‘Algorithmic Impropriety’ in UK Policing Contexts: A Developing Narrative?

GRACE, Jamie <http://orcid.org/0000-0002-8862-0014>

'Algorithmic impropriety' in UK policing contexts: A developing narrative?

Jamie Grace

Abstract:

There is an increasing use of algorithmic or machine learning-based intelligence analysis in the UK policing context. Two of the most high-profile types of intelligence retention and analysis practices used by the Metropolitan Police have recently been found to be unlawful. Notably, these were i) the indefinite retention of a peaceable individual's records on a specialist domestic extremism database, and ii) the overly-lengthy retention of disproportionately BAME citizens in London on a 'Gangs Matrix'. These two findings, from the European Court of Human Rights and the UK Information Commissioner's Office, respectively, have been indications that forces that would heed the call of Her Majesty's Chief Inspector of Constabulary in 2018 to devote more resources toward investment in 'AI' for policing purposes must do so carefully. Indeed, the new National Data Analytics Solution (NDAS) project, based within West Midlands Police, has recently been the subject of critical ethical scrutiny on a number of fronts. The West Midlands force has had its own offering of a data-driven 'Integrated Offender Management' tool delayed by the demands for more clarity from a bespoke ethics committee. This has possibly headed off a later finding of unlawfulness in the courts, as there could possibly have been a challenge by way of judicial review on administrative law principles as well as data protection law and human rights and equality law. As a result, this chapter seeks to draw out lessons for policymakers from these early skirmishes in the field of 'predictive policing'. This piece also concludes with some observations about the need for a set of minimum standards of transparency in a statutory authorization process for algorithmic police intelligence analysis tools (APIATs), in a mooted Predictive Policing (Technology) Bill.

Jamie Grace is a Senior Lecturer in Law in the Department of Law and Criminology at Sheffield Hallam University. Jamie is also vice-Chair of the West Midlands Police and Crime Commissioner's Independent Data Analytics Ethics Committee. Jamie's co-authored research on the legalities of the algorithmic analysis of police intelligence led to a model regulatory framework for algorithmic decision-making in UK policing (known as 'ALGO-CARE'), that has been promoted for use by the National Police Chiefs' Council. 'ALGO-CARE' was cited favourably in a report on artificial intelligence by the Lords Select Committee on Artificial Intelligence in April 2018. Jamie's evidence to a separate Parliamentary inquiry on algorithmic decision-making in the UK public sector was cited in the final inquiry report in May 2018.

Introduction

The inevitable greater use of algorithmic police intelligence analysis tools (APIATs) by UK police forces, as something that is to be done for the 'public good', is a conflicted and unproven stance, or at least is a matter of perspective and priority (O'Neil, 2009). Each element of AI in policing that is a 'public good' to one person (resources more accurately targeted, recidivists deterred, and potential harms prevented) can entail for some people a more obvious opportunity to see controversy (a sticking plaster for austerity cuts to criminal

justice agencies, stigmatisation and the unfair labelling of some offenders, and the erosion of civil liberties). This tension over whether AI in policing is a 'public good' is the essential question in the use of the technology in policing contexts, and specifically the use of machine learning in offender risk prediction. It is sometimes immediately apparent to see the potential unfairness or risks of biases in a paper proposal from a police force as to how they would like to use data (and as is discussed below), but less obvious to see potential risks to public health, or access to justice, in uses of the same type of data technologies in practice, since these wider risks may take a longer time to emerge or to be verified.

This chapter seeks to draw out lessons for policymakers from early skirmishes in the field of 'predictive policing', where the hope is that 'algorithmic impropriety' could be better avoided. 'Algorithmic impropriety', as I have written elsewhere (Grace, 2019: 1), "is something that could be seen as [a breach of] the combination of administrative law grounds of review as part of a bundle of accountability standards that draw on wider bodies of law", in 'calling out' the mis-use or unlawfulness of particular deployments of machine learning technology in a specific policing context. This unfairness or mis-use could be as a result of one or more of 'data inequalities', 'accuracy biases' and 'decisional opacity' (Grace, 2019).

This chapter takes as its theoretical frame an important chapter by Mireille Hildebrandt (2009) on the theory of group and individual profiling. Hildebrandt first points out that profiling can be biometric, location-based/spatial, web-based/click-based, and individual versus group, etc. This piece focuses, below, on the proposal and developmental work from West Midlands Police, at the time of writing, to develop a risk profiling tool using machine learning approaches (a new APIAT), and which could identify more than 8,000 'High Harm Offenders' to be managed more intensively, police resources permitting, and out of nearly half a million known offenders living in the West Midlands area in the last 20 years. (It is unfortunate that a term like 'High Harm Offender', used by West Midlands Police, is however a subject internally labelling, and does not appear to be a widely transferable terms in policing.) The predictive power of police professionals, in this context meaning experienced officers in offender management teams, may be overwhelmed by the scale of data that it is possible to access in current WMP systems as a whole, and so would seek to work in a qualitative and responsive fashion with offenders in their charge, as opposed to more proactively, preventively and quantitatively, as they might be better able to do when aided by an APIAT. Hildebrandt (2009: 284) reminds us, after all, that "Data mining… provides the data controllers with patterns that are invisible to the naked human eye." What is not clear is how police officers, and offender managers themselves, will resist or otherwise fail to adhere to a quantitative turn instigated by an APIAT having a greater influence over their practices than other tools or processes currently.

