Artificial Intelligence and the Administration of Criminal Justice:

Predictive Policing and Predictive Justice. Australia Report.

By Professor Rick Sarre & Assoc Professor Ben Livings[1]

Overview

In general, the use of Artificial Intelligence (AI)-based systems in Australia's business community is growing. Machine-learning practices and the use of AI more generally are now commonplace.

The Australian Government is building policies and regulatory frameworks to pursue the goal of positioning Australia as a global leader in AI, recognising the value of harnessing 'big data' across the spectrum of governmental responsibilities.

Initiatives include:

- the recent launch of Australia's 2021 AI Action Plan as part of the Digital Economy Strategy
- the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Artificial Intelligence Roadmap
- the Artificial Intelligence Ethics Framework, which includes eight principles designed to ensure AI is safe, secure and reliable.
- the government committing over A$100 million in investment pledged to develop the expertise and capabilities of an Australian AI workforce and to establish private–public partnerships to develop AI solutions to national challenges.

AI algorithms are an attractive solution to social problems because they promise to:

- enable high volumes of data processing at speed
- enable high volumes of data processing at speed, while identifying patterns human judgement is not capable of
- supercharge knowledge management while (supposedly) removing human bias from that process
- operate with ethical principles coded into their decision-making.

Technologies based on AI are slowly entering the various stages of the criminal processes, too, and have the potential to deter crime, to investigate criminal activity, and sentence offenders.

Predictive policing is one such stage. A useful definition is as follows:

> The use of dynamic prediction models that apply spatio-temporal algorithms to core business data supplemented by secondary data sources, including internal corporate data and external environmental and socio-economic data, with the purpose of forecasting areas and times of

[1] University of South Australia. May 2023. The authors are grateful to Professor Lorraine Mazerolle (UQ), Drs Tim Cubitt and Anthony Morgan (AIC) and Frank Schiliro (AFP) for their input and information.

increased crime risk, which could be targeted by law enforcement agencies with associated prevention strategies designed to mitigate those risks.[2]

Those academics who are investigating predictive policing (for example, Jerry Ratcliffe in the US)[3] are hoping to engage AI to predict or forecast future trends and to identify future or existing perpetrators. They do so by processing 'big' data using algorithms set according to their parameters. They thus attempt to use AI to align their policing strategies with these predictions.

The purpose of algorithm-based predictive policing is to determine the locations where and times of day (or night) when crimes are most likely to be committed. This does not fundamentally differ from an officer's intuition about the probable behaviour of offenders, except that the calculation is made much more quickly, and can be applied, it is said, on a broader scale.

Predictive policing is designed to enable the police to target groups of individuals who might be engaging in criminal activity, for example by pointing to digital social networks, voice recognition, facial recognition, social media and online transactions. AI technologies are designed to provide investigative assistance, for example, by analysing 'big' data. These technologies alert police regarding whom to watch, and where, before crimes are committed. They toss up suspicious behaviours for the examination of police. In other words, AI is predicated on what is expected of human behaviour, based upon past activity.

Moreover, 'big' data labs amass information that goes beyond 'crime' data in an effort to determine the locations and times of different crimes. For example, weather conditions and transport options can be interpolated into the data sets.[4] These data are being used to assess the daily activities of people positioned next to where and when crime is most likely.

Artificial intelligence thus challenges the strict law of evidence because AI-based systems produce evidence themselves. The question for us is whether evidence produced by AI is reliable in a criminal trial. Moreover, which categories of evidence will such information fall into: witness testimony or technical expert evidence? It may be necessary to create new categories or concepts for implementing *ad hoc* rules on the admissibility of evidence in a trial. An example is the drowsiness detection and distraction warning system embedded in an automated vehicle, which monitors driver behaviour to enhance safety. Under what conditions may judicial authorities use the information given by the software robot to substantiate a charge against a particular driver? The answers are not obvious.

Part I of this report reviews predictive police practices. AI-enhanced practices to assist in justice processes more generally is dealt with later in **Part II.**

Potentially reconstructing the law of evidence in Australia (**Part III**) is not dealt with in this report as the authors cannot find any Australian literature that deals with this subject.

[2] Daniel Birks, Michael Townsley and Tim Hart (2023) Predictive Policing in an Australian Context: assessing viability and utility, *Trends and Issues in Crime and Criminal Justice,* 666, Canberra: Australian Institute of Criminology, p.2.

[3] Jerry Ratcliffe, et al (2021) The Philadelphia predictive policing experiment. *Journal of Experimental Criminology*, 17: 15–41.

