live cases, either through interventions by a human programmer or through machine learning approaches that allow the algorithmic system to automatically adapt its equations to incorporate new data.[10]

PredPol is a predictive policing system used across the United States. It generates predicted crime hotspots for the following day using three data points drawn from police data sets: crime type, crime location, and date/time of the criminal activity.[11] It employs a self-learning algorithm based on seismic activity models, and is one of the few systems that has made its algorithm accessible to researchers.[12]

Corporation) available at <https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf>.

[3] See Andrew G. Ferguson, "Crime Mapping and the Fourth Amendment: Redrawing High-Crime Areas" (2011) 63 Hastings Law Journal 179, at p. 187.

[4] Robert Muggah, "Does Predictive Policing Work?" (*Canadian Global Affairs Institute*, December 4, 2016) available at <https://www.cgai.ca/opeddecember42016>.

[5] Meaghan Craig, "Saskatoon police lead the country with Predictive Analytics Lab" (*Global News*, January 15, 2016) available at <https://globalnews.ca/news/2455063/saskatoon-police-lead-the-country-with-predictive-analytics-lab/>.

[6] Matt Meuse, "Vancouver police now using machine learning to prevent property crime" (*CBC News*, July 22, 2017) available at <https://www.cbc.ca/news/canada/british-columbia/vancouver-predictive-policing-1.4217111>.

[7] Carmen Cheung, "Making Sense of the Black Box: Algorithms and Accountability" (2017) 64 Criminal Law Quarterly 540, at p. 2.

[8] *Ibid.*

[9] *Ibid.*

[10] *Ibid.*

[11] PredPol, "How PredPol Works", available at <http://www.predpol.com/how-predpol-works/>.

[12] Carmen Cheung, "Making Sense of the Black Box: Algorithms and Accountability" (2017) 64 Criminal Law Quarterly 540, at p. 2.

The Human Rights Data Analysis Group recently conducted a study on PredPol's algorithm using police-recorded drug crime data from Oakland, California.[13] Not surprisingly, the study found that when police-recorded data are used to train an algorithmic system, the algorithms will detect and reproduce the patterns in the data.[14]

This can have the dangerous effect of perpetuating bias in policing. Police data sets reflect the choices that police have made in the past in terms of which communities and individuals should be the subject of law enforcement attention.[15] When this data represents the key input for predictive policing algorithms, it can exaggerate the criminality risk in certain neighbourhoods and groups that have historically been targeted by police.[16] Police will then be increasingly likely to patrol or monitor these same neighbourhoods and groups, and observe new criminal acts confirming their prior assumptions about criminality.[17] These newly-observed criminal acts will then feed into the predictive policing algorithm to create a feedback loop: the model becomes increasingly confident of its predictions, compounding the initial biases in the police data sets.[18] This creates confirmation bias at two levels: at the level of individual police officer decision-making, as well as at the level of the algorithm itself.

In August 2016, 17 civil rights organizations in the United States released a joint statement on the civil rights concerns of predictive policing. They emphasized the possibility of racially biased outcomes, as well as the lack of transparency, public debate, and attention to community needs.[19] The Los Angeles Police Department (LAPD), in particular, stopped using many of its predictive policing programs because of inherent racial biases targeting black and Latino neighbourhoods. Local community groups challenged the program LASER (Los Angeles' Strategic Extraction and Restoration) in court, but LAPD discontinued the program before trial (which had been scheduled for August 2019).[20]