Algorithms in the UK public sector

Algorithmic risk prediction practices, while still in their infancy (particularly so in policing circles), might be representative of a greater shift toward algorithmic governance in the longer term. In the short term, these practices can often be portrayed as problematic. For example, Philip Alston, the UN Special Rapporteur on Human Rights and Extreme Poverty at the time of writing, has argued that the "British welfare state is gradually disappearing behind a webpage and an algorithm, with significant implications for those living in poverty" (Alston, 2019: 13). For governance to be fair, then data governance must be fair, because data enables governance. Hildebrandt (2009) argues that democracy depends on both individuals being able to formulate their own ideas in a private sense, and yet also depends on a public degree of accountability and transparency in relation to state institutions in the way that they

use data about individuals to make decisions affecting their lives. Taking the development of the printing press as a transformative technological tool of governance comparable in its revolutionary power to the potential of 'big data' twinned with machine learning tools, Hildebrandt reminds us that:

> *"While the printing press first allowed the rule by law (the sovereign using written codes as a means to rule his subjects), it later enabled the rule of law (the internal division of sovereignty that separates the enactment of legal rules by the legislator from their interpretation in a court of law)." (2009: 300)*

For algorithmic police intelligence analysis tools to be moved in their use, from not just simply tools to more efficiently rule *by* law, but to become tools to better secure the rule *of* law by, through and against law enforcement bodies, methods must be developed to enhance the deliberative democratic basis of these APIATs. Ten years ago, Hildebrandt herself wrote that (2009: 307): "Creating transparency and privacy-enhancing tools should enable a citizen to contest the application of (group) profiles, rejecting the idea that one can be judged on the mere basis of a correlation." This is because, of course, and as Holmes reminds us: "Problems arise with false correlations…" (2017: 58). In policing contexts these false correlations may be discriminatory, or be a factor in depriving a person of their liberty, or facilitating a disruption of their family life, and so on.

Worryingly, in our contemporary society we run the risk of predictive analytics supported by 'big data' being *too* influential on the humans 'in the loop' (Keats Citron and Pasquale, 2014). But linguistically, at least, it was ever thus, it would seem. As noted by Holmes (2017: 3), there was a historical usage of the term 'data' to mean an article of faith:

> *"The Oxford English Dictionary attributes the first known use of the term ['data'] to the 17th-century English cleric Henry Hammond in a controversial religious tract published in 1648. In it Hammond used the phrase 'heap of data' in a theological sense, to refer to incontrovertible religious truths."*

And yet the same data science that builds an algorithm for use in a policing context can highlight for us the degree of expected (in)accuracy - what the algorithm alone cannot do is recommend that its predictions can be trusted, or should be. Some writers have set out examples of why contextual knowledge is vital in determining the ability of an algorithm to get at 'the truth'. For example, the acceptability of the balance between 'statistical sensitivity' and 'statistical specificity' of an algorithm, and thus the overall number or rate of 'false positives', will depend very greatly on the prevalence of risk in the overall population whose data are processed by the tool concerned. Haigh gives the following hypothetical example of a terrorist profiling tool in an airport, where an emphasis has been placed on statistical sensitivity (and an algorithm emphasising detection over the avoidance of false positives) (Haigh, 2012: 101-102):

> *"...suppose that the probability a real terrorist evades [profiling in airport] checks is tiny, 1/10,000, while the chance that an innocent person is led away for an intensive interrogation is a miniscule 1/100,000. How likely is it that someone picked out is guilty? ... We cannot answer the question without some idea of the would-be passengers who are terrorists. Try one in a million - frighteningly high, given that Heathrow handles… fifty million passengers a year. But the figures assure us that, even with fifty potential terrorists, it is overwhelmingly likely that all will be detected. ... Unfortunately, five hundred innocent passengers will also be detained! Among those stopped by this system, fewer than 10% are terrorists. And if there are fewer*

> *than fifty terrorists, the chance that someone who is stopped is indeed guilty are even lower. Detection methods must have much better performance figures if they are to be useful."*

As well as substantive issues of accuracy and false positive rates, there are questions of procedural justice posed by the use of AI in policing contexts. Boden (2018) acknowledges that something in meaning and nuanced can be lost in explaining or investigating issues by using a chat-bot, even a state of the art one. Is it merely enough that information is gleaned by the police from a victim of crime, using an AI-powered simulacrum of an investigating officer? Isn't something of police legitimacy lost along the way?

These questions of the type of legitimacy fostered by the use of algorithmic police intelligence analysis tools are crucial ones to be posed, in this pivotal era for 'big data' in policing contexts, since much the data drawn upon is indicative of the discriminatory treatment of ethnic minorities and the working class (and members of the ethnic minority working class) as disproportionately vulnerable citizens and residents in the UK. Akala reminds us of this set of issues, in asking rhetorically (2019: 205): "Who attacked the miners at Orgreave? Who lied after Hillsborough? The job of the police is to protect the state and working class people obviously do not control the state in any meaningful sense." Now we must pose awkward questions around class and race in terms of the use of big data in policing, too, following the criticism on equality grounds of the use of the Gangs Matrix by the Met in London in recent year (ICO, 2018), discussed below.

