INSTITUT BROADBENT INSTITUTE

BIG DATA AND CRIMINAL JUSTICE — WHAT CANADIANS SHOULD KNOW

Authors

DANIEL KONIKOFF
Centre for Criminology and Sociolegal Studies
University of Toronto

AKWASI OWUSU-BEMPAH
Department of Sociology
University of Toronto
Broadbent Policy Fellow

HUMANS HAVE BECOME DATA-PRODUCING MACHINES.

Every Google search, credit card purchase, social media interaction, and doctor's visit leave traces of information about you, where you've been, who you've interacted with, and what you like. What's more, advertisers, data brokers, and government agencies can collect and analyze the digital breadcrumbs you leave behind as you go about your day. Welcome to the world of 'big data.'

While data-driven technologies may be used for the benefit of individuals and society as a whole, they run an equal risk of entrenching discrimination and exacerbating various forms of inequality. The realm of criminal justice is no exception; big data has both the potential to infuse fairness into the administration of justice, and, more worryingly, expedite the reproduction of existing biases. Below, we outline what 'big data' is, how it is used in the context of criminal justice in Canada and beyond, and how we might think about the potential beneficial and detrimental effects of these technologies on our society.

THE CANADIAN AND INTERNATIONAL CONTEXTS

On its surface, the term 'big data' refers both to very large data sets, as well as the tools used to manipulate and analyze them. This concept, however, does not just refer to the harvested information – it also refers to the motivations behind what harvesting that information is supposed to achieve. When data is collected en masse, and algorithms (a series of instructions that tell a computer what to do) cross reference data both within and between datasets, the computational software processing the data identifies patterns within them. It is this notion of *"identifying patterns"* that serves as the backbone of predictive justice.

Predictive justice uses data on past occurrences or behaviours to make decisions about the future, such as who and where will be policed, how an individual should be sentenced given the risk they pose to others, and when someone should be released from prison. Though the international scope of predictive justice technology's use is currently unclear, research and discourse has arisen primarily in the United States, the United Kingdom, and Europe regarding predictive technologies' presence in law enforcement and the justice system.

Unfortunately, there has been a lack of both awareness and scholarship regarding how this technology is being employed across Canadian police departments, justice agencies, and courts. Journalists have reported that cities such as Vancouver, British Columbia and London, Ontario have adopted predictive policing software within the last five years, and have also pointed to Ottawa Police's Strategic Operations Centre - which monitors protests on social media - as an example of how Canadian police are using big data and predictive policing. Motherboard also recently reported that police agencies in Ontario and Saskatchewan have been using a "Risk-driven Tracking Database" (RTD), which combines information collected by the police, schools, social workers, and other community agencies to track "negative behaviour", identify potentially at-risk people, and to deploy resources for "proactive intervention." The RTD is just one aspect of a "Hub model" presently employed in more than 100 cities across Canada, and is a technological offshoot of the model's aim to encourage social service agencies to collaborate and share sensitive information about "vulnerable" people. Attention has also recently been drawn to Project Wide Awake, an RCMP-led operation involving wide-scale monitoring of individuals' social media activity. The initiative began at least two years ago as a reactive measure designed to help the RCMP analyze social media accounts linked to specific criminal investigations. Since then, the project has shifted toward more proactive use, as the RCMP has started scouring social media to identify crimes in progress or stop planned offences before they occur.

However, little else is publicly known about the full extent of these technologies, the particulars of their use, and specifics about the impact that they have had on Canadian law enforcement. The same goes for sentencing algorithms and data-driven risk assessment tools, whose presence in Canadian courts and corrections (with the exception of the Level of Service Inventory Revised (LSI-R), which will be discussed in more detail later) is minimally reported on, let alone fully understood. While journalists and commentators alike have pointed to the potential risks to privacy these invasive and predictive technologies pose, the full extent of their repercussions is presently shrouded in questions.