---

[13] Kristian Lum and William Isaac, "To predict and serve?" (*Significance*, October 2016) available at <https://rss.onlinelibrary.wiley.com/doi/epdf/10.1111/j.1740-9713.2016.00960.x>.
[14] *Ibid.,* at p. 15.
[15] Andrew Selbst, "Disparate Impact in Big Data Policing" (2017) 52 1 Georgia Law Review 109, at pp. 119-123 (noting that many of the criticisms that apply to risk assessment algorithmic tools apply equally to predictive policing tools).
[16] *Ibid.*, at pp. 18-19.
[17] Carmen Cheung, "Making Sense of the Black Box: Algorithms and Accountability" (2017) 64 Criminal Law Quarterly 540, at p. 3.
[18] *Ibid.*
[19] "Statement of Concern About Predictive Policing by ACLU and 16 Civil Rights Privacy, Racial Justice, and Technology Organizations", available here <https://www.aclu.org/other/statement-concern-about-predictive-policing-aclu-and-16-civil-rights-privacy-racial-justice>.
[20] "LAPD ends another data-driven crime program touted to target violent offenders", *Los Angeles Times,* April 12, 2019, available at:

Defence counsel need to be mindful of these concerns, and should prepare to challenge the use of predictive policing where it entrenches on *Charter* rights. As the Supreme Court of Canada held in *R. v. Le,* "research now shows disproportionate policing of racialized and low-income communities".[21] Where a predictive policing algorithm has, by virtue of its design, the effect of aggravating the over-policing of racialized and low-income communities, counsel may wish to challenge its use as a violation of the right to equality under s. 15 of the *Charter*.

In addition, where police are approaching individuals in such communities, individuals are more likely to be "detained" within the meaning of s. 9 of the *Charter*. This is because, as the Supreme Court held in *Le,* a reasonable person in the individual's shoes will be aware of the history of over-policing and that will affect his perception of whether he can simply walk away from police.[22] He will be more likely to treat police questions as demands that must be obeyed rather than as requests that may be ignored.[23] Therefore, defence counsel should carefully scrutinize the reasons given by police for any such encounter/detention. The high crime nature of a neighbourhood cannot itself be a basis for detaining individuals.[24] And as the Ontario Court of Appeal recently clarified in *R. v. Dudhi,* race or racial stereotypes cannot play *any* role in the decision to detain, regardless of whether there exist other objective grounds for police action.[25]

B. Using Algorithms to Assess Risk

The second way in which algorithms can be used in criminal justice is to help the courts assess risk. This is most relevant in bail and sentencing hearings.

Predictive algorithms are commonly used as risk assessment tools for bail and sentencing determinations in the United States.[26] COMPAS (Correction Offender Management Profiling for Alternative Sanctions) is perhaps the most widely used software system to assess the risk of recidivism.[27] COMPAS generates risk scores from answers given to over 100 questions, which may be

---

<https://www.latimes.com/local/lanow/la-me-laser-lapd-crime-data-program-20190412-story.html>.

[21] *R. v. Le,* 2019 SCC 34 at para. 95.

[22] *Ibid.*

[23] *R. v. Grant* (2006), 81 O.R. (3d) 1 (C.A.), at para. 24.

[24] *R. v. Mann,* 2004 SCC 52 at para. 47.

[25] *R. v. Dudhi,* 2019 ONCA 665 paras. 62-63.

[26] Angèle Christin, Alex Rosenblat, and Danah Boyd, "Courts and Predictive Algorithms" (*Data & Civil Rights: A New Era of Policing and Justice*, October 27, 2015), at pp. 2-3, available at <https://www.law.nyu.edu/sites/default/files/upload_documents/Angele%20Christin.pdf>.

[27] Carmen Cheung, "Making Sense of the Black Box: Algorithms and Accountability" (2017) 64 Criminal Law Quarterly 540, at p. 3.

provided either by the defendant or derived from criminal records.[28] COMPAS will compare the defendant's information with group data comprised of the recidivism rates collected from multiple sample populations of released offenders for a specific period of time.[29] The risk score reflects the likelihood that those with a similar history are either more or less likely to commit another crime following release from custody.[30]

ProPublica, an investigative non-profit news organization in the United States, has studied the efficacy of and biases inherent in COMPAS. The results are discouraging. The rate of false positives for high risk of recidivism was nearly double for black defendants (45%) than for white defendants (23%).[31] At the same time, the rate of false negatives for low risk of recidivism was nearly double for white defendants (48%) than for black defendants (28%).[32]