Nevertheless Her Majesty's Chief Inspector of Constabulary (Sir Thomas Winsor), argues that (2018: 34): "…there is real potential for technology, such as artificial intelligence (AI) and machine learning, to make the police more effective and efficient. [If] the police are going to be able to prevent and detect crime in the future – particularly technology enabled crime – they need to invest now in the technology and training to do so." Sir Thomas' call for investment and emphasis one the cost-effectiveness of policing through data analytics (not yet well-substantiated in practice, to the knowledge of this author, but often cited in theory), bears repeating in detail here (2018: 35):

> *"The opportunity here is not only to get machines to do faster what the police already do. It is also to use technology to achieve police objectives in ways we have not even thought of yet, and might never. Instruments and technology exist today which can process information far faster, more efficiently and more reliably and effectively than any human could. But, even more significantly, the capability exists now to devise ways of learning – of machines thinking for themselves – which no person has ever achieved, and perhaps no person ever could. Preventing, investigating and determining the causes of crime all involve numerous complex factual permutations, unpredictable human behaviour and random as well as intended events. This is the perfect field for the application of smart [analytics]. Of course, this technology costs money. But if the police invest now, working with the experts who have created and are developing these capabilities, this powerful technology has the potential to make them more efficient, and achieve huge advances in public safety and security, and timely justice."*

The Gangs Matrix case study

While the Gangs Matrix used recently in London is not a machine learning-based tool, it is a scoring system which gives rise to problems that are indicative of the general trend. The system, operated by the Metropolitan Police over a period of years, is a worthy case study that shows that police investment in database technologies must be done with requisite care, and the use of predictive tools undertaken with requisite due process.

In November 2018, the Information Commissioner's Office (ICO) published an Enforcement Notice which highlighted failings on the part of the Metropolitan Police Service in operating its Gangs Matrix (ICO, 2018). The Enforcement Notice found that there were breaches of the data protection principles under the Data Protection Act 1998 (which was in force for much of the life of the Gangs Matrix from inception) as well as suggestions of a breach of the public sector equality duty under the Equality Act 2010. In December 2018, the Mayor's Office for Policing and Crime (MOPAC) published a review of the use of the Gangs Matrix over a five-year period, June 2013 to May 2018 (MOPAC, 2018).

In the period June 2017 to May 2018, 82.3% of people on the Matrix were BAME, 99% Male, and 55.6% under 18 years of age. (MOPAC, 2018: 25). One cannot escape the fact that the Matrix had to an extent become a means of managing and calibrating the criminal punishment of thousands of black teenage boys in London. As such, it is no surprise that in considering the need for particular reforms to the use of the Matrix, the MOPAC report concluded "there is no room for complacency or rashness." (MOPAC, 2018: 41)

It is worth setting out here the scoring and grading system of 'nominals' (individuals) for gang-related offending. Individuals graded 'red' in terms of risk of gang-related offending would see 'daily activity' aimed at them, based on a multi-agency plan put in place for them - while an assessment is made for a judicial order known as a Criminal Behaviour Order. Amber-graded individuals have the same sort of plan drawn up, but 'activity' to disrupt their criminality will be less than daily, and an assessment for the potential for a Criminal Behaviour Order would not be a strict necessity. Finally, green-flagged individuals would receive 'diversion or engagement activities' from primarily just one key agency (MOPAC, 2018: 21).

In September 2018, 4% of individuals on the Matrix were scored red, 31% Amber and 65% Green (MOPAC, 2018:20). MOPAC were concerned that individuals were still maintained as listed on the Matrix when they posed little risk of engagement with gang-related activity, recommending "a thorough reappraisal of the individuals in the Green category, with a focus on: those that currently score ‘zero-harm’; those that have never had a harm score or have remained in the Green category for their entire time on the Matrix; and those under the age of 18." (MOPAC, 2018:40) There was a serious problem with the Matrix in terms of what is known in European human rights law terms as the 'accessibility' principle, concerning a lack of transparency and thus a shortfall in the publically-available information provided to those included on the Gangs Matrix. This meant that "the MPS’ lack of a clear, publicly available policy document specifically setting out how the Matrix operates [was] an important shortcoming." (MOPAC, 2018: 42)

However, not all legal principles support the idea of greater accountability over how the Matrix operated. In 2017, the High Court had rejected the idea that sporadic or occasional police monitoring of social media accounts containing 'intelligence' made public by an individual could engage the right to respect for private life in European human rights law

(under Article 8 of the European Convention on Human Rights). A particular element of this case, brought by Salman Butt, had turned on whether a person who places evidence of their own harmful views or lifestyle online, in a publically-accessible manner, had a 'reasonable expectation of privacy'. The High Court (later supported by a judgment from the Court of Appeal in the same case) found that they would not, where police surveillance of those online expressions was light-touch, irregular or one-off profiling. The *Butt* case will then have ramifications for the use of 'SOCMINT' (social media intelligence), which can be fed into scoring and risk prediction systems such as the Gangs Matrix (Grace, 2017; 2019). As the MOPAC report on the Gangs Matrix explained:

> *"It has also been suggested that where the police view public profiles and access open source material in order to help inform a decision as to whether an individual should be listed on the Gangs Matrix, then in every case this should require authorisation for directed surveillance under RIPA [the Regulation of Investigatory Powers Act]. It is however doubtful that viewing and considering material that has been placed online by an individual and made publicly available by them would usually constitute surveillance for RIPA purposes. If the viewing was intensive and repeated in relation to a specific target individual, then this might perhaps cross the line into being directed surveillance [requiring more formal authorisation], but it is not easy to define where the border might be." (MOPAC, 2018: 46)*

In conclusion, the MOPAC report recommended that (MOPAC. 2018: 55):

