CHAPTER 19

ALGORITHMS IN THE JUSTICE SYSTEM: CURRENT PRACTICES, LEGAL AND ETHICAL CHALLENGES

Sophia Adams Bhatti and Holli Sargeant

TABLE OF CONTENTS

I. INTRODUCTION

If we do not maintain justice, justice will not maintain us (Francis Bacon) **19-001**

Algorithms play an increasingly important role in all aspects of society—from simple choice selection in retail decisions to fundamental pillars of our democratic societies. The potential to create net benefits, as well as long-term harms, co-exists with a technology that has immense computing power, which is general purpose in its design and application, and which is global and operates at scale. The pace of change in technology that is at the disposal of law enforcement is considerable. In the past two years, there has been a rapid emergence of new generative AI systems and large language models, like GPT. Similar to algorithmic tools already being deployed, generative AI may also offer potential benefits to the criminal justice system but poses a serious risk of exacerbating existing social and legal challenges.

Concerns of potential unfairness, discrimination and opacity are relevant in many applications of AI but are profoundly problematic when set within the context of the justice system. As a fundamental pillar of a well-functioning democracy, complacency in how our justice system evolves is likely to set a course for the widening of the democratic deficit to a chasm- like scale. It is for this reason that it is vital that the concerns and potential benefits are carefully examined, understood and explained.

In the main, the response from governments globally has been to allow the market to lead the way by applying a basic free market set of principles, in the hope that the supply and demand equation will even out to deliver high quality and trusted AI-based tools.

Increasingly,[1] it is being acknowledged that this laissez-faire approach is not go- **19-002**

[1] There are many examples of increasing concern about the potential harms caused by AI tools used poorly in the justice system. A high profile example includes this statement by leading civil rights activists in the US at *https://civilrights.org/2018/07/30/more-than-100-civil-rights-digital-justice-and-community-based-organizations-raise-concerns-about-pretrial-risk-assessment/*.

ing to deliver the outcomes which society is willing to accept. Despite all the hype, and the constant overuse of the concept that AI has all the answers, that it can gaze into the future with pinpoint accuracy, and that it is free of error—for those who understand the tools available these myths are often far from the truth of what is plausible and accurate. In many cases, there are also trade-offs to be made. These trade-offs are ones which engage important and deep-rooted values. And yet, government policies seem to be happy with allowing these trade-offs—such as right to a fair trial versus efficiency, or accuracy versus efficiency—to be bartered through the lens of economics alone.

In many deployments of AI in criminal justice, there is a reliance on the purchasing of tools developed by private companies whose driving force is not the rule of law, but rather efficiency of market solutions which can deliver profits. For many developers, the very concept of the rule of law will be alien to the development process. Profitability can be in direct conflict with tenants such as openness and transparency, or even rigorous testing. In the case of *State of Wisconsin v Loomis*,[2] this tension is clearly seen, with the Wisconsin Supreme Court ruling against the defendant's right to examine the underlying algorithm used in the case against him. The implications of such decisions speak directly to Frank Pasquale's concerns about the "social implications of these invisible practices".[3] The duality of the very notion of the "black box" is very relevant in these settings—instead of recording every decision and explanation, with the aim of delivering clarity, transparency and knowledge, the emerging "black box society" approach in criminal justice is one lacking in all these traits.

19-003 Obvious conflicts such as market dynamics aside, the use of AI in the field of criminal justice raises other concerns. These include the biases in the data used, the continuation of discriminatory practices, the false legitimacy awarded to decisions by virtue of being "statistical", lack of skills and understanding to deploy the tools appropriately, human rights tensions, privacy issues and social licence.

Does this mean that there is no room in the system for such tools? That rather depends on mitigating the significant risks and issues associated with their use. Important measures include having suitable standards of care and regulatory thresholds (such as the ability to separate good data from bad[4]); having clearer standards and certification of quality; and increasing training and competency of law enforcement personnel, judges, parole officers, legal advisers and other participants. However, currently, there is no concerted effort to implement these mitigations.

II. IDOLATRY OF DATA AND AI

19-004 The use of AI in criminal justice is varied and continuing unabated. This is happening notwithstanding thematic and recurrent concerns expressed in numerous studies, reports and investigative journalism. It seems that "the allure of the technol-

---

[2] *State of Wisconsin v Loomis* 881 N.W.2d 749 (Wis. 2016). The defendant, Eric Loomis filed an appeal to his sentencing based on due process concerns due to the lack of opportunity to examine the AI tool used to assess his "risk" level. The court found against Loomis, arguing that it did not believe that the inability to examine the "black box" led to a lack of due process. See *https://www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690*.

[3] F. Pasquale, *The Black Box Society, the secret algorithms that control money and information* (Harvard University Press, 2015).

[4] P. Jeffrey Brantingham, "The Logic of Data Bias and Its Impact on Place-Based Predictive Policing" (2018) 15 Ohio St. J. Crim L. 473, 485.

ogy is clear—the ancient aspiration to predict the future, tempered with a modern twist of statistical sobriety"[5] is too strong for many in the criminal justice system to resist.