[4] See in the US context, for example, Ran An, Renee Zahnow, Dorina Pojani, Jonathan Corcoran (2019) Weather and cycling in New York: The case of Citibike, *Journal of Transport Geography,* 77, May, pp 97-112. Regarding mass transit data see Gary Higgs, Renee Zahnow, Jonathan Corcoran, Mitchel Langford, & Richard Fry (2017) Modelling spatial access to General Practitioner surgeries: Does public transport availability matter? *Journal of Transport & Health*, 6, Sept, pp 143-154. Regarding commuter travel see Hexia Zhang, Renee Zahnow, Yan Liu, Jonathan Corcoran (2022) Crime at train stations: The role of passenger presence, *Applied Geography,* 140, March, Article no. 102666.

Part I

Practices in Australia regarding predictive policing

The most common reference to predictive policing in Australian literature (prior to 2020) is the work of Bennett Moses and Chan.[5] Their article cited primarily US practices and research. They also included some material emanating from the UK. They did not refer explicitly to any Australian policing practices, essentially because this field in Australia was in its naissance.

Since the Bennett Moses and Chan article, there has been some development in Australia regarding predictive policing and the use of big data.

There has been further work in predictive policing by Australian academics. In 2017, Griffith University set up a Social Analytics Lab. This Lab allows sensitive, de-identified large complex government agency data to be stored and studied to reveal patterns and insights. One of the first projects to be run in the lab was a scan of crime data recorded over the previous decade to identify patterns in burglary and car crime that could inform operational policing.[6]

Using the Griffith Criminology Institute's Social Analytics Lab, three researchers, Daniel Birks, Michael Townsley and Tim Hart, set out to test the predictability of three crimes (burglary, theft of a motor vehicle and theft from a motor vehicle) in three locations in Queensland.[7] Their results showed all three crime types were able to be forecast with varying degrees of accuracy. But they suggested that tailoring parameters and methods to the location of interest, based on local patterns regarding the volume, diffusion and concentration of crime was imperative.

The Birks *et al* study concluded as follows:

> [P]redicting the most likely location for crime occurrence in the short term seems possible, and the two prediction algorithms were able to forecast at higher rates (or equivalent rates with less data) than our idea of current approaches. Moreover, these results were achieved for multiple crime types in different study regions. The second half of the predictive policing enterprise—crime reduction—has mixed evidence of effectiveness. We only evaluate the prediction component of the predictive policing enterprise here. The next step is to consider and design effective tactical responses to preventing crimes based on the identified patterns. There is considerable commentary in the academic literature on the organisational factors that inhibit effective crime reduction and problem-solving. Police leaders should be mindful of these when considering implementing such approaches.[8]

Some police programs now boast of AI-based systems that are currently being used for predictive activities. The following paragraphs provide some useful examples:

Child exploitation

According to researchers from the Australian Institute of Criminology, machine learning analytics have allowed police to analyse complex datasets for the purpose of tracking those who would exploit children for prurient purposes. Banking transactions data have provided a useful tool for the prevention of crimes against children associated with sexually explicit materials. These data interrogate non-linear interactions among a number of variables including browsing history, time online and financial transactions.

---

[5] Lyria Bennett Moses & Janet Chan (2018) Algorithmic prediction in policing: assumptions, evaluation, and accountability, *Policing and Society*, 28:7, 806-822.

[6] https://news.griffith.edu.au/2017/06/08/million-dollar-data-labs-sheds-light-on-qld-crime/

[7] Daniel Birks, Michael Townsley and Tim Hart (2023) Predictive Policing in an Australian Context: assessing viability and utility, *Trends and Issues in Crime and Criminal Justice,* 666, Canberra: Australian Institute of Criminology.

[8] Daniel Birks, Michael Townsley and Tim Hart ((2023) Predictive Policing in an Australian Context: assessing viability and utility, *Trends and Issues in Crime and Criminal Justice,* 666, Canberra: Australian Institute of Criminology, p. 19.

In 2021 Timothy Cubitt, Sarah Napier and Rick Brown evaluated a police practice that employs AI algorithms. This study considered whether machine learning analytics could offer insight into the transaction and offending behaviours of prolific live-streamers of child sexual abuse material.

> This is an emerging body of work in which the characteristics of offenders are largely unknown. The frequency and monetary value of transactions among these individuals are particularly important and have implications for identifying these crimes among financial transactions data. Offenders did not appear to have engaged in violent offending; rather, a history of low-harm offending was common, although the under-reporting of sexual offences among children and adults is an important consideration.[9]

The analysis provided a better understanding of and ability to identify offenders who pay to watch the abuse of children via live stream.