> *"Both the Operating Model and the training should have a particular focus on ensuring:*
> *• that the right people are on the Matrix;*
> *• that people are added and removed in a standardised, evidence-based manner;*
> *• that they can be removed and that the 'gang' label will not 'follow' them;*
> *• that local Matrices are refreshed regularly so that individuals don't stay on any longer than necessary;*
> *• that the guidance on the use of social media for intelligence purposes is updated; and*
> *• that the Data Protection principles and legislation are fully applied."*

The problem of course, is what we mean by 'the right people': who deserve, we might decide, because of the risk *they* pose (and not because of their family or the estate they grew up on) to be listed for a particular period on the Gangs Matrix. The Gangs Matrix, and any successor system in London, is now regulated as with all data processing by law enforcement bodies in the UK, by the Data Protection Act 2018. The MOPAC Report does not mention S.47 DPA 2018, despite the report being published after the DPA 2018 had come into force. Section 47(3) of the DPA 2018 requires that data processing should be restricted (i.e. not used in an organisation) where "it is not possible to ascertain whether it is accurate or not". And under S.205 DPA 2018, 'inaccurate' data is data which is 'incorrect or misleading'. As a result, incorrect of misleading data should be restricted in its use by a police force. So what about all those 'green'-flagged individuals or 'nominals' with no real risk score at all? Or with zero scores? They are patently not gang members to the best knowledge of the MPS, and their profiles arguably should have been deleted immediately following the publication of the MOPAC report, since the 2018 Act was in force from May 2018.

Timothy Pitt-Payne QC has suggested that the transparency of the Gangs Matrix should be augmented (and in a way that furthers the duty on the MPS under human rights law, following the 'in accordance with the law' criterion) (Pitt-Payne, 2018: 77) with "a public-facing document covering topics such as the following":

> *"The purpose of the Gangs Matrix, including: encouraging individuals to divert from gang membership; managing the risks presented by the individuals listed; and managing the risk that those individuals will themselves be victims of violent crime.*
>
> - *The criteria for inclusion on the Gangs Matrix.*
> - *The basis on which an individual is scored.*
> - *The practical consequences of being listed on the Gangs Matrix with a particular score.*
> - *The circumstances in which an individual's listing will be changed, or in which an individual will be removed from the Gangs Matrix altogether.*
> - *The arrangements for sharing information from the Gangs Matrix on a London-wide level.*
> - *The arrangements for sharing information at borough level."*

Shortcomings in transparency around police uses of APIATs are perceived with regularity now. In general, says Peter Yeung (2019), "…a lack of transparency continues to plague the field." Readers will note that the Enforcement Notice from the ICO in relation to the Gangs Matrix led to a commitment by the MPS to operate the Matrix lawfully by the end of 2019, and in line with the recommendations from the MOPAC report. This would include a greater degree of transparency. And yet, this greater transparency has already been missed in relation to the newly-revealed 'Concern Hub' - a predictive analytics tool, purportedly once more set to be used to "safeguard young people at significant risk of becoming involved in violence, drugs, or gang activity" (Yeung, 2019). Wil Crisp, writing in *The Independent* in March 2019, noted that in the drive to replace the Gangs Matrix with a newer system, senior Metropolitan Police Officers were assuring him that "across the board lessons have been learned" (Crisp, 2019). And yet confirmation that the recommendations from the MOPAC report on the Gangs Matrix have been progressed has not been forthcoming, nor has the ICO explained any element of progress on the reform of the Matrix (although the ICO has released general guidance to police forces on the operation of gangs databases (ICO, 2019a)).

As the MPS has been set the task of eventual greater transparency over the Gangs Matrix, and rapidly developing the capacity of their new Concern Hub project, the ICO has also been investigating the London Borough of Newham - which it fined £145,000 for unlawful sharing of un-redacted Gangs Matrix content with partner organisations (ICO, 2019b). Some of the details of more than 200 individual profiles on the Matrix were leaked as a result of this breach to gang members on Snapchat. Some of the 200 individuals identifiable in the material, leaked in 2017, went on to become victims of violent gang-related crime.

In the West Midlands, in the last two years, there has been work done by the police, funded by the Home Office through the Police Transformation Fund, in building a true machine learning-based predictive tool, unlike the cruder Gangs Matrix in London, that will be used to identify 'High Harm Offenders'. By way of distinct contrast, this WMP predictive policing project has made efforts, at least in comparison with the older and less sophisticated Gangs Matrix, to be more transparent from the start.

The West Midlands case study

A revealing piece in the *New Scientist* in November 2018 highlighted the way that the National Data Analytics Solution (NDAS) had determined to use its £2m funding from the Home Office Police Transformation Fund to draw on an enormous amount of data from across a number of force areas, to become the largest single UK predictive policing data project to date (Baraniuk, 2018). According to Baraniuk the NDAS was designed to draw on:

> *"…local and national police databases, including records of people being stopped and searched and logs of crimes committed. Around 5 million individuals were identifiable from the data… Looking at this data, the software found nearly 1400 indicators that could help predict crime, including around 30 that were particularly powerful. These included the number of crimes an individual had committed with the help of others and the number of crimes committed by people in that individual's social group… The machine learning component of NDAS will use these indicators to predict which individuals known to the police may be on a trajectory of violence similar to that observed in past cases, but who haven't yet escalated their activity. Such people will be assigned a risk score indicating the likelihood of future offending." (Baraniuk, 2018)*

