November 18, 2020

Harnessing the Power of AI Technology; A Commentary on the Law Commission of Ontario Report on AI and the Criminal Justice System

By: Lisa Silver and Gideon Christian

Report Commented On: Law Commission of Ontario, *The Rise and Fall of AI and Algorithms In American Criminal Justice: Lessons for Canada*, (Toronto: October 2020).

The Law Commission of Ontario (LCO) recently released its Report on the use of artificial intelligence (AI) and algorithms in the Canadian criminal justice system. The Report, which is the first of three papers on the issue, is one of the most comprehensive discussions of the use of AI and algorithmic technologies in the criminal justice system to date. In Canada, AI use in the criminal justice system is limited and not easily subject to in-depth review. In the United States, however, AI and algorithms are used extensively throughout the justice system, particularly in pre-trial release decision-making. Not surprisingly, then, the Report draws from this American experience to arrive at a number of recommendations for application to the Canadian context. Based on those lessons learned, the LCO Report warns of "the risk of adopting unproven and under-evaluated technologies too quickly to address long-standing, complex and structural problems in the justice system" (at 7). Yet, in the midst of this cautionary tone, the Report also recognizes that AI use in the criminal justice system will likely increase in the future. The Report proactively outlines a framework for such use by urging AI regulation, the application of legal protections to AI, and community involvement in developing AI best practices. All of these warnings and recommendations are extremely useful but the Report begs the basic question of whether the justice system should be using machine intelligence, with its embedded biases, in matters that can profoundly change people's lives. Ultimately, the Report should stand as a timely reminder of the unharnessed power of technology and the realistic potential for injustice when it is used without restraint.

Defining AI

Artificial intelligence is a term used to refer to broad range of technological methods and tools that "learn" from the performance of tasks, thus exhibiting intelligence similar to cognitive intelligence. These include machine learning, facial recognition, and natural language processing technologies. The Report adopts a definition that extends the concept of AI to include dominant social practices of individuals who design the technological system, and the industrial power that runs the system. Algorithm is defined as the mathematical logic that enables the system to perform tasks or make decisions (at 8).

The Extent of AI Use in Canada

AI and algorithms have been used to automate decision-making process in public and private settings. The Report notes the increasing use of AI in government decision-making in the US.

Unfortunately, there is no clear indication as to the extent of AI use by the government in Canada. The only information on AI use in Canada is contextual and case specific, arising from disclosure in litigation, access to information requests, or investigative news reports. For example, Canadians were unaware that Clearview AI facial recognition technology was being used by various police departments in Canada, including the RCMP, until it was made public by the Canadian press.

In the US criminal justice system, AI and algorithms have been used to automate decision-making in the context of risk assessment in bail, sentencing, inmate housing classification, and parole. These tools have been used in bail hearings to predict whether the accused would reoffend if released from custody pending trial; in sentencing, they have been used to recommend the appropriate sentence for a convict taking into consideration whether the convict has a low or high risk of re-offending; in the context of inmate housing classification, they have been used to recommend security classification of inmate e.g. high, medium, or low; and in the parole context, they have been used to determine whether an inmate should be released prior to the completion of their criminal sentence.

Report Themes

The Report identifies eleven themes in its review of AI use in the criminal justice system (at 3–4). The first theme sets the tone for the entire Report by describing the primary challenges with AI's use as involving a "significant new frontier in human rights, due process and access to justice" (at 3). This theme focuses on the human issues engaged by this "new frontier" such as "equality, bias, access to justice and due process ... affecting fundamental rights." This first theme should be the overarching one and should receive primary attention, considering that *Charter* rights and values lay at the core of these fundamental rights.

The next ten themes identify the myriad of other concerns raised by AI and algorithms such as providing "simple solutions" for "complex problems" (at 3). Although predictive analytics appear "objective" and "evidence-based", this veneer of neutrality may in fact hide the risks of using "unproven" technologies or even proven technologies that are based upon flawed and biased data. Compounding this concern is the use of the statistical data or AI output, which is subject to manipulation, by those making policy choices. This reveals the dark side of human interaction with machine intelligence; AI systems are merely data points requiring human interpretation.