Another report appeared in the literature on this subject in 2022. The authors conclude as follows:

> The proliferation of child sexual abuse material (CSAM) is outpacing law enforcement's ability to address the problem. … Software tools using biometric systems have shown promise in this area but are limited in their utility due to a reliance on a single biometric cue (namely, the face). This research seeks to improve current investigative practices by developing a software prototype that uses both faces and voices to match victims and offenders across CSAM videos.[10]

Their paper describes the development of this prototype and the results of a performance test conducted on a database of CSAM and ends with a caveat and a call for more research.

> Given the nature of CSAM, we were unable to ensure that various ethnicities and genders were represented equally in our test data. Therefore, it is possible that certain biases could exist. Future research should attempt to test these, and other incorporated algorithms, for such biases.[11]

Family and Domestic Violence

A study was conducted in 2020 on the design and implementation of an automated text-mining method that extracted information from a large-scale (almost half a million) set of family and domestic violence police records. The authors designed a predictive analytics approach to breaches of apprehended violence orders based on the extracted information regarding mental health. Their findings indicated not only that mining the free-text family and domestic violence police records can yield substantially useful and previously unknown information but also that text mining can fuel predictive analytics that can indicate high-risk offenders in the family and domestic violence area, impacting early prevention and intervention policies in family and domestic violence cases.

The method was suitable, they said, to examine the relationship between abuse types and victim injuries; the relationship between gender and abuse types; and the risk of escalation for victims of domestic violence. Potential also exists for this extracted information to be linked to other information sources on diagnosis of mental health problems, and for these data to be used as inputs into models that can predict future offending by repeat family and domestic violence perpetrators.

The authors offered this view of the potential for other avenues for predictive policing:

> The utility of the WebCOPS data, particularly the free-text police narratives, has never been examined in a public health paradigm. The proposed exploratory study will, for the first time, take a 'big data' approach to increase our understanding of a pernicious social problem … by

[9] Timothy Cubitt, Sarah Napier and Rick Brown (2021) Predicting prolific live streaming of child sexual abuse, *Trends and Issues in Crime and Criminal Justice,* 634, Canberra: Australian Institute of Criminology, p 18.
[10] Bryce Westlake *et al.* (2022) Developing automated methods to detect and match face and voice biometrics in child sexual abuse videos. *Trends & Issues in Crime and Criminal Justice*, No. 648. Canberra: Australian Institute of Criminology, p. 1.
[11] Ibid p. 10.

> improving our knowledge about the characteristics and patterns of these related events. There is significant potential for the approach taken in this study to be applied to other areas such as sexual offences, fraud, and other violent offences, and for the scope of the data linkage to be expanded to include other information sources (e.g., housing, welfare, and Medicare data).[12]

Queensland Police is trialling the adoption of such an approach. The police are using an AI algorithm in an attempt to flag high-risk family and domestic violence offenders. The statistics provide compelling evidence that the AI program can reduce offending by high-risk, high-harm family and domestic violence perpetrators. This has potential to significantly reduce deaths, given that 30 percent of family and domestic violence homicides in Queensland are perpetrated by offenders already known to police for family and domestic violence incidents, and that known offenders are significantly overrepresented in family and domestic violence-related suicides.

One should also note the possibilities identified by researchers regarding the police analysis of social media in order to identify potential offenders.[13]

Fortunately, there is evidence of efforts to include the ethical principles of accountability, reliability (operating in accordance with the intended purpose of risk assessment); fairness (only known, repeat offenders are assessed by the AI); and human and social wellbeing (via the effort to reduce violence by habitual offenders). This is being done, for example, by employing humans, not robots to decide whether a person is sufficiently high risk to warrant pre-emptive police intervention in the home. Queensland Police maintains that they are developing the policy and regulatory frameworks necessary to guide the ethical, effective and democratically legitimate use of AI algorithms by the public sector.

Nevertheless there are concerns

> Due to existing over-policing trends, people of colour and low-income communities tend to be overrepresented in historical data that is used to train predictive policing algorithms. This can cause the algorithm to erroneously or unfairly identify these communities as high risk.[14]

Police misconduct

Factors that protect against serious misconduct in police have been interrogated using 'partial dependence plots' or PDPs. PDPs analyse the contribution of the variable to the probability of classification to the dependent variable (i.e., serious misconduct) at different points within the range of that variable.

A study on this subject by Timothy Cubitt was published by the Australian Institute of Criminology in 2021. A machine learning analysis, 'random forest,' was utilised to produce a robust predictive model, with PDPs employed to demonstrate within-variable interaction with serious misconduct. The random forest algorithm establishes an outcome based on the predictions of the decision trees making its predictions by taking the average of the output from various decision trees. Increasing the number of trees increases the precision of the outcome.