The Alan Turing Institute and the Independent Digital Ethics Panel for Policing concluded in their *Ethics Advisory Report for West Midlands Police* that the necessarily multi-agency repercussions of adopting a National Data Analytics Solution that facilitated predictive interventions would see data-driven policing put at the heart of public protection work of that state, a "significant change with profound ethical, institutional and policy implications."(Alan Turing Institute, 2017: 3) The National Analytics Solution Project Team then gave a *Response to the Alan Turing Institute and IDEPP*, explaining that ethical safeguards including data protection impact assessments, proportionality reviews of pilot phases of the work, and independent ethical panel scrutiny would all be implemented. The Project Tem also noted that to "fail to take the opportunity that technology offers is to fail the public we serve by misusing the resources they provide policing with." (National Analytics Solution Project Team , 2017: 4). It was then that the Police and Crime Commissioner for the West Midlands (WMPCC) and WMP jointly established a scrutiny panel and procedure in the form of an independent Data Analytics Ethics Committee, in late 2018 and early 2019 (WMPCC, 2019a). The WMPCC ethics committee can give advice on project, and can urge caution to the extent of rejecting outright those proposals which are flagrantly unethical or distinctly unjustified. The Chief Constable of West Midlands police is as a matter of law not bound by the advice or recommendations of the WMPCC ethics committee - albeit in administrative law terms a chief constable who failed to take into account at all the relevant considerations of ethics committee advice may act unlawfully in taking a decision against such advice.

The first task of the WMPCC ethics committee was to consider a working proposal for an 'Integrated Offender Management' (IOM) tool, which would score more than 8,000 individuals as prospective 'high harm offenders', in order for the resources available to teams of police offender managers to be better targeted (WMPCC, 2019b). This proposal was not rejected entirely (WMPCC, 2019c), but was delayed in its move to an implementation stage, despite some praise from the ethics committee (for a careful weighting of the statistical model underpinning the tool to one that was more statistically *specific* as opposed to statistically *sensitive* (Grace, 2019). During the first design of the IOM tool, WMP had placed emphasis on a number of possible versions of the tool that preferred statistical *specificity*, ensuring a

caution around generating 'false positives' for individuals flagged as high risk. The final model chosen had a healthy respect for statistical specificity in this way.

But the ethics committee felt that more could be done to explain the safeguards around possible ethnicity biases against black offenders residents in the West Midlands, arising from using historic data available to the force - concerning the best part of half a million individuals, going back, in some case, 20 years (WMPCC, 2019c). it was also not clear from the first iteration of the model as to whether the data drawn upon would be stronger in terms of criminal process provenance (such as convictions, or charges) or whether records of arrests would also be fed into the algorithm too. Accordingly, the ethics committee required that the police 'Data Lab' working on the IOM tool should clarify and revise their processes around these issues. (The WMP 'Data Lab' later clarified for the ethics committee that the only individuals profiled by the IOM tool were those with convictions and/or charges, but upon whom there might also be police intelligence fed into the model (WMPCC, 2019d).) Tom McNeil, a strategic advisor to the West Midlands Police and Crime Commissioner and driving force behind the establishment of the ethics committee concerned, was quoted in a *Guardian* piece as saying (Marsh, 2019): "The robust advice and feedback of the ethics committee shows it is doing what it was designed to do. The committee is there to independently scrutinise and challenge West Midlands police and make recommendations to the police and crime commissioner and chief constable… This is an important area of work, that is why it is right that it is properly scrutinised and those details are made public."

There are other models of regulation emerging in the United States that local governments in the UK might admire and aspire toward if they were very wary of predictive policing practices, but politically and constitutionally might struggle to develop. The City of Oakland in California has a Privacy Advisory Board (much as if West Midlands Police operated an ethics committee for data analytics in tandem with all local authorities in the West Midlands region) (City of Oakland, 2019), but local government and PCCs do not have the legal power in English law to fetter the discretion of chief constables to purchase something like facial recognition technologies - unlike in San Francisco where expenditure by the police on such tech is now more tightly controlled by local government officials due to privacy concerns (Paul, 2019). Rashida Richardson and her colleagues have also reported how New York City has founded a predictive technology decision-making committee; and that in Vermont an AI task force (and so more like an inquiry panel than an ethics committee) has been established to investigate ethical/unethical practices in the area of predictive policing (Richardson et al, 2019).

Meanwhile, some actors in the private sector are starting to realise that there might be a place in the market for ethical-only tech. Axon (formerly Taser) have declared publically that their body worn video camera products for law enforcement officers will not be compatible with automated or live facial recognition technology currently used or promoted for law enforcement use by other tech companies. Axon even have an ethics board, established with assistance from academics at US and UK universities, to enhance and define the company's public position on what it seeks to pursue, exactly, by developing ethical tech (Quach, 2019).

General legal points on algorithmic or predictive policing tools: and practical ramifications

The general common law basis of predictive policing data analysis tools to be used in the UK, such as the 'Concern Hub', can be acknowledged as lawful for the purposes of Article 8 ECHR, given the recent decision in *Catt v UK* (43514/15) (24 January 2019). The decision in

Catt by the European Court of Human Rights established the idea that while the UK legal framework concerned would tolerate the recording of textual intelligence data about an individual for interim analysis by the police, in order to make predictions operationally about risk, the longer the retention of that intelligence continues where an innocent, or not very harmful, person is concerned, the greater the likelihood that this retention will go on to be seen as disproportionate, and thus unlawful.