Whether we like it or not, big data and algorithmic decision-making have become embedded within processes of justice administration around the world. These predictive technologies are appealing because they claim to make justice a speedier, more egalitarian affair; they take complicated and potentially-biased discretionary decisions – such as who to police and who to assess as "higher-risk offenders" - and reduce these decisions to scores, numbers, or dots on a map. In so doing, these technologies can incur greater costs than benefits; as will be explored in detail below,

predictive justice technologies, even with the noblest of aims, stand to increase the inequalities that already exist in how justice is carried out. Though it is unclear whether the Canadian justice system has embraced these technologies to the extent that the US system has, Canadians should nevertheless maintain some skepticism about the use of big data and algorithmic decision-making in the administration of justice. Below are just some of the many ways in which big data and predictive justice have been incorporated into the pillars of criminal justice process.

POLICING

Policing is a traditionally reactive enterprise, with officers responding to service calls and dealing with situational demands as they arise. When police rely on big data and predictive analytics, however, their policing becomes more proactive, sometimes aggressively so.

The basis for this development is "hot spots policing," an ideology in which police deploy greater resources to segmented geographic zones where crime is supposedly concentrated. Like hot spots policing, *predictive policing* software employs a variety of variables to identify specific areas in which future crime is most likely to occur. Predictive policing algorithms crunch large swaths of data from diverse sources to uncover patterns in criminal behaviour and develop maps of where different types of offences are clustered. From there, departments deploy officers to areas that algorithms have flagged as high-risk; "predictive policing," writes Andrew Guthrie Ferguson, "literally changes where police go."

Police departments – predominantly in the US, but increasingly in Canada – can choose from a variety of predictive policing software packages available on the market, each with their own algorithms and input variables. HunchLab, a predictive policing model implemented in St. Louis, Missouri, after the Ferguson protests of 2014, draws its predictive power from crime data, census data, and population density, and factors in other variables such as the location of schools, churches, bars, clubs, and transportation centres. Conversely, PredPol, the most widely used predictive policing software in North America, is far less granular than HunchLab, and only uses past crime type, location, and time of crime to generate predictions about where, when, and what type of crime is most likely to occur.

On top of helping police target particular neighbourhoods, predictive analytics can help police direct their resources towards 'gangs' and even individuals. One such example of this method of predicting crime using individual-focused modeling is Chicago's Strategic Subjects List (SSL) pilot project. Developed by the Chicago Police Department and the Illinois Institute of Technology, and funded by the National Institute of Justice, the SSL's prediction model generates a risk score based on past criminal history, arrests, parole status, gang affiliation, and social ties to previous homicide victims, in order to identify individuals who are most likely to be involved in a shooting as either a victim or a perpetrator. Police then use this information to dispense "custom notification letters" to the people on their "heat list," warning them about the future and offering them violence prevention services – essentially, letting them know they are being watched.

WHEN POLICE RELY ON BIG DATA AND PREDICTIVE ANALYTICS, HOWEVER, THEIR POLICING BECOMES MORE PROACTIVE, SOMETIMES AGGRESSIVELY SO.

COURTS

Risk and needs assessment scores generated by algorithms are instrumental in the evidence-based sentencing movement. Advocates of this ideology believe that embracing algorithms is the key to tackling some of the justice system's most sustained problems. By using predictive technology in the courts, risk assessment scores may help to 1) reduce judicial disparity; 2) promote consistent sentencing; 3) better prioritize and allocate correctional resources; 4) adjust punishments for certain categories of offenders; 5) reduce prison overcrowding, and; 6) encourage the use of alternative, non-incarceration sanctions.

In the courtroom, judges can use risk assessments generated by predictive algorithms to determine a defendant's likelihood of reoffending. Judges may then use the risk assessment to influence their sentencing decisions. For example, the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is a risk assessment tool that generates a risk score based on a defendant's responses to a 137-question survey. This survey screens for a wide swath of variables, including criminal history, residential stability, education, vocation, recreational habits, and social isolation. From their responses, defendants are assigned a score between 1 and 10, with the higher numbers suggesting a greater likelihood of engaging in future crime.