Cubitt wrote:

> PDPs demonstrate the relationship between the outcome variable, in this instance serious police misconduct, and the independent variables within the model. Random forests detail the importance of independent variables in predicting the outcome but provide little information regarding specific individual points within those variables that facilitate a strong prediction

[12] Armita Adily, George Karystianis and Tony Butler (2021). Text mining police narratives for mentions of mental disorders in family and domestic violence events. *Trends & Issues in Crime and Criminal Justice* 629. Canberra: Australian Institute of Criminology, p 11.

[13] Jia Xue, Junxiang Chen & Richard Gelles (2019) Using Data Mining Techniques to Examine Domestic Violence Topics on Twitter, *Violence and Gender*, https://doi.org/10.1089/vio.2017.0066.

[14] Lauren Solomon and Nicholas Davis (2023) *The State of AI Governance in Australia,* Human Technology Institute, UTS, Sydney, p 21.

> rate. PDPs are a valuable technique for interpreting the random forest, as they provide insight into the point within variables that were most and least important in making predictions …. [The study] analysed a sample of 600 sworn police officers with substantiated instances of serious misconduct and a matched sample of 600 control officers.[15]

He concluded with a caveat:

> The ability to identify where an effect peaks and troughs for predictors of misconduct is important in an applied setting. … While the analytics used here provide notable insights, they must be viewed in the context of low reporting rates and the barriers associated with reporting police misconduct.[16]

A year earlier Cubitt published a similar paper with two colleagues, Ken Wooden and Karl Roberts. Again, the data-driven approach proved fruitful in predicting misconduct, but with the same caveat:

> The findings of this research support the use of data driven analytics in the analysis of police misconduct, however many of these results adhere to conventional wisdom. The finding that prior behaviour is predictive of future behaviour, particularly regarding deviance, was not novel.[17]

This finding supported the findings of their prior research.

Closed Circuit Television (CCTV) analysis

The New South Wales (NSW) Police Force is using AI to analyse CCTV data. Currently in trial, the Face Matching Service program (somewhat controversially) has access (by virtue of all Australian governments agreeing to participate in the program) to government databases of existing government-issued documents with photo ID to enable comprehensive ID verification that organised crime groups cannot alter and thereby falsify. In the NSW trial, 'privacy protection and security' are aspired to by housing photo IDs in separate State and Territory databases to which all States and Territories have agreed to grant each other access via a central hub, rather than housing all the data in one database. This 'one stop shop' poses great cybersecurity risks to all the data housed there, a theme explored by two American scholars in a 2021 contribution.[18]

Human resource information for selection of police applicants

There has been a broad adoption of AI screening of police job applications by human resources professionals to whittle down the number of applicants to a strong but manageable pool. Even within this type of application, there is huge variation in how and what AI data-processing and or decision-making could be (and in some cases is) used for internationally.

The future of predictive policing in Australia

There is little evidence in Australia that the results provided by AI-based systems (other than the cases described above) has led to any wholesale changes in policing methods. The reason is probably because there are too many unknowns about the use of the data, its efficacy, its effects on privacy, its potential to breach human rights standards, and its costs relative to the value of the information provided.

In April 2022 a report was published by the Australian Strategic Policy Institute (ASPI) on this subject. Written by Teagan Westendorf, it was designed as a snapshot, indeed a meta-analysis, of the phenomenon. The author, while highlighting its possibilities, raised number of concerns and caveats.

---

[15] Timothy Cubitt (2021) Effective management of serious police misconduct: A machine learning analysis, *Trends and Issues in Crime and Criminal Justice*, 633, Canberra: Australian Institute of Criminology, p 3.

[16] Ibid p. 11.

[17] Timothy Cubitt, Ken Wooden and Karl Roberts (2020) A machine learning analysis of serious misconduct among Australian police, *Crime Science* 9, 22.

[18] Janne Gaub & Marthinus Koen (2021) Cameras and Police Dataveillance: A New Era in Policing. In Bruce Arrigo & Brian Sellers (Eds), *The Pre-Crime Society: Crime, Culture and Control in the Ultramodern Age,* Bristol: Bristol University Press (Chapter 9)

Westendorf concludes that, while useful, AI in the predictive policing space has its limitations. AI can never replace the humanity of individuals working and engaging with their communities.[19]

The following paragraphs contain information drawn from this report.

The ASPI report notes that, for 'deep learning' types of AI, the difficulty in use of the data ranges from 'extremely difficult' to 'completely impossible,' given the 'black box' nature of decisions made by artificial neural networks learning from large bodies of data. The current limitations in this area are the key problem in figuring out how to use safely and ethically the data that are collected. Moreover, there is a need for policing agencies to regulate ethically the use of AI. If one cannot be certain of what correlations an AI is independently developing to inform data screening and decision-making, one cannot be certain that these functions are complying with appropriate and ethical principles.