Another theme raises the continual problem of legal protections lagging behind the technology (at 3). This lag, the Report urges, must be acknowledged and gaps must be closed before AI is used or introduced – not after. Although the Report labels this concern as matters of "due process", legal protections involving evidentiary rules and *Charter* protections are more than due process. Legal rights protect individuals in our system from the power of the state and are fundamental to our human dignity and self-autonomy. Applying constitutional and *Charter* protections as the oversight tool to ensure equality, fairness and justice in AI use is required but what is really at risk here is the violation of an individual's human dignity, personal autonomy and self-worth. AI can be used to dehumanize and this must be acknowledged at the outset.

The last 4 themes identify other legitimate concerns with AI use while also framing it as an opportunity that the legal system should prepare for by developing best practices. This contention deserves a pause. It is not necessarily correct to assume that because we have technology and because it presents opportunities that we should use AI in the justice system. Even the call for "broad participation in the design, development and deployment" of AI assumes that society needs AI or should use AI to restrict a person's liberty interests. Even "thoughtful, deliberate and incremental" reforms to the law to ensure protections are in place cannot take the place of thoughtful and deliberate conversations as to whether we should legally, morally and ethically use AI to determine if a person enters custody or leaves it.

Lessons Learned and Identified Issues

Traditionally, bail decisions are often made in summary proceedings by human judges. The Report noted that this process could be flawed as a result of reliance on intuition and personal preferences by human judges. While it might seem that reliance on AI to make these decisions could help overcome the inconsistency, human bias, and sometimes outright prejudice that can arise from intuitive human judgement, the Report is very critical of this automated process. The Report notes that these AI tools are trained on data that reflects structural racism and institutional inequity evident in our court system and law enforcement. The ability of these AI technologies to perpetuate the bias and inequity prevalent in the real world raises some serious concerns. In the light of this problem, the Report highlights some important issues that must be considered or addressed before the use of AI is embraced by the Canadian criminal justice system. Below, we highlight some of the issues identified.

Report Issue # 1: Bias In, Bias Out

Algorithmic tools are being used in criminal justice risk assessment to determine the likelihood that an offender in the criminal justice system will reoffend. This is an important determination, as it serves to balance public safety concern with the *Charter* rights of the offender. However, the Report notes that algorithmic risk assessment tools could also function as "a sophisticated form of racial profiling" (at 7 and see "Not In It For Justice", *Human Rights Watch* (11 April 2017)). David Robinson and Logan Keopke, in their article on *Civil Rights and Pretrial Risk Assessment Instruments,* raised similar concerns (see *Civil Rights and Risk Assessments* at 4 and the LCO Report at 20–21). In their view, this technology has inherent legitimacy issues as "the world of mass incarceration and racially inequitable criminal law" provide the data for the risk assessment. This is evident from the fact that the training data or inputs used to train the algorithmic tools are the product of racially disparate practices. Thus, the biased data fed into the system will inevitably result in biased output from the algorithmic system.

Research confirms this bias. In Broward County, Florida, for example, the risk assessment AI tool for sentencing "proved remarkably unreliable" and "likely to falsely flag black defendants as future criminals" (at 12). The American Pretrial Justice Institute (PJI) recently remarked there is "no pretrial justice without racial justice" (at 13). The PJI has taken the position that AI risk assessment tools "can no longer be a part of our solution for building equitable pretrial justice systems" (at 13). Another important point noted in the Report is the fact that although race is not included as an explicit variable in these algorithms, this does not imply that the tool is raceneutral. Factors that may correlate heavily with an individual's race as well as factors that disparately impact on race such as arrests and criminal records are not excluded from the algorithm. (at 21). This may result in some form of race-based discrimination arising from the assessments made by such tools. This possibility should give rise to serious concern in the Canadian criminal justice system, which is characterised by disproportionate representation of Black and Indigenous peoples.

Thus, it appears that while the use of algorithmic risk assessment tools may result in consistent, "evidence-based" and efficient predictions, the Report noted that they could potentially result in data discrimination and a basis for section 7 and 15 *Charter* challenges. This is an important legal issue that will need to be comprehensively addressed before developing or implementing the use of AI and algorithmic tools in the Canadian criminal justice system. (at 22).

Report Issue #2: Data Transparency

The lack of transparency relating to how AI and algorithmic tools work is another significant issue identified by the Report. This is related to the "black box" concept associated with these tools. The lack of transparency here arises in three ways: the data used in the analysis by these tools, including data used to train the system; the weight attached to the data by the algorithm; and whether specific factors or combination of factors used in the analysis are proxies for problematic variables, e.g. race and poverty (at 23).