Similar algorithms are beginning to be used in bail decisions, too. California recently passed a bill that would eliminate the cash bail system and replace it with one in which algorithmic risk assessments recommend whether an accused should remain in jail or be released. Though the State of California has not specified which algorithms its counties must adopt for its revised bail process, there already exists some algorithmic models designed for this very purpose. The Public Safety Assessment (PSA) tool, for example, advertises itself as capable of "[improving] pretrial decision making by providing judges with more information" without having to "rely on factors such as race, ethnicity, or geography." Adopted in dozens of US jurisdictions, the PSA uses nine factors (age at current arrest, current violent offence, pending charge at the time of the offence, prior misdemeanor conviction, prior felony conviction, prior violent conviction, prior

failure to appear in the past two years, prior failure to appear older than two years, and prior sentence to incarceration) to predict each of the following three pretrial outcomes: Failure to Appear, New Criminal Activity, and New Violent Criminal Activity. However, since some of these factors – such as age at current arrest, those pertaining to prior convictions, and those pertaining to prior failure to appear charges – are likely skewed by race and socioeconomic status, whether or not the PSA actually fosters more equitable outcomes for defendants remains to be proven.

CORRECTIONS

Big data and predictive algorithms do not seem to factor into corrections as ubiquitously as they do in policing and the courts. Predictive justice's usage in corrections, however, is nevertheless important to raise, for it may help make predictions about individuals both behind bars and beyond them

Canadian correctional institutions, for example, have utilized the Level of Service Inventory-Revised (LSI-R) – originally developed in Canada – to make decisions about the level of supervision and treatment services required for a given individual. Though not algorithmic in the technological sense, the LSI-R is actuarial, in that it generates scores using responses to a 54-item survey grouped into 10 subscales: Criminal History, Education/Employment, Finances, Family/Marital, Accommodations, Leisure/Recreation, Companions, Alcohol/Drug, Emotional/Personal, and Attitude Orientation. Unlike most risk assessment tools, the LSI-R's doesn't just estimate recidivism; it is also used to predict parole outcome and the likelihood of success in community supervision, and to aid in decisions concerning bail, program designation, and security rating. However, when it comes to predictive accuracy, the LSI-R – as well as COMPAS, mentioned earlier – has questionable reliability along racial and ethnic lines. Two studies have shown that, across both LSI-R and COMPAS, Black offenders were more likely to be "overclassified" (predicted to be rearrested when they actually were not) than White or Hispanic offenders. Because judges view higher scores as suggestive of a higher likelihood of reoffending, it is likely that the LSI-R and COMPAS are giving these very officials the algorithm-derived permission to treat Black offenders more harshly than others.

Other software exists that streamline specialized services to inmates who need them the most. HarrisLogic, currently used in Dallas, Texas, is a new software used to determine which inmates suffer from mental illness. Pooling and cross-referencing data from jails, police departments, emergency services, mental health and social services, courts, and hospitals, HarrisLogic helps prison officials determine which prisoners need more specifically designed mental health services. It is also used to determine which prisoners shouldn't be in prison at all, providing information on who should be redirected away from jails and towards more therapeutic forms of care.

Predictive justice algorithms can also be used to evaluate offenders' post-release needs and anticipate what sort of intervention plans may best suit them upon re-entry into the community. For example, if the data shows that a particular inmate scored low on an education subscale, correctional professionals could use this information to make sure that the individual receives education-geared interventions once on probation or parole. Most important for prediction software used in the community supervision context is that the software is intended to be *dynamic*, adjusting risk and needs assessments according to how the ex-offender is performing out in the community.