By way of example, the Australian *Privacy Act 1988* controls access to and use of citizens' data (there must be no cross-matching without a court order) but it does not limit non-human decision-making that uses those data.

As the ASPI report outlines:

> For police, it could be argued that the inability to know all the variables and correlations factored into an AI insight or decision is balanced out by the ability of the AI to factor vastly more data points into its calculations than an individual or team of human analysts or police officers could. That is to say, the margin of error possible due to limited AI functionality is cancelled out by the margin of error possible due to the limited data-processing capacity of human brains. But this is a question for assessing appropriate implementation, and similarly requires an accurate understanding of the limitations of both AI and human analyses. It also risks discounting the value and agency of individual and collective human knowledge in organisations such as police forces, instead of using that strength as a starting point for AI tools' development for law enforcement.[20]

Without transparency between computer scientists, data scientists and app developers and police, it would be difficult for police to then add the human lens of receiving an AI insight or decision after the fact and then having the choice of validating and actioning it or not.

This means the limitations on transparent and comprehensible use of AI decision-making need to be sufficiently mitigated in policing scenarios, such that they do not compromise the right to privacy and equitable treatment by law enforcement. Pre-emptive policing almost by definition assumes an error rate in predicting future crime, and the act of pre-emption is itself an exercise of law enforcement power over individual citizens that has consequences and implications for how our society operates.

It is thus ASPI's view that we need to be cautious in our reliance on AI and not to overstate its potential. Police must act legally, maintain trust, and support criminal justice proceedings to the highest standards of proof. If we are relying on opaque tools such as those that employ AI, those imperatives might be at risk.

This caution is echoed by Bennett Moses and Chan.

> Predictive policing is … premised on the assumptions that it is possible to use technology to predict crime before it happens …, that forecasting tools can predict accurately, and that police will use this knowledge effectively to reduce crime. But such positive beliefs around predictive policing are often based on a mythological and unrealistic view of actual capabilities and practices.[21]

---

[19] Teagan Westendorf (2022) *Artificial intelligence and policing in Australia,* Australian Strategic Policy Institute (ASPI) Report.
[20] Ibid. p. 6.
[21] Lyria Bennett Moses & Janet Chan (2018) Algorithmic prediction in policing: assumptions, evaluation, and accountability, *Policing and Society*, 28:7, 806-822, at 807.

Bennett Moses and Chan point in particular to accountability when it comes to the decisions made by AI, and issues relating to the comprehensibility and transparency of decisions made by the software. There are also concerns about storage of the data harvested and created by AI technologies. Bennett Moses (who is a Professor of Law at UNSW and Director of the Allens Hub for Technology, Law and Innovation) set out concerns as follows:

> It goes back to any automated system – it should be transparent, it should be fair, it should be accountable, it should be evaluated and tested, and the predictive policing software industry should be doing all of those things, but most of them are not doing any of them … There's nothing specific in the law that says the police can use software to make predictions, but there's also no law saying they can't. The idea of a program running in the background which takes in diverse data on us … the rules on data sharing are jurisdiction by jurisdiction, and some don't even have proper privacy legislation.[22]

Moreover, in a 2023 Report prepared by the Human Technology Institute of the University of Technology Sydney, the authors noted a similar concern

> AI systems cause harm when they are overused, used inappropriately or deployed recklessly without regard to their second- and third-order effects. This category includes when the excessive use of AI technologies such as facial recognition or predictive policing at scale severely limits human rights. In this category are also so-called 'unintended consequences' – social, political, economic, and environmental impacts of AI systems that developers, deployers or users fail to account for or recognise.[23]

Australia does not have laws that specifically address AI, its use and data collection.

Media coverage of the use of AI in policing in Australia has not been widespread and is generally benign, but some concern has been expressed around particular initiatives. In 2021, for example, newspaper *The Guardian* reported on the trial by Queensland police (referred to earlier in this report) of an algorithmic AI tool designed to predict and prevent domestic violence incidents.[24] The report noted that the trial had been welcomed by some domestic violence campaigners, who recognised the usefulness of early – or even pre-emptive – detection and action. The possibility for early action is potentially useful in domestic violence cases, due to the combination of offending that may escalate in seriousness, alongside victims who may be reticent to report. There was hope that the 'proactive' policing enabled by the AI tool would ameliorate some of the deficiencies in the current policing approach. But the author of the article noted concerns around the use of AI in policing:

> The use of artificial intelligence in policing remains fraught – evidence that predictive policing systems ultimately reduce crime is thin – and experts warn there are significant potential pitfalls.[25]

Notably, the research cited in support of predictive policing emanated from the US (again demonstrating the relative paucity of material to draw from in the Australian context), but the article also quoted Australian-based experts who expressed concern about the technology. These experts pointed to a danger that existing biases could be reinforced through the creation of 'feedback loops' within the operation of the AI software.