ProPublica also found that the algorithmic system's predictive value varied depending on the type of future crime at issue.[33] When the recidivism risk categories predicted by COMPAS were compared to the actual recidivism rates of defendants, COMPAS had a 61% accuracy with respect to overall recidivism within a two-year period, but only a 20% accuracy with respect to violent recidivism.[34] When a full range of crimes were considered — including minor offences like driving with an expired licence — the algorithm was just slightly more accurate than a coin toss.[35]

Some of the miscalculations of risk are attributable to incomplete or inaccurate inputs, such as missing prison records.[36] However, some of the miscalculations are attributable to the algorithm itself, including the way in which variables like socioeconomic status, employment status, and family background are weighed in the calculation of risk scores.[37]

---

[28] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine Bias", (*ProPublica* May 23, 2016), available at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
[29] Sara M. Smyth, "Can We Trust Artificial Intelligence in Criminal Law Enforcement?" (2019) 17 Canadian Journal of Law and Technology 99, at p. 105.
[30] *Ibid.*
[31] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine Bias" (*ProPublica* May 23, 2016), available at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
[32] *Ibid.*
[33] *Ibid.*
[34] *Ibid.*
[35] Sara M. Smyth, "Can We Trust Artificial Intelligence in Criminal Law Enforcement?" (2019) 17 Canadian Journal of Law and Technology 99, at p. 108.
[36] Carmen Cheung, "Making Sense of the Black Box: Algorithms and Accountability" (2017) 64 Criminal Law Quarterly 540, at p. 3.
[37] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine Bias", (*ProPublica* May 23, 2016), available at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

The choice of variables may also have constitutional dimensions since algorithmic systems may rely on variables such as race, gender, or religion.[38] And even in systems that exclude such sensitive variables, proxy variables (such as postal code, income, or educational background) may be used.[39] These variables are considered proxies because they correlate strongly with race and reflect racial bias.[40] Depending on the function of these variables in the algorithmic system, equality guarantees may be implicated.[41]

These concerns are compounded by a lack of transparency and accountability. Sometimes, the proprietors of the software will resist disclosure of the underlying algorithms on the basis of trade secret. But even if the algorithm is disclosed, that may not present the entire picture. Algorithms that use machine-learning techniques evolve as new data is fed into the system. This means that even the computer programmers who designed the algorithm may not understand the specific procedures by which a given result is generated.[42] It is hard enough to cross-examine a computer science expert who has designed a complex algorithm. But it is impossible to cross-examine a computer that has evolved the algorithm on its own through the process of machine-learning.

Wisconsin v. Loomis[43] is an example of how these concerns can affect due process. In that case, the trial court relied on a COMPAS risk score during sentencing. The defendant claimed a violation of his right to due process because he could not obtain disclosure of the underlying computer code and therefore could not challenge the scientific validity of the COMPAS risk assessment.[44] The Wisconsin Supreme Court rejected this argument on the unconvincing basis that the COMPAS score was only one of several factors considered by the trial court at sentencing. Nevertheless, the Court outlined a series of concerns about the COMPAS algorithm, including those set out in the ProPublica report:

---

[38] Carmen Cheung, "Making Sense of the Black Box: Algorithms and Accountability" (2017) 64 Criminal Law Quarterly 540, at p. 3.
[39] Angèle Christin, Alex Rosenblat, and Danah Boyd, "Courts and Predictive Algorithms", *Data & Civil Rights: A New Era of Policing and Justice*, at p. 5.
[40] *Ibid.*
[41] Carmen Cheung, "Making Sense of the Black Box: Algorithms and Accountability" (2017) 64 Criminal Law Quarterly 540, at p. 3.
[42] Angèle Christin, "Predictive Algorithms and Criminal Sentencing", at p. 283 in Daniel Bessner and Nicolas Guilhot, *The Decisionist Imagination: Sovereignty, Social Science and Democracy in the 20th Century*, 1st ed. (2018, Hart Publishing).
[43] 2016 WI 68, Case No. 2015AP157-CR (Wis. S.C., 2016). The Supreme Court of the United States declined to hear Loomis' appeal in 2017.
[44] Carmen Cheung, "Making Sense of the Black Box: Algorithms and Accountability" (2017) 64 Criminal Law Quarterly 540, at FN 22.