Key legal frameworks for predictive policing analytics tools are: the EU Data Protection Directive concerning law enforcement processing (2016), as transposed into Part 3 of the Data Protection Act 2018; the Public Sector Equality Duty (PSED) and the requisite 'due regard' standard for decision-making that must be cognisant of equalities issues, therefore, under the S. 149 Equality Act 2010 (in England and Wales, for example); as well as positive and negative obligations under the ECHR, taking effect through the duty on UK public bodies under S.6 HRA 1998; and the Rehabilitation of Offenders Act 1974, including its section 4 duty to treat a person with a spent conviction/caution 'for all purposes in law' as not committing the offence concerned. Particular challenges for police operators of algorithmic or predictive analytics tools that arise from this legal framework are outlined below.

These particular challenges relate to the legal intricacies of processing certain types of 'sensitive' personal data; the handling of victims' and witnesses' personal data; the exact degree of automation used by a predictive policing tool; the accuracy of a chosen model in terms of its preference for identifying all high risk individuals versus avoiding 'false positives'; the extent of data protection impact assessments as integrated with other impact assessments focusing on human rights and equality issues; the proportionality of decisions taken in an algorithmically-informed way; the extent of the transparency concerning the tool and public engagement and consultation over its development; and the way that spent convictions are, or are not, taken into account and/or processed in the model concerned. These challenges can then be grouped into the following categories: i) data scope issues; ii) process issues; iii) issues of human rights impact.

<u>i) Data scope issues</u>

The scope of the data retained and that is drawn upon by an algorithmic police intelligence tool (APIAT) must be data processed that is necessary for that purpose. Good practice might then be to determine which pieces of data about an individual 'nominal' are not predictively powerful, during the design process for a tool, and screen them out. The retention period for items of data drawn on by the tool is another factor to take into consideration. In terms of categories of sensitive personal data: if the use of some sensitive data (caught by section 24 of the Data Protection Act 2018) is implied and so the data cannot be readily removed from the model, then particular care should be taken to justify the necessity of the use of that point of data in the model in the requisite data protection impact assessment (see below), e.g. 'health data' if a person in the model is a victim of violence, or the political beliefs of a person if they are described in arrest data or intelligence as having been arrested at a particular protest.

Importantly, the results or outputs of a tool that draws on intelligence reports, rather than hard legal 'facts' such as records of arrests or charges or convictions will need to be presented to officers that use the tool in such a way that indicates that, in the language of Section 34 of the DPA 2018, these outputs are at least partly based on personal assessments" of other officers earlier in time. Likewise, the outputs of an APIAT will need to be similarly flagged if they draw upon victims' and witnesses' data as when "processing personal data for… law

enforcement purposes, a clear distinction must, where relevant and as far as possible, be made between personal data relating to different categories of data subject".

There are also issues with determining the exact application of the Rehabilitation of Offenders Act 1974, concerning spent convictions and cautions. But in some good news for police forces in the UK, the application of the 1974 Act may not actually be too problematic in the context of working with APIATs, assuming the validity of one judicial interpretation of the 1974 Act. Importantly, the s.4 duty to treat a person with a spent conviction/caution 'for all purposes in law' as though they have not committed a particular offence does not always include public protection purposes as 'legal purposes', following *N v Governor of HMP Dartmoor* [2001] EWHC Admin 93. So an APIAT with clear predictive purposes in the public protection field might benefit from this judge-made exemption from a statutory framework for spent convictions.

ii) Process issues

Section 64 of the Data Protection Act 2018 provides that a data protection impact assessment is undertaken by a body considering deploying an APIAT before any processing of data of 'nominals' takes place, and where " a type of processing is likely to result in a high risk to the rights and freedoms of individuals". In essence, this might mean that an integrated impacts assessment process is required under the combined measures of the DPA 2018, the Human Rights Act 1998 and the Equality Act 2010. (MOPAC produced an 'integrated impact assessment' (MOPAC, 2018b) in addition to its internal review of the operation of the Gangs Matrix over a five-year period, but this did not explicitly incorporate particular human rights as themes in the assessment process in the way that it did with 'protected characteristics' in it equality strand, although human rights-related values of protecting communities from harm while improving their cohesion and health were included.)

In terms of a requisite degree of public engagement and transparency over the inception, development and deployment of an APIAT, the public sector equality duty might require evidence gathering from communities that would disproportionately be engaged through the use of the planned APIAT, and certainly it is the view of notable barrister Timothy Pitt-Payne QC that human rights standards of 'accessibility' of legal information now require that there is public engagement over predictive policing issues, chiefly in the release of information notices to the public concerning the relevant APIAT (Pitt-Payne, 2018). There is an advantage to such an approach, in relation to what is known as the common law 'duty to give reasons', since in this context, the degree of information that has to be provided to those individuals affected by decisions informed by the APIAT should be underpinned, and rendered more clearly lawful, through the notification of communities about the use of the APIAT 'up front'.

A vital consideration in the deployment and use of an APIAT is the extent to which there is automation of police operational decision-making following the onset of the use of the tool concerned, in a particular way. Sections 49 and 50 of the DPA 2018 between them give a particular set of safeguards in connection with a fully automated decisions that would have an impact on the rights of an individual in the criminal justice context - firstly a person is to be informed of the fully-automated decision about them, and secondly, should they then take this opportunity to object, then the automated decision concerned will need to be re-made by a human officer and decision-maker. Automated decision-making is avoided through there being a 'human in the loop' - and so a careful and detailed process map of sorts might really help inform a future and more detailed stance by a police organisation on this issue, when

determining the nature of a partial intervention with a 'nominal' individual based on a fully-, partly- or initially-automated decision that draws on the outputs of an APIAT.