Using the example from Queensland and its use of AI to tackle family and domestic violence, the ASPI report highlighted a good practice model for others (implementing AI approaches) to follow. ASPI

---

[22] (https://www.unsw.edu.au/news/2020/06/predictive-policing--will-you-do-time-before-the-crime

[23] Lauren Solomon and Nicholas Davis (2023) *The State of AI Governance in Australia,* Human Technology Institute, University of Technology Sydney, p 16.

[24] https://www.theguardian.com/australia-news/2021/sep/14/queensland-police-to-trial-ai-tool-designed-to-predict-and-prevent-domestic-violence-incidents

[25] Ibid.

noted that Queensland data scientists were employed by Queensland (Qld) Police to work closely with police officers at all stages of the AI development, training and deployment.

> Owning the whole supply chain gives Qld Police *as comprehensive as possible* (given the tech limitations discussed above) understanding and oversight of the processes by which the AI is developed, trained and then deployed and monitored by the service's in-house data scientists. This includes understanding what possible human bias has been coded into the AI, what mitigation strategies have been used, what AI biases might develop through its operation on live datasets and what should be guarded against via monitoring once the AI is deployed into live datasets. Qld Police ownership also seems to have provided an opportunity for authentic policing knowledge and judgement to be fed in at the design stage, rather than just as a retrofit after the proprietary development of a product. Critically, it also means Qld Police was able to choose what datasets the AI was trained on, which held, exclusively, historical Qld Police data. This means that, despite it not being possible to avoid coding human bias into an AI, Qld Police could be certain that the bias being coded in was that of its own historical data, and therefore known and understood. This means Qld Police data scientists can use the same training datasets as a historical resource from which to glean information on the historical human bias of the police force and try to code safeguards against it into the AI.
>
> Second, comprehensive knowledge of the AI's decisions is similarly increased by Qld Police owning and developing it in house, because the data scientists using and monitoring it and the police officers employing it in their wider work have all the same information about it as those who developed and trained it. That said, it remains impossible to comprehensively know how AIs make decisions once deployed into live datasets, as they develop more and more correlations that aren't rendered visible to monitors.
>
> Third, eligibility for assessment by the AI requires subjects to be already evidenced as high risk and high harm by their previous interactions with police in DV incidents. This means that the problem for which the AI is proposed as a solution is helping police know which of all the homes experiencing repeated family violence incidents, they should doorknock to deter high-harm violence, given that police don't have the resources to doorknock all homes that have a demonstrated record of repeated DV incidents.[26]

In all, says Westendorf, one must be cautious before placing all one's 'eggs' in one AI 'basket.' AI is a helper. It is not a complete answer. Indeed, AI may obfuscate the root causes of a problem by over-policing. It also may deflect resources away from the solutions. For example, are there opportunity costs? Are there other ways to expend resources to support victims of repeat offenders in the family court system than focusing attention and resources into AI 'solutions'?

Some mention needs to be made here on the subject of human rights concerns arising from predictive policing. These concerns arise from the paucity of information that is publicly available regarding how risk assessment tools are framed and operate, and how they are utilised to make decisions. Without this information it is impossible for individuals subject to these tools to challenge decisions about them, decisions that may affect their rights to liberty and freedom of movement.

By way of example, in 2015, the New South Wales Civil and Administrative Tribunal accepted a request from NSW police force not to release information relating to one of its risk assessment tools. In *DEZ v Commissioner of Police*,[27] the NSW Civil and Administrative Tribunal upheld a refusal by NSW Police to disclose information relating to the risk model. The refusal was upheld on the basis that the information was protected by public interest immunity.[28]

---

[26] Teagan Westendorf (2022) *Artificial intelligence and policing in Australia* ASPI Report. Pp 9-10.
[27] *NSW Police Force* [2015] NSWCATAD 15.
[28] See Australian Human Rights Commission, *Human Rights and Technology,* at page 41: https://humanrights.gov.au/our-work/rights-and-freedoms/publications/human-rights-and-technology-final-report-2021

It is impossible to say at this stage how AI-based systems for predictive policing are being or will be perceived by the public in Australia. We can go some way to determining how they will be perceived by reference to the critiques of critical criminological commentators to whom we turn now.