> … [a] recent analysis of COMPAS's recidivism scores based upon data from 10,000 criminal defendants in Broward County, Florida, concluded that black defendants 'were far more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism.' Likewise, white defendants were more likely than black defendants to be incorrectly flagged as low risk … this study and others raise concerns regarding how a COMPAS assessment's risk factors correlate with race.[45]

The Court went on to mandate that pre-sentence reports incorporating a COMPAS assessment be accompanied by five warnings for judges: (1) the "proprietary nature of COMPAS" prevents the disclosure of how risk scores are calculated; (2) COMPAS scores are unable to identify specific high-risk individuals because these scores rely on group data; (3) although COMPAS relies on a national data sample, there has been "no cross-validation study for a Wisconsin population"; (4) studies "have raised questions about whether [COMPAS scores] disproportionately classify minority offenders as having a higher risk of recidivism"; and (5) COMPAS was developed specifically to assist the Department of Corrections in making post-sentencing determinations.[46] In issuing these warnings, the Court made clear its desire to cast doubt on the tool's accuracy and reliability. Nonetheless, it appears that other states are far too quickly accepting the accuracy of these proprietary algorithms without first validating their accuracy.[47]

Research shows that it is psychologically difficult and rare for prosecutors and judges to "override" the recommendations provided by an algorithm. It is difficult even for highly trained professionals to challenge numbers and equations if they have no formal training in statistics.[48] Thus, when the algorithm provides a "high" estimate of risk, the judicial tendency may be to incarcerate regardless of countervailing factors.[49]

To the extent that judges override the algorithmic information, studies have shown that they err on the side of detaining rather than releasing individuals. A 2006 report on juvenile justice in California showed that "detain overrides" (*e.g.*, a judge's decision to incarcerate a defendant when the algorithm

---

[45] *Wisconsin v Loomis,* 2016 WI 68, Case No. 2015AP157-CR (Wis. S.C., 2016), at para. 63.
[46] *Wisconsin v Loomis,* 2016 WI 68, Case No. 2015AP157-CR (Wis. S.C., 2016), at para. 100.
[47] Alyssa Carlson, "The Need for Transparency in the Age of Predictive Sentencing Algorithms" (2017) 103 1 Iowa Law Review 303, at pp. 323, 329.
[48] Angèle Christin, Alex Rosenblat, and Danah Boyd, "Courts and Predictive Algorithms", *Data & Civil Rights: A New Era of Policing and Justice*, at p. 7.
[49] *Ibid.*

provides a low risk estimate) are much more frequent than "release overrides" (*e.g.*, the decision to release a defendant when the algorithm provides a high risk estimate).[50]

In addition, judges may adjust their approach to sentencing over time in order to more closely follow the predictions being generated by algorithmic risk-assessment tools.[51] This cognitive bias, known as "anchoring", involves people drawing on the very first piece of evidence at their disposal, however weak, when making subsequent decisions.[52] Consequently, judges may be prompted by the recommendation of an algorithm to increase the sentences they are inclined to impose without realizing that they are following the algorithm and being influenced by their cognitive bias in favour of lengthier periods of incarceration.[53]

Also concerning is the theory of justice embedded in the algorithms.[54] Section 718 of the *Criminal Code* sets out five sentencing objectives: denunciation, deterrence, separation of offenders from society where necessary, rehabilitation, reparations for harm done to the community, and promoting a sense of responsibility in offenders. As currently designed, algorithms privilege a view of justice based on estimating the "risk" posed by the offender when deciding on a sentence with the goal of incapacitating dangerous individuals. This can lead to over-emphasis on the objective of separation at the expense of other equally important (if not more important) sentencing principles.[55]

For all of these reasons, defence counsel should be careful to scrutinize any algorithmic evidence offered by the Crown at the bail or sentencing stage. Defence counsel should seek disclosure of the specific algorithmic tool used, including the underlying computer code. Where the tool is widely known and has been criticized in studies (*e.g.,* COMPAS), defence counsel may consider objecting to admissibility on the basis that the evidence is not credible and trustworthy.