Furthermore, a key issue in process terms is the way that an APIAT is developed with a particular emphasis on the 'trade-off' sought between two types of accuracy that can be sought in the chosen predictive model: statistical sensitivity (which *in extremis* would make a tool as strong as possible at predicting high risk offenders, for example, but with a high 'false positive' rate down the line) or statistical specificity (an emphasis in the trade off toward the desire to correctly sort low-, medium- and high-risk offenders, with a correspondingly lower 'false positive' rate for its outputs - albeit with more 'false negatives', or high-risk offenders 'missed'). With regard to the statutory bar on the processing of data in ways that is 'misleading', given the language of S.205 DPA 2018, a careful focus on statistical *specificity* over *sensitivity* is to be applauded, as it will clearly in most instances be less 'misleading' to use a model weighted toward the former (Grace 2019).

iii) Issues of human rights impacts

Issues of human rights impacts typically will boil down to the application of a 'proportionality analysis', in addressing the interference by the processing of data on an individual through a machine learning-based tool on Article 8 ECHR, that is to say, the right to respect for private life under the Convention. Given that the processing of most data by such a tool will be confidential data privy to criminal justice bodies and partner agencies, as opposed to SOCMINT, and so not readily in the public domain, these will be data that if processed, will typically give rise to a 'reasonable expectation of privacy' on the part of an individual or 'nominal'. As such, UK common law in interpreting the Convention requires there is an overall *fair balance* "between the rights of the individual and the interests of the community" (as per *R. (on the application of Quila) v Secretary of State for the Home Department* [2011] UKSC 45). Following the now-standard approach from the UK courts to this proportionality analysis, what will help determine a truly 'fair balance' will be the extent of a rational basis for the processing overall, and whether the processing is 'no more than necessary' to achieve an objective of sufficient importance. Other 'qualified' rights under the ECHR, such as freedom of expression, will also instigate such a proportionality analysis if they are engaged; while the right to freedom from discrimination in the enjoyment of the right to respect for private life (engaged if the decision-making supported by an algorithm leads to indirectly discriminatory outcomes, for example) is violated should there be a pattern of discriminatory data governance practice that is manifestly without reasonable foundation (Grace, 2019; Raine, 2016).

Conclusions

There is clearly a developing narrative concerning the police practice of developing APIATs across the UK. The public, the media, and certainly human rights NGOs are demanding greater transparency over the concentration and use of such tools. Perhaps a fundamental problem is the common law basis of the creation and operation of (algorithmic) intelligence databases, combined with the generalist underpinning of Part 3 of the Data Protection Act 2018 only. That largely common law basis means that there has yet to be any specific Parliamentary authorisation in statute of machine learning-based APIATs in the field of police practice in the UK, despite the number of forces now wrangling with the legitimacy, both legal and democratic, of these technologies and their ramifications and opportunities.

In their letter to the *Guardian* newspaper in April 2019, Prof. Ruth Gilbert and Matthew Jay (2019) argued that "A wider public debate needs to be informed by research into how effectively the use of people's data predicts and reduces criminality, who else experiences targeting and privacy intrusion due to prediction errors, and whether better use of data could reduce such collateral harm… The public has a right to know how data about them is being used." Public legal education on an issue is well-complemented by the deliberation by Parliament over a new statutory framework for a controversial trend in the justice sector. Such a statutory position would be an opportunity to clarification in the law, based on a platform of more public, and more rigorous, discussion of the issues in policy circles.

As such, acknowledgment that a new predictive tool is potentially a new, more intensive form of police 'dataveillance' (Clarke, 1997) in a city or region would be best reflected in the legislature debating a Predictive Policing (Technology) Bill. In many ways, one thing such a Bill could do is set up a regulator/authorisation process via the courts, as per surveillance warrants under the Investigatory Powers Act 2016. The nature of APIATs focused on supporting offender management processes are not particularly secretive, and secrecy only lends them an ominous air that could be reduced (and the ability of citizens to understand and challenge them) if their process of development were also more commonly scrutinised in the open. A statutory authorisation process for such APIATs could then take (the lack of) such transparency into account in the designated method of giving approval for new tech, or new purposes for analytics in data-driven policing.

References

Alan Turing Institute and Independent Digital Ethics Panel for Policing, *Ethics Advisory Report for West Midlands Police*, 28th July 2017.

Alston, P. (2019), *Visit to the United Kingdom of Great Britain and Northern Ireland: Report of the Special Rapporteur on extreme poverty and human rights*, United National Human Rights Council, 23 April 2019.

Baraniuk, C. (2018) 'EXCLUSIVE: UK police wants AI to stop violent crime before it happens', *New Scientist*, 26th November 2018, from https://institutions.newscientist.com/article/2186512-exclusive-uk-police-wants-ai-to-stop-violent-crime-before-it-happens/

Boden, M. A. *Artificial Intelligence: A Very Short Introduction*, OUP: London, 2018, 2nd edition

City of Oakland, *Privacy Advisory Commission*, from https://www.oaklandca.gov/boards-commissions/privacy-advisory-board (accessed 02.07.2019)

Clarke, R. (1997), 'Introduction to Dataveillance and Information Privacy, and Definitions of Terms', from: <http://www.rogerclarke.com/DV/Intro.html> accessed 25th April 2019.