Critical criminology

The idea of "predictive policing" was theoretically first raised by critical criminologists exploring the notion of what they referred to as "pre-crime." This involves police and other administrative bodies with or without the use of algorithmic predictions, engaging in interventions designed to disrupt, incapacitate or restrict those deemed to embody future crime threats, especially terrorism, as explained by Jude McCulloch and Dean Wilson in their book *Pre-crime: pre-emption, precaution and the future.*[29] Jude McCulloch and Sharon Pickering have presented similar views.[30]

Here is their assessment of the problems spawned by those engaging in any 'pre-crime' assessments (in this case, in counter-terrorism strategies).

> The failure to distinguish sufficiently between evidence and intelligence and unlawful processes associated with the gathering of intelligence or the deployment of coercive powers by intelligence agencies has led to numerous failed or aborted terrorism prosecutions … The whole pre-crime project of accurately predicting threat through intelligence relies on accurate information on the variables associated with increased threat. Preventing terrorism and the pursuit of security has led to a growing and profitable field of 'crime science' that sees prediction and risk management as entirely feasible and objective ...[31]

But we shouldn't get ahead of ourselves.

> However, there has been little headway made in efforts to establish relevant and meaningful variables contributing to the risk of terrorism. Effective profiling has been deemed difficult, if not impossible …, and no statistical link has been demonstrated between 'psycho-sociological features, nationality or birthplace' and the risk of terrorism. … Reviews of the effectiveness of counter-terrorism tactics based on 'racial', ethnic and religious profiling since 11 September have found no positive results in identifying potential terrorists ... Despite this, 'race', religion and ethnicity continue to be seen and used as proxies for risk under counter-terrorism frameworks.[32]

Moreover, there are other problems associated with pre-crime that they have identified too:

> Pre-crime laws and the coercive measures that travel with them mobilize prejudice around identity and lead to intensified politicization of policing and law.[33]

A possible future

The ASPI report addresses the future need for Australian government policy and regulatory frameworks to guide police towards safe and ethical use of AI in ways that caution against following market pressures to use AI whenever possible. Here is how they present the tasks and challenges for governments seeking to build policy and regulatory frameworks. They must:

> • account for technological limitations that have negative ethical implications, and legally require the mitigation of the latter for the public and private sectors such that unethical uses are ruled out despite efficiency gains.

[29] Abingdon: Routledge, 2016.
[30] Jude McCulloch and Sharon Pickering (2009) Pre-crime and Counter-Terrorism, *British Journal of Criminology* 49, pp 628-645.
[31] Ibid p. 635
[32] Ibid.
[33] Ibid.

- use AI where it is appropriate and where there is an evidenced, net benefit; that is, in any given scenario, begin with the question 'Why use AI?' rather than adopting a baseline assumption that AI presents the right solution to all challenges.
- consider when *not* to use AI, based on ethical, legal and net benefit considerations.
- support and incentivise law enforcement and other security agencies to use AI in ways that maintain the democratic balance between civil liberty and security, just as the current electronic surveillance law reforms endeavour to do, in a legally binding way.[34]

The task is to navigate the opportunities AI presents to policing that point us towards a net benefit for Australian values, security and community safety. We commend the researchers who are currently pursuing these paths, such as Clare Southerton and Emmeline Taylor, who explore the effect of pre-crime assessments of young people who come to the attention of police using algorithmic data.[35]

Part II

Predictive justice in Australia

Introduction

The interaction of AI and the law is firmly established as a topic of scholarly research in Australia.[36] Furthermore, some limited academic work is being done in Australian universities on the use of AI in the criminal justice space.[37]

In 2009, Anna Ferrante wrote of the relatively slow take-up of data linkage in Australian criminal justice, when compared with its take-up in health sciences and medical research.[38] Since the data are de-identified in such data-linkage exercises, they have limited use for predictive policing, but Ferrante suggests that their utility for criminological research is much greater. She states:

> Knowledge about crime and what influences offending (and re-offending) feeds directly into crime prevention policy and practice which, in turn, deliver positive outcomes for individuals and communities.[39]

Hence there has been a relatively slow use of AI in Australian justice practice, when compared to jurisdictions such as the US and the UK, but the above references suggest an awareness of the potential for greater use of AI, an understanding of the uses to which it could be put, and the advantages and shortcomings of such deployment.

Phone-detection

One area in which AI seems to have been readily accepted is in relation to the use of mobile phone-detecting cameras, targeted at phone use by drivers. New South Wales was the first Australian jurisdiction (and claims to have been the first in the world) to implement the technology in 2019. The States of Victoria and Queensland have joined it. There are plans for its introduction in South Australia in September 2022. The enabling technology is produced by an Australian company called Acusensus.