---

[50] "Juvenile Detention Risk Assessment: A Practical Guide to Juvenile Detention Reform", A Project of the Annie E. Casey Foundation, available at <https://www.aecf.org/resources/a-practice-guide-to-juvenile-detention-reform-1/>. See also Tom Simonite, "Algorithms should've made courts fair. What went wrong?" *Wired* (September 5, 2019), online: <www.wired.com>.

[51] Angèle Christin, Alex Rosenblat, and Danah Boyd, "Courts and Predictive Algorithms", *Data & Civil Rights: A New Era of Policing and Justice*, at p. 7.

[52] Amos Tversky and Daniel Kahneman, "Judgment Under Uncertainty: Heuristics and Biases" (1974) Science 185 (4157), at pp. 1124–1131.

[53] Angèle Christin, Alex Rosenblat, and Danah Boyd, "Courts and Predictive Algorithms", *Data & Civil Rights: A New Era of Policing and Justice*, at p. 7.

[54] *Ibid.*

[55] Harcourt, Bernard E, *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age,* 1st ed. (2007, University of Chicago Press), at p. 39.

Guidance can be drawn from *Ewert v. Canada*.[56] While *Ewert* is a correctional placement case rather than bail or sentencing case, it is nonetheless instructive for counsel seeking to challenge the admissibility of risk assessment tools. Ewert, who identified as Métis, challenged the use of five psychological and actuarial risk assessment tools used by the Correctional Service of Canada (CSC) to assess an offender's psychopathy and risk of recidivism. He argued that they were developed and tested on predominantly non-Indigenous populations and that no research confirmed that they were valid when applied to Indigenous persons. The trial judge agreed and found that the CSC's use of these tools — in the face of long-standing concerns about their applicability to Indigenous offenders — represented a violation of s. 24(1) of the *Corrections and Conditional Release Act* (which requires the CSC to "take all reasonable steps to ensure that any information about an offender that it uses is as accurate, up to date and complete as possible"). Similar reasoning can be applied to exclude questionable risk assessment tools at the bail or sentencing stage of a criminal proceeding.

Defence counsel may also find it helpful to compare the risk assessment tool in question to New Jersey's Public Safety Assessment (PSA) tool. The New Jersey PSA attempts to avoid the problem of structural bias against certain demographic groups by eliminating risk factors that might act as proxies for race or gender.[57] The algorithm considers just nine risk factors:

(i) the person's age at the current arrest;
(ii) whether the current offence is violent;
(iii) pending charges at the time of the offence;
(iv) prior misdemeanor convictions;
(v) prior felony convictions;
(vi) whether those prior convictions were for violent crimes;
(vii) prior failure to appear in the past two years;
(viii) prior failure to appear instances that are older than two years; and
(ix) prior incarceration sentences.[58]

Importantly, the New Jersey PSA does not take into account factors like education, income, or employment which might serve as proxies for race.[59] A spokesperson for the developer of the PSA

---

[56] *Ewert v. Canada*, 2018 SCC 30.
[57] Issie Lapowsky, "One State's Bail Reform Exposes the Promise and Pitfalls of Tech-Driven Justice" (*Wired*, September 5, 2017) available at <https://www.wired.com/story/bail-reform-tech-justice/>.
[58] *Ibid.*
[59] *Ibid.*

also cautions against over-reliance on the scores generated by the algorithm, encouraging them to be used merely as a baseline for further reasoning rather than in isolation.[60]

C. Using Algorithms to Establish Identity

The third, and perhaps most consequential, way in which algorithmic evidence can be used in criminal proceedings is to establish the identity of the perpetrator. Two types of algorithmic evidence are increasing in prominence in this regard: (1) probabilistic genotyping of DNA samples; and (2) facial recognition technology. A proper discussion of the former is beyond the scope of this paper. This paper will focus on the latter.