The Pitfalls of Big Data, Algorithms, and Predictive Justice

Big data and algorithmic decision-making may make the administration of justice more efficient and more equal. But with these potential benefits come a high likelihood of negative outcomes, as these technologies

BLACK OFFENDERS ARE MORE LIKELY TO GET HIGHER RISK SCORES, WHICH SUGGEST TO JUDGES A HIGHER LIKELIHOOD OF REOFFENDING. AS A RESULT, THESE PREDICTIVE TECHNOLOGIES GIVE JUDGES THE ALGORITHM-DERIVED PERMISSION TO TREAT BLACK OFFENDERS MORE HARSHLY THAN OTHERS."

and algorithms, when exploited, may jeopardize the basic human rights of those within or beyond the justice system.

Algorithms, big data, and predictive justice may help police departments allocate resources better and redistribute their officers to focus their patrols on areas the data deems most "in need of policing." They can hasten the time it takes for officers to make an arrest by cross-referencing databases far quicker than police professionals ever could. These technologies can be used to cut court backlogs, draw snap predictions about someone's likelihood of reoffending, assign appropriate risk scores to offenders, and help reduce prison overcrowding in the process.

With these outcomes as possibilities, it's no wonder that governments and justice agents view this technology as desirable. The circumstances vary as to what inspires municipalities or regions to embrace predictive technology; for example, Andrew Guthrie Ferguson, author of *The Rise of Big Data Policing*, believes that it is the confluence of police frustration with dwindling resources and Black people's dissatisfaction with police discrimination that has led cities where both of these issues are prominent to turn toward more seemingly objective, resource-savvy policing technologies. More universally, predictive justice technologies' appeal lies in their capacity to maximize productivity while minimizing resource expenditure. Through the Department of Justice's Bureau of Justice Assistance, the US federal government has already provided police, courts, and corrections with millions of dollars in innovation grants in order to fund both the research and implementation of data-driven justice technologies in key municipalities. In applying for federal funding (usually without public consultation), these institutions are given the freedom to play around with new ways to police and prosecute, without stressing their organizations' operational budgets. At present, the Government of Canada's Justice Policy and Innovation Program (JPIP) offers funding opportunities to agencies exploring progressive justice projects, but it is unclear whether this program has attracted technological, data-minded innovations similar to those of its neighbours south of the border. Regardless of this ambiguity, the presence of predictive policing software in certain Canadian cities mentioned earlier, as well as Corrections Canada's use of the LSI-R, suggest that Canada has a vested interest in the efficiency and expediency that big data and algorithmic models bring to bear on matters of justice

Faster decisions aren't always better decisions

However, just because predictive justice technologies allow criminal justice actors to make decisions faster

IF THE DATA THAT CRIMINAL JUSTICE ACTORS COLLECT AND FEED INTO THEIR DATABASES ARE BIASED, THE RESULTS OF THE PREDICTIVE ANALYSIS USING THESE DATA WILL BE BIASED, TOO. PREDICTIVE JUSTICE TECHNOLOGIES ARE AT HIGH RISK OF SIMPLY REPRODUCING THE BIASES AND PARTIALITIES THAT ALREADY OPERATE IN THE CRIMINAL JUSTICE SYSTEM.

doesn't mean that they make decisions better. One study, for example, found that COMPAS, the above-mentioned risk assessment tool, is less accurate than humans in predicting a defendant's likelihood of reoffending (humans in COMPAS-using states were able to predict recidivism with 67% accuracy - 2% greater accuracy than COMPAS). Similarly, police relying on big data may have greater access to information about the citizens they are sworn to protect, but that information might be incorrect or outdated. Police responding to a 911 call at a residence, for example, may be notified by a quick database check that the occupants of that residence have a criminal record – a notification that may no longer hold true if officers only have access to dated housing records.