There are many problems arising from the lack of transparency evident in these tools. First, it is difficult to test the tools for accuracy and bias, and second, it is difficult to legally challenge the use of these tools in the criminal justice context because an offender bears the burden of proof, which will be more difficult to discharge as a result of lack of transparency surrounding their operation. The Report argues that any introduction of these tools in the Canadian criminal justice system will inevitably raise important questions surrounding data transparency and accountability (at 24).

Another important issue relating to transparency is the proprietary nature of these AI tools. If the AI tools used in the criminal justice system are developed by private corporations, there is the tendency to licence the tools for use to the relevant government departments. This license to use does not entitle the relevant department to any right of access to proprietary trade secrets. The implication is that the accused/offender, the prosecution, and the court have no ability to review how the tools work. This problem was evident in the 2016 Wisconsin state court decision in *State v Loomis,* 2016 WI 68. The offender challenged the use of COMPAS, an algorithmic risk assessment tool, in his criminal sentencing decision. In imposing the sentence, the judge relied on the COMPAS algorithmic risk assessment tool that suggested that the offender had an extremely high risk of reoffending. In challenging the length of the criminal sentence, the offender sought access to proprietary information in the COMPAS software. The software developer refused, which refusal was surprisingly upheld.

The *Loomis* decision is unique and has not been followed by any other court. Until similar case emerges again in any US jurisdiction, it would be difficult to predict whether other courts will follow the *Loomis* decision. In Canada, the position appears to be different, bearing in mind the

Supreme Court of Canada's decision in *May v Ferndale Institution*, 2005 SCC 82 (CanLII), confirming the right of an offender to access proprietary software information relevant to their *Charter* right challenge. Hence, in considering the adoption and implementation of AI-powered tools in the Canadian criminal justice system, it is important to note that offenders may have a *Charter* right to access proprietary information where it is relevant to their defence in the criminal proceeding. This would require that AI tools used in the Canadian criminal justice system be developed either by the relevant government department, by a private corporation that would be willing to grant access to their proprietary information, or developed on open source software.

The right of access to this information is very important to an accused person. The Report notes that "a criminal accused confronting an algorithmic risk assessment faces even more difficulty in presenting a full answer and defence to the charges against them" (at 37). Depriving them access to proprietary information relevant to their criminal case will only present additional hurdles that "may actually compound the over-representation of low-income and racialized communities already present in the criminal justice system" (at 37).

Report Issue # 4: Data Accuracy, Reliability and Validity

This also goes to the legitimacy of AI and algorithmic tools. Development and implementation of these tools will require serious considerations of "choices, consequences, best practices and requirements inherent in data practices" (at 24). Considering the important practical and legal consequences that may arise from these issues, the Report suggests a public debate on these issues rather than leaving them up to developers or statisticians.

Report Issue #5: Data Literacy: Risk Scores and Automation Bias

Another important issue related to the use of AI tools in the criminal justice system relates to the interpretation of the risk assessment made by the tools. Risk scoring could be misleading and prejudicial where the user lacks understanding of what the scores really mean and how they are determined. Kelly Hannah-Moffat has noted the tendency by lawyers and probation officers to interpret high risk score by individuals to mean high risk of offending rather than simply connoting that the individuals share similar characteristics with average members of the group with that score. She noted that "Instead of being understood as correlations, risk scores are misconstrued in court submissions, pre-sentence reports, and the range of institutional file narratives that ascribe the characteristics of a risk category to the individual" (Kelly Hannah-Moffat, "Actuarial Sentencing: An "Unsettled" Proposition" (2012) 30:2 Justice Q 270 at 12 - 13).

Also related to this is the problem of automation bias, which arises from the human tendency to believe that any machine-processed information is inherently accurate, trustworthy and flawless. To address automation bias, AI tools deployed in the criminal justice system should be able to make predictions that leave the user knowledgeable as to how the predictions are made, and the results presented to judicial officers should be easily understandable and not misleading.

Issues # 6 to 10: Due Process, Public Participation and Law Reform

An important distinction was made in the Report between the statistical predictions made by the algorithmic tools and the policy decisions that transform these predictions into an "action directive". The difference between the two is that the tools measure the risk while the policy decision determines how the risk is managed (at 27). Hence, it is important for policy makers to ensure that important policy goals are appropriately reflected in the design of these tools.

The rest of the issues shift from the AI as data producers to AI as an expression of policy. In this part of the Report, the LCO tackles issues of policy bias, transparency and accountability, legal rights and remedies as well as the creation of regulatory frameworks.