Crisp, W. (2019) 'Concern hub: New Metropolitan Police gang database sparks privacy and profiling fears', 13th March 2019, from https://www.independent.co.uk/news/uk/crime/concern-hub-metropolitan-police-gang-matrix-database-a8812371.html

Gilbert, R. and Jay, M. (2019) 'We need debate on data-driven policing', *The Guardian*, Tuesday 23rd April 2019, from: https://www.theguardian.com/uk-news/2019/apr/23/we-need-debate-on-data-driven-policing

Grace, J. (2017) 'Countering extremism and recording dissent: Intelligence analysis and the Prevent agenda in UK Higher Education' (2017), *Journal of Information Rights, Policy and Practice* Vol. 2(2) (online)

Grace, J. (2019) 'Algorithmic impropriety in UK policing?', *Journal of Information Rights, Policy and Practice*; Vol, 3 Issue 1, from: https://jirpp.winchesteruniversitypress.org/articles/abstract/23/

John Haigh, *Probability: A Very Short Introduction*, OUP: London, 2012

Hildebrandt, M. (2009) 'Profiling and AmI', in Rannenberg, K., Royer, D.and Deuker, A. (eds.), *The Future of Identity in the Information Society: Challenges and Opportunities*, 2009, Springer.

Holmes, D. E. *Big Data: A Very Short Introduction*, OUP: London, 2017.

Information Commissioner's Office (2018), *Enforcement Notice - Data Protection Act 1998 - Metropolitan Police Service*, from: https://ico.org.uk/action-weve-taken/enforcement/metropolitan-police-service/

Information Commissioner's Office (ICO) (2019a), 'Processing gangs information: a checklist for police forces', from: https://ico.org.uk/for-organisations/in-your-sector/police-justice/processing-gangs-information-a-checklist-for-police-forces/

Information Commissioner's Office (ICO) (2019b), 'London council fined by the ICO for disclosing sensitive personal data about alleged gang members', from: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/04/london-council-fined-by-the-ico-for-disclosing-sensitive-personal-data-about-alleged-gang-members/

Keats Citron, D., and Pasquale, F. (2014) 'The scored society: due process for automated predictions', *Wash. L. Rev.* 89 (2014): 1.

Marsh, S. (2019) 'Ethics committee raises alarm over 'predictive policing' tool', *The Guardian*, Saturday 20th April 2019, from https://www.theguardian.com/uk-news/2019/apr/20/predictive-policing-tool-could-entrench-bias-ethics-committee-warns

Mayor's Office for Policing and Crime (MOPAC) (2018), *Review of the Metropolitan Police Service Gangs Matrix*, December 2018, from: https://www.london.gov.uk/sites/default/files/gangs_matrix_review_-_final.pdf

Mayor's Office for Policing and Crime (MOPAC) (2018b), *Gangs Matrix Review - Integrated Impact Assessment*, December 2018

National Analytics Solution Project Team, *Response to the Alan Turing Institute and IDEPP*, 2017.

O'Neil, C. *Weapons of Maths Destruction: How Big Data Increases Inequality and Threatens Democracy*, 2009.

Paul, K. (2019) 'San Francisco is the first US city to ban police use of facial recognition tech', The Guardian, Wednesday 15th May 2019, from: https://www.theguardian.com/us-news/2019/may/14/san-francisco-facial-recognition-police-ban (accessed at 02.07.2019)

Quach, K. (2019) 'US cop body cam maker says it won't ship face-recog tech in its kit? Due to ethics? Did we slip into a parallel universe?', *The Register*, 28th June 2019, from: https://www.theregister.co.uk/2019/06/28/axon_facial_recognition/

Richardson, R., Schultz, J. and Crawford, K. (2019), Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice (February 13, 2019). New York University Law Review Online, Forthcoming. Available at SSRN: https://ssrn.com/abstract=3333423

Pitt-Payne, T. (2018) 'Appendix 1 - Legal Assessment', from the Mayor's Office for Policing and Crime, *Review of the Metropolitan Police Service Gangs Matrix*, December 2018, from: https://www.london.gov.uk/sites/default/files/gangs_matrix_review_-_final.pdf

West Midlands Police and Crime Commissioner, (2019a): Ethics Committee, from https://www.westmidlands-pcc.gov.uk/transparency/ethics-committee

West Midlands Police and Crime Commissioner, (2019b), Ethics Committee briefing note, from: https://www.westmidlands-pcc.gov.uk/media/514522/Ethics-Committee-03042019-IOM-MODEL.pdf

West Midlands Police and Crime Commissioner, (2019c), Ethics Committee minutes, April 2019, from: https://www.westmidlands-pcc.gov.uk/archive/april-2019/

West Midlands Police and Crime Commissioner, (2019d), Ethics Committee minutes, July 2019, from: https://www.westmidlands-pcc.gov.uk/archive/ethics-committee-meeting-july-2019/

Raine, T (2016) 'The Value of Article 14 ECHR: The Supreme Court and the "Bedroom Tax"' U.K. Const. Law Blog, available at https://ukconstitutionallaw.org/) (accessed at 28.07.2018)

Winsor, T (2018), *State of Policing - The Annual Assessment of Policing in England and Wales 2017*, HMICFRS: London.

Yeung, P. (2019) 'The grim reality of life under Gangs Matrix, London's controversial predictive policing tool', *Wired*, 2nd April 2019, from https://www.wired.co.uk/article/gangs-matrix-violence-london-predictive-policing