With that said, it is important to highlight that, unlike the human decision-makers driving the criminal justice system, risk assessment algorithms, analytical software, and databases aren't flesh and blood. Impartial technologies will supposedly have a neutral approach to predictive justice, removing the aspect of "relying on one's gut" and personal experiences that make evaluations and predictions by criminal justice actors so prone to inconsistency.

Biased data generates biased predictions

But just because these technologies are not subject to the vicissitudes of the human mind does not mean that they are perfect. One of the biggest shortcomings of big data and algorithmic decision making in criminal justice lies in the quality of the data used. If the data that criminal justice actors collect and feed into their databases are biased, the results of the predictive analysis using these data will be biased, too. Predictive justice technologies are at high risk of simply reproducing the biases and partialities that already operate in the criminal justice system.

Take predictive policing, for example. Predictive models of policing need data; but if that data is compiled by a police force that disproportionately targets minority populations, the model will *learn* to generate predictions that simply perpetuate the data it's received, thereby upholding racial bias (and violating any claim the technology can make toward supposed neutrality). Furthermore, if certain minority groups are more likely to be convicted because of their race or the neighbourhood in which they live, then sentencing algorithms that use prior convictions or home addresses as key factors to determine the length of the sentence will exacerbate

existing inequalities. This is particularly problematic, considering that many social, cultural, political, and economic factors contribute to the over-policing and excessive criminalization of minority populations clustered in disadvantaged neighbourhoods in the first place.

For example, in her article about the Los Angeles Police Department's highly technologized policing practices, Sarah Brayne discusses how predictive policing is caught in a racialized feedback loop of its own making. As part of recently suspended Operation LASER (LA Strategic Extraction and Restoration), LAPD officers were informed of a problem crime in a particular division. They were then instructed to gather intelligence on their patrols by stopping suspected individuals and filling out field interview (FI) cards replete with personal information. From these patrols and FI stops, officers generated a list of "offenders" who were each assigned a point value and given a numerical rank. This number system, however, was cyclical: having a high point value was predictive of future police contact, but each police contact further drove up the individual's point value. If officers were driven to disproportionately stop people of colour or individuals in disadvantaged neighbourhoods, and if someone got a point for every time they'd been stopped by the police, then those with the greatest likelihood of being stopped *again* were simply those who had been stopped before – regardless of whether they had any criminal involvement. Thus, when it comes to race and socioeconomic status, predictive justice technologies run the risk of continually disfavouring the most vulnerable of populations, producing justice outcomes far less equitable than one would hope.

Proprietary software hides algorithms from scrutiny

What's more, these algorithms used in criminal justice decision making remain the property of the private companies that develop them. Software companies usually prevent outsiders from "opening up the black box" of predictive algorithms, under the guise of protecting their intellectual property. However, without cracking these black boxes open, it is difficult to discern how these algorithms actually function – neither the governments who buy them, the justice officials who use them, nor the people upon whom they are used, necessarily know how they work. Not only does this raise questions regarding transparency, but it also threatens procedural justice; if we don't know how these algorithms crunch the data they are fed, it may be more difficult to say with certainty whether these algorithms are, themselves, actually as fair as they purport to be.

Decisions based on categorization instead of individual behavior

This unfairness also manifests in the way predictive justice technologies dehumanize their subjects through categorization. One of the pitfalls of big data is its tendency to view people not as complex individuals but rather as "parts of categories." Predictive software analyzes individual data against other like individuals, and uses these comparisons to make decisions, identify patterns, and see how individuals align with the broader category in which they've been lumped. This poses a particular problem for risk assessment tools and their related algorithms; they may purport to take into account theoretically relevant variables about a given offender, but make decisions based on aggregates (and how the offender compares to others) rather than on individual behaviour. Such a consequence may resonate more with select minorities, who may be deemed more risk-prone because of their "algorithmic alignment" with a category preordained as "risky" or "dangerous."