Facial recognition technology involves a computer program analyzing one's unique facial structure, such as the distance between the nose and lips, and mapping those key features onto an existing image — or, commonly, against a database of existing images — for comparison.[61] The technology can also be used in "real-time."[62] For example, some police departments can scan the faces of passersby using a surveillance camera.[63]

Studies have raised concerns about the accuracy of some of the major facial recognition systems in use. For example, Amazon's facial recognition software misidentified 28 members of Congress and matched them with criminal mugshots.[64] Computer scientists, Joy Buolamwini (MIT Media Lab) and Timnit Gebru (the technical co-lead of Google's Ethical Artificial Intelligence Team), have shown that facial recognition systems have a harder time differentiating between men and women of darker skin tone.[65] About 130 million US adults are already in facial recognition databases, but the original

---

[60] *Ibid.*
[61] "Why AI Is A Growing Part Of The Criminal Justice System" (*Science Friday*) available at <https://www.sciencefriday.com/segments/artificial-intelligence-is-a-growing-part-of-the-criminal-justice-system-should-we-be-worried/>.
[62] *Ibid.*
[63] Stephen Gaines, "The Perpetual Line-Up: Unregulated Police Face Recognition in America" (Georgetown Law Center on Privacy & Technology, October 18, 2016) available at <https://www.perpetuallineup.org/>.
[64] Jacob Snow, "Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots" (*American Civil Liberties Union*, July 26, 2018) available at <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>.
[65] Joy Buolamwini, Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification", Proceedings of Machine Learning Research 81: 1-15, 2018, available at: <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>.

datasets are mostly white and male. This creates a bias against those with darker skin tones, giving rise to significant error rates.[66]

Because of these concerns, San Francisco recently banned the use of facial recognition by transport and law enforcement agencies.[67] Other cities in the US and other countries around the world, however, are testing the technology in trials. In the UK, for example, police forces in South Wales, London, Manchester, and Leicester have been testing the technology[68] in the face of concerns raised by civil liberties organizations about the number of false matches the systems create.[69] In South Wales, the technology had a 90% error rate.[70]

Lynch v State (Florida)[71] is a troubling example of facial recognition programs being used to identify the perpetrator of a crime despite concerns about its reliability. Lynch was accused of selling 50 dollars' worth of crack cocaine to undercover officers. Unlike many undercover drug operations, however, Lynch was not arrested at the scene of the transaction. Instead, police surreptitiously took a cell phone picture of the suspect, which they forwarded to a crime analyst. The crime analyst uploaded the cell phone picture and ran it through a facial recognition program called FACES (Face Analysis Comparison Examination System). FACES draws from a database of more than 33 million driver's license and law enforcement photos. The crime analyst's search resulted in four possible suspects, including Lynch: Lynch was given one star for the quality of the match, while the other three suspects received no star. The reliability of FACES' star system was unknown.

At Lynch's trial, he was convicted and sentenced to 8 years in prison. The prosecution did not, however, disclose the photographs of the three other individuals who matched the suspect's cell phone pictures. On appeal, Lynch argued that this failure of disclosure warranted a new trial. The Florida