---

[34] Teagan Westendorf (2022) *Artificial intelligence and policing in Australia*, ASPI Report, p 11.
[35] Clare Southerton and Emmeline Taylor (2021) Dataveillance and the Dividuated Self: The Everyday Digital Surveillance of Young People, In Bruce Arrigo & Brian Sellers (Eds), *The Pre-Crime Society: Crime, Culture and Control in the Ultramodern Age*, Bristol: Bristol University Press (Chapter11).
[36] See, for example: Michael Guihot & Lyria Bennett Moses (2020) *Artificial Intelligence, Robots and the Law*, Lexis Nexis.
[37] Stacey Hannem, Carrie B. Sanders, Christopher J. Schneider, Aaron Doyle, Tony Christensen (eds), *Security and Risk Technologies in Criminal Justice*, Canadian Scholars, 2019.
[38] Anna Ferrante (2009) The Use of Data-Linkage Methods in Criminal Justice Research: A Commentary on Progress, Problems and Future Possibilities, *Current Issues in Criminal Justice*, 3, 378-392.
[39] Ibid at p 389.

The cameras work by constantly monitoring road traffic, and capturing high-resolution images, using infrared light to penetrate through the windscreen. This image is then run through AI software, which makes an initial determination on whether phone use or a seatbelt offence has been detected. If the AI determines an instance of phone use or a seatbelt offence, the image is then passed to a human member of staff, who looks at the image to determine if there has indeed been an offence committed. The Acusensus technology therefore operates as a screening tool, picking up potential offences. The actual determination on whether to proceed to enforcement is in the hands of a human decision-maker.

The implementation of mobile phone detection cameras does not seem to have caused a public backlash in the jurisdictions in which it has been implemented, or where it is proposed. Indeed, the NSW government department Transport for NSW points to 'strong community support for using cameras to enforce illegal mobile phone use while driving or riding'. Transport for NSW cites research commissioned before implementation, during pilot operation of the technology and after full implementation, that demonstrates 74%, 80% and 79% public approval, respectively. Moreover, there does not appear to have been significant negative media coverage of the use of AI in this context.

One area where there is potential for concern when it comes to the deployment of the mobile phone detection cameras is in relation to privacy and the security of the data captured by the cameras. Since the cameras that are deployed to detect mobile phone and seatbelt offences run continually, they capture a high number of images. The potential for this to cause disquiet has been recognised, with government agencies keen to provide reassurance. Transport for NSW states that where the AI software deems there to have been no offence committed, the image will be 'permanently and irretrievably deleted, typically within an hour'. Where the AI does indicate a potential offence, the image is passed to a human decision-maker. At this point, the image is 'cropped and pixelated to remove information that would identify the vehicle or the vehicle location'. If the human decision-maker determines that there has been no offence, the image is deleted within 72 hours. It should be noted that there is no promise that the deletion of these images will be 'permanent and irretrievable', although there are assurances about the robustness of data protection measures and about the screening and training of staff. The approach taken by NSW appears to be the approach also taken in other jurisdictions.[40]

Concerns: Ethics and safety

These key initiatives all mention ethics and safety being important, especially in balancing commercial interests and incentives, but none of the government documents cited above mentions how the current limitations on AI technology may compromise those principles of ethical, safe, and explainable AI. AI, which is often cited as a solution to remove intentional and unconscious biases, can, in fact, learn such biases and propagate them. This means that the transparency of and ability to explain AI decisions (necessary to understand AI decisions) are not yet available. In other words, in Australia we are not yet able to make AI functionality sufficiently transparent and comprehensible to have confidence in all its applications. These and other concerns have been recently raised by two Australian scholars, Pat O'Malley & Gavin Smith (albeit in an International volume).[41]

Conclusion

In summary, without a detailed understanding of the strengths and weaknesses of AI, all governmental capacity to develop policy and training guidance and regulatory frameworks for AI use is constantly under challenge. Whether it be in justice processing (including in law enforcement) or any other governmental interaction with the public, the imperative to uphold democratic rights and ethical AI standards (that Australian policy mandates) should never diminish. It is not safe nor sufficient to use

[40] (https://roadsafety.transport.nsw.gov.au/stayingsafe/mobilephones/technology.html

[41] Pat O'Malley & Gavin Smith (2021). Pre-crime and the 'Control Society': Mass Preventive Justice and the Jurisprudence of Safety. In Bruce Arrigo & Brian Sellers (Eds), *The Pre-Crime Society: Crime, Culture and Control in the Ultramodern Age,* Bristol: Bristol University Press (Chapter 3).

human validation belatedly. To be sure, AI may promise but it can never deliver an entirely problem-free operation.