Correlations that aren't always real

Predictive justice technology's decisions can also be, by virtue of the vastness and depth of big data, almost entirely misguided. One of the great myths of big data is that "more is better" – the belief that with *more* data, we can identify *more* patterns and solve *more* of the world's problems. But what big data calls "identifying patterns" is what most statisticians call merely "drawing correlations," which are oftentimes spurious or entirely irrelevant. (Click here to see how flexible correlations can be!). If not assessed with a critical eye, predictive justice algorithms could lead police, judges, and prisons to draw correlations that simply aren't there, and make life-altering decisions for people based on spurious statistical relationships.

THE EXTENT TO WHICH PREDICTIVE JUSTICE TECHNOLOGIES ACTUALLY REDUCE, PREVENT, OR ANTICIPATE CRIME IS PRESENTLY UNKNOWN.

Effectiveness is unknown

Worse yet, for all of big data's vastness and depth, and for all the objectivity with which algorithms are supposed to operate, the extent to which predictive justice technologies actually reduce, prevent, or anticipate crime is presently unknown. Articles about data-driven policing and risk assessment tools, for example, have thoroughly explored the social implications of their usage, but have, so far, devoted little time to evaluating their bona fide effectiveness at achieving their aims. Predictive justice is attractive because it allows justice agencies to "do more with less," but whether these technologies are actually "doing more" is, as it stands, a question in need of an answer.

The Canadian Context: Revisiting the Hub, RDT, and Project Wide Awake

Between Project Wide Awake, the RDT, and the Hub model in which the database is embedded, Canadians should be cognizant of the potential ramifications of these technologies. When it comes to the RDT and the Hub model, representatives from community agencies can enact interventions – such as house check-ins, hospitalization, or even arrest – without individual consent. Further, each Hub (comprised of representatives from police, mental health, welfare, and housing agencies) is locally driven, meaning they are not overseen by or held accountable to larger provincial branches of government. Furthermore, the Hub model involves both the collection and sharing of personal information; though the RTD is supposedly de-identified and does not include individuals' names and addresses, the database involves the collection of information about 100 risk factors and 51 protective factors. Not only does aggregating that many data points stand to paint a pretty detailed picture of people whose identities are supposed to be blurred, but this information can also be cross-referenced with other databases and sources of information to re-identify – and de-anonymize – the RTD. In so doing, personal information can be linked to the names and addresses of the persons to whom it belongs, rendering already vulnerable people even more vulnerable to privacy violations.

Project Wide Awake poses similar threats to privacy. Though the social media data that the project's software analyzes is technically open to the public, journalists and scholars alike have expressed concerns that the project represents a foray into mass surveillance, with the RCMP becoming "a fly on the wall in the homes of consumers." At present, the public has little option but to speculate on the pervasiveness of this technology's usage, as the RCMP has refused to release its policies on how social media information is gathered and monitored, and have refused to divulge the findings of a privacy assessment conducted before the software's implementation.

CONCLUSIONS

Big data, predictive software, and risk assessment algorithms have already significantly shifted the criminal justice landscape. These seemingly handy pieces of technology have the capacity to expedite and streamline the work of criminal justice actors, but may do so at the expense of those entering or already entrenched in the justice system. Though it is unclear the full extent to which the Canadian justice system has embraced these technologies, recent reporting on Project Wide Awake and Risk-driven Tracking Databases suggests that Canadian law enforcement has already started to accept them, and that their implementation can bring consequences for vulnerable populations. Canadians should maintain a healthy skepticism about the use of big data and algorithmic decision-making, which are only likely to grow more ubiquitous as we head forth into an increasingly technologized future.

LIST OF TECH IN NORTH AMERICA

POLICING

Operations INTERSECT

Risk-driven Tracking Database

Project Wide Awake

Hub Model

HunchLab

PredPol

CrimeScan (more here)

COURTS

COMPAS (more here)

The Public Safety Assessment

CORRECTIONS

HarrisLogic

Indiana Risk Assessment System

*List is not exhaustive