The Report pointedly highlights the potential bias inherent in the policy choices flowing from the use and interpretation of the AI data. Risk assessments are composed of tolls that measure the risk and manage that risk (at 18). The measurement is based on algorithms created by a number of differing statistical measures that produce statistical data. That data is then used in another set of policy-driven frameworks to indicate how that risk can be managed. If the risk cannot be managed in the community in accordance with the policy framework, the accused person would not be a candidate for bail. This shows the intricate relationship between AI data and policy, which is often hidden behind and obscured by the concept of "evidence-based" predictions.

Any one point in the risk assessment can create a flawed outcome (at 19). Avoiding this requires the measurement data, the measurement tool, the data from the accused person, the interpretation of the measurement, and then the subjective policy management tools to be relevant, unbiased and impartial toward the individual whose life may be altered by the decision (at 26). This means that racial neutrality is not enough to create a reliable assessment algorithm. We need to recognize the injustices perpetrated against Indigenous people and the Black community to ensure fair and just outcomes. The justice system has a duty to consult with those communities most at risk (at 30). Even the approach toward the assessment can make a difference in the outcome. For instance, risk assessments tend to focus on the individual's failures while in the system as opposed to their successes. A tool geared toward failure creates an environment where personal failure is the expected norm (at 25 and see also *R v Zora,* 2020 SCC 14 (CanLII), Martin J at paras 26, 57, 79 & 87 making similar comments in the context of failure to comply with bail release offences).

The Report also raises legal capacity issues. Does the rule of law permit review of AI and algorithms and if so, what remedies are available? This is a pressing issue considering we do have AI use in Canada, yet, according to the Report, this use is not documented and therefore not open to scrutiny (at 10). The issues flowing from this question are innumerable, opening a virtual Pandora's box of complex problems. The issues start with procedural and legal rights both in common law and the *Charter* and run through evidentiary rules before landing on remedies.

In criminal law terms, these are legal rights issues ranging from investigation to punishment, which challenge every point of contact between the individual and the criminal justice system. For instance, as raised in the Report, right to counsel under section 10(b) of the *Charter* may apply when risk assessments are used. In broader terms, the case law on informational privacy and section 8 of the *Charter* will impact AI use. Procedurally, disclosure questions will arise as they have with the use of simple technology such as breathalyzer devices. The admissibility of AI metrics will bring into question the presumptions for admissibility of electronic documents under the *Canada Evidence Act*, RSC 1985, c C-5 (*CEA*) (at 31 and see sections 31.1 to 31.8 of the *CEA*). All of these legal issues require financial resources to ensure the accused person has means to raise these challenges (at 37). Access to justice is therefore a crucial component of AI use. Undeniably, AI will change our criminal law and, as the Report suggests, the legal community must be prepared for it.

In the end, the LCO response to the bias, policy, and legal rights concerns relies on the creation of best practices with the dual objectives of transparency and accountability. These are important values in implementing technology and parallel our societal expectations from our decision-makers. However, while transparency and accountability give people access to how decisions are made and why, this information does not protect people from the imposition of those decisions in the first instance. Transparency and accountability are oversight tools that apply while AI is being used or after use. Rather than a discussion predicated on its use, a broad-based discussion is needed on why we should use AI and predictive analytics at all. Transparency and accountability can inform the discussion but should not drive it. This preliminary discussion is vitally important considering predictive AI, as suggested earlier in this article, can be "a sophisticated form of racial profiling."

Report Conclusion

In the final part of the Report, the LCO posits four preliminary questions in the use of AI and algorithms in the criminal justice system (at 41). The first, "should there be a moratorium on algorithmic risk assessments or similar tools in the Canadian criminal justice system?", appears to question AI use in the first instance, albeit not as clearly as it should be framed. The real question is "why do we need AI?" In other words, it is not enough that the technology is available to do this, what we need to decide is whether we want/need to use it. The second question, "what is the potential for algorithmic risk assessments?" should be framed more neutrally and requires a cost-benefit analysis. We often praise evidence-based inquiry but here we must ask the fundamental question of how that evidence is created. This question would then lend meaning to the third inquiry as to whether AI can "advance equity, access to justice and systemic efficiency." Notably, missing in this third question is the overall question of whether AI can tangibly and measurably achieve justice. The final question is the most challenging one posed: "what is the path forward?" This is the ultimate question, the answer of which is not clear and will take all of our human acumen to answer.