---

[66] Matthew Wall, "Biased and wrong? Facial recognition tech in the dock" (*BBC News*, July 8, 2019) available at <https://www.bbc.com/news/business-48842750>.
[67] Dave Lee, "San Francisco is first US city to ban facial recognition" (*BBC News*, May 15, 2019) available at <https://www.bbc.com/news/technology-48276660>.
[68] Matthew Wall, "Biased and wrong? Facial recognition tech in the dock" (*BBC News*, July 8, 2019) available at <https://www.bbc.com/news/business-48842750>.
[69] Chris Fox, "Face recognition police tools 'staggeringly inaccurate'" (*BBC News*, May 15, 2018) available at <https://www.bbc.com/news/technology-44089161>. See also Ian Sample, "Facial recognition tech is arsenic in the water of democracy, says Liberty" (*The Guardian*, June 7, 2019) available at <https://www.theguardian.com/technology/2019/jun/07/facial-recognition-technology-liberty-says-england-wales-police-use-should-be-banned>.
[70] David Davis, "Facial recognition technology threatens to end all individual privacy" (*The Guardian*, September 20, 2019) available at <https://www.theguardian.com/commentisfree/2019/sep/20/facial-recognition-technology-privacy>.
[71] *Lynch v State,* No. 1d16-3290 (Fla Dist Ct App, Dec 27, 2018).

First District Court of Appeals rejected this argument and affirmed the conviction on the basis that Lynch could not show that the other photos would have supported his misidentification defence. Lynch appealed again, but the Florida Supreme Court denied discretionary review.

One can only hope that a similar result is not reached in Canadian courts. Defence counsel should be on high alert whenever police use facial recognition algorithms to identify the perpetrator of a crime. At minimum, the defence should request disclosure of the algorithm, the underlying computer code, and all matches made by the program. Other matches may be exculpatory.

Should the Crown seek to adduce evidence of the algorithmic matches in court, defence counsel may need to launch a full scale attack on the reliability (and therefore admissibility) of the algorithm. This will require the calling of expert evidence to examine the merits of facial recognition technology in general as well as the specific algorithmic tool at issue. The latter is only possible where the defence has access to the underlying computer code.

In the event that the Crown resists disclosure on proprietary grounds, the defence should object to the admissibility of the evidence on the basis that its reliability cannot be properly tested. In the appropriate case, where exclusion alone will not remedy the injustice, the defence should also consider asking for a stay of the proceedings. The proprietor of an algorithm may have a legitimate claim of confidentiality over the underlying computer code. This may justify a confidentiality order in respect of this evidence to prevent it from being disclosed to any non-parties or the general public.[72] But the proprietor's claim of confidentiality cannot justify an unfair trial by allowing the Crown to benefit from the inculpatory value of the algorithm without disclosing it to the defence so that its exculpatory value can be extracted. If the Crown is prevented from disclosing relevant information to the defence, it is up to the court to ensure the fairness of the trial. Depending on the circumstances, this could mean excluding the resulting evidence or even staying the proceedings.[73]

III. Conclusion

Innovation often comes with unintended consequences. When innovation happens in the criminal justice system, those consequences can be among the most severe: the unfair deprivation of liberty.

[72] *Sierra Club of Canada v. Canada (Minister of Finance),* 2002 SCC 41.
[73] *R. v. Ahmad,* 2011 SCC 6.

Rightly or wrongly, the frontline responsibility for guarding against this consequence often falls on defence counsel.

In order to properly discharge their duties, defence counsel need to be mindful of the various ways in which computer algorithms can be used in the criminal process. This paper has highlighted three methods of use with examples of each, but this is by no means an exhaustive list. As the technology continues to develop, new forms of use — including those that may be currently unimaginable — will spring forth. Defence counsel need to be ready to understand, use, and/or object to these algorithms as they arrive.

In rising to this challenge, defence counsel should return to what they know best: the first principles of criminal, evidence, and constitutional law. Disclosure. Other *Charter* protections like the right to equality and the right not to be arbitrarily detained. The rules of admissibility for evidence of questionable reliability, especially expert evidence. These are the tools that criminal defence lawyers have used to defend their clients for decades. As increasingly complicated algorithms supplement or replace different parts of the criminal process, these tools will remain just as if not more relevant than ever. How we make use of these tools is up to us.