Our Conclusion

As society considers the costs and benefits of AI in the criminal justice system, we must be cognizant that racial profiling, carding, and the over-incarceration of Indigenous people and members of the Black Community are embedded into the AI debate. Much of the Report critiques the heavy reliance on AI by the United States in pretrial custody decision-making. According to the Report, the United States implemented AI with good intentions, trying to find a bail release regime that was more objective and consistent than the judicial discretionary outcomes. Unfortunately, this false premise of neutrality, combined with a lack of regulation and oversight, resulted in a tool that amplified racism and bias. The LCO seems to suggest that knowing these errors is the first step to ensuring AI use and implementation is fair and just. This may be so but we must beware of our Canadian bias, which is based on a general attitude that what happens in the United States cannot happen in Canada. We must remember that history can indeed repeat itself no matter the country. The altruistic premise for the use of predictive analytics does not answer the question as to why the technology was not more carefully conceived and regulated in the United States. The lesson learned here is that good intentions cannot be a substitute for careful consideration.

Whether it is by machine or humans, pre-trial custody interferes with a person's life and liberty. It is a form of state sanctioned coercion. AI, used in this context, may prove to be "weapons of math destruction," a term coined by Cathy O'Neil to illustrate the potential devastating effect of algorithmic decision-making. The use of AI in bail release runs contrary to recent Supreme Court case law in the area, which looks to an individualized approach to bail informed by *Charter* rights and values *(*See *R v Antic*, 2017 SCC 27 (CanLII) at para 67 and *R v Zora,* 2020 SCC 14 (CanLII) at paras 6, 22, 29, 46, 47, 52, 75, 80 & 100). Those decisions have brought the criminal justice system into a more mindful space in which every person in that system is worthy of just consideration. Review and restraint by decision-makers are the twin mantras of the bail system in the *Criminal Code*. To operationalize these tools, the bail system requires informed, impartial and unbiased decision makers. The use of AI might deflect from these key requirements and distract the decision makers from the real issues. As with sentences, bail is to be tailor-made for the individual. AI would not permit this or, even worse, may appear to be individualized when it is not. In these circumstances, justice would not be done nor would it be seen to be done.

There are also capacity issues with AI and whether our system can respond to the many valid concerns raised in the Report. For instance, if we live in a "scored society" (at 25) then everyone in the legal system must understand how that score is tallied. Training for lawyers and judges must be available to ensure data literacy but other principles may interfere with this response. Judicial independence, for instance, may run against mandatory training, leaving literacy up to the individual judge. In the courtroom, training may raise bias concerns and lead a judge to make decisions based on untested judicial knowledge. Sometimes, a little knowledge may go too far resulting in less scrutiny of the evidence as opposed to more vigilance. Similar to the expert witness experience, where the Supreme Court created an encompassing judicial oversight framework, the courts must embed into the AI legal framework a robust mechanism for judicial oversight. Our evidentiary rules must take notice of this form of unique technology to safeguard against potential miscarriages of justice. It will require fundamental changes to the way we perceive justice in Canada to adequately respond to all of these issues.

In conclusion, there are indeed many complex legal and policy issues that will arise from the adoption of AI and algorithmic tools in the Canadian criminal justice system. The existing legal framework does not adequately address these issues. Hence, the LCO Report demonstrates the urgency and importance of addressing these issues to ensure that our legal standards and rules are at pace with technological development in this area of our criminal justice system (at 38).

Based on their study of the use of AI and algorithmic tools in the US criminal justice system, as well as the analysis of the criminal justice system in Canada, the LCO Report concluded that a deployment of algorithmic risk assessment tools (which it referred to as "unproven and under-evaluated technologies" at 7) in the Canadian justice system at this point would be a mistake. Before we embrace AI technology and embed it into our system of justice, the various issues and concerns raised in the Report must be addressed. This will require all stakeholders in the justice system to work together to harness the power of AI technology to promote equity, fairness and justice. Until then, AI technology must be approached with caution and concern.

This post may be cited as: Lisa Silver and Gideon Christian, "Harnessing the Power of AI Technology; A Commentary on the Law Commission of Ontario Report on AI and the Criminal Justice System" (November 18, 2020), online: ABlawg, http://ablawg.ca/wp-content/uploads/2020/11/Blog_LS_GC_LCO_Report.pdf

To subscribe to ABlawg by email or RSS feed, please go to http://ablawg.ca

Follow us on Twitter @ABlawg