The rise of AI in criminal justice is not an isolated development, but rather is set within the context of a much broader growth of AI across many elements of our daily lives. These systems are built on foundational layers of data, and set within a growing belief, amongst some, that data is everything. Yuval Noah Harari describes such "Dataism" as a belief that "the universe consists of data flows, and the value of any phenomenon or entity is determined by its contributions to data processing".[6] This is not a niche or fringe notion, but one which is gaining mainstream support. Meanwhile, the allied concept of *datafication* refers to the quantification of human life through digital information, very often for economic value, not just the making of information.[7] This process has major social consequences. Its appeal lies in the "offer", the *promise*, of a world in which things which hitherto have been too hard to decipher and understand by the human mind, can be unlocked, in an empirical data driven way, one might even go so far as to draw a resonance with the traditional ontological arguments made by Anselm in *Proslogian*.[8] This idolisation of data, of the algorithm, entices and encourages the users to "trust" the data and the algorithms in an unquestioning way, deferentially adhering to its outputs. But the promise is just that—an "I.O.U" for a potential yet to be reached. Whilst it is true that there are AI breakthroughs almost on a daily basis—whether it be detecting heart disease,[9] helping in pre- and post-natural disaster events,[10] detecting wildlife poaching[11]—there is still plenty of room for improvement. Whilst the idea that AI might be able to one day help to accurately predict who will or will not reoffend is hugely appealing to some, the truth is that the AI tools on the market today are far from omniscient or reliable. There are deep flaws in the data, application and understanding of these technologies.

Criminal justice as a sector is not unique in offering problematic trade-offs, but it does offer a useful case study of the implications of AI in an acute setting. One within which the wrong decisions can lead to individual and collective human rights infringements, but also importantly encourage the slow corrosion of society's trust in the system itself. As Professor Sylvie Delacroix argues, the justice system as a whole, and the values we have come to associate with it, might be transformed beyond recognition by uncritical reliance on data-intensive technologies, and that transformation comes with high individual and collective stakes.[12] A parallel concern is that without rigour, standards, transparency or democratic licence these systems seep into the fabric and infrastructure of our justice system and corrode the trust which the system relies upon, without most people even knowing about it. And

---

5 Pasquale, *The Black Box Society, the secret algorithms that control money and information* (2015).

6 Y.N. Harari, *Homo Deus: A Brief History of Tomorrow* (Harvill Secker, 2016), p.428.

7 A.U. Mejias and N. Couldry, "Datafication" (2019) 8(4) *Internet Policy Review*, *https://policyreview.info/concepts/datafication*.

8 See *https://jasper-hopkins.info/proslogion.pdf*.

9 See *https://www.nature.com/articles/s41746-019-0130-0*.

10 See *https://www.forbes.com/sites/cognitiveworld/2019/03/15/how-ai-can-and-will-predict-disasters/*; *http://aidr.qcri.org*; *https://www.lexalytics.com/lexablog/artificial-intelligence-disaster-relief*.

11 See *https://www.resolve.ngo/trailguard.htm*.

12 S. Delacroix, "Computer Systems Fit for the Legal Profession?" (2018) *Legal Ethics*, *https://papers.ssrn.com/sol3/,papers.cfm?abstract_id=3158132*.

by the time the collective recognises the harms, it will be all too late to reverse or undo the damage.

19-005 A further consideration is the role of different actors in the system as a whole, and where responsibility ought to rest. Maintaining justice is not only about the decisions we take, but the maintenance of the justice *system* itself—and like all systems there are a variety of inputs and outputs. In criminal justice, there are a range of disciplines which play a part, and in this data driven context these now include a supply chain which previously we may not have seen as relevant to the rule of law. Responsibility spans public and private sectors, from the coders in Silicon Valley, to the police forces deploying tools locally, through to the judges—each has a role and a responsibility. To use the frame of economics, this duty, to protect the rule of law, is one which rests both at the supply and demand side of the equation.

III. How widespread is the use of AI in the criminal Justice system?

19-006 What is the extent of the use of AI in the UK criminal justice sector? How is it being used? Who is deploying it? In short, no one really knows. There have been numerous headlines and a few attempts to catalogue the position within specific instances, but nothing is complete or comprehensive. In this jurisdiction, this is because there is no requirement for police forces to disclose this information, no national register and no way of undertaking a systematic audit. As such, there are no definitive figures as to the number of forces using algorithmic tools in the United Kingdom (UK) and in what context. A similar lack of transparency is experienced in other jurisdictions, such as the United States (US). It is clear from public statements[13] that there is recognition that the lack of transparency causes some to have concerns about the impacts of deploying such technologies. That said, an appropriate legal policy response has yet to be put in place.

What is clear from the studies which have been conducted both in the UK and US, is that there is an upward trend in adoption of these tools. A 2016 study in the UK based on Freedom of Information (FOI) requests asked all police forces whether they used any sort of computational or algorithmic data analysis or decision making in relation to the analysis of intelligence, and to confirm the nature and purpose of any such algorithms. Of those which responded, 14% reported affirmatively.[14] The civil society organisation Liberty similarly made 90 FOI requests to police forces in 2018, with 14 returned affirmatively.[15] The Law Society of England and Wales Commission on the use of AI in Criminal Justice, of which the author of this chapter was the Director, sought to also map the known uses of AI across England

---

[13] Home Office (UK), "*Home Office—Written evidence (NTL0055)*" (2021), *https://committees.parliament.uk/writtenevidence/40975/pdf/*; and Forensic Science Regulator, "*Conference Questions and Answers*" (2023), *https://www.gov.uk/government/publications/forensic-science-regulator-2023-conference-questions-and-answers/forensic-science-regulator-2023-conference-questions-and-answers*.

[14] M. Oswald and J. Grace, "Intelligence, Policing and the Use of Algorithmic Analysis: A Freedom of Information-Based Study" (2016) 1(1) *Journal of Information Rights, Policy and Practice*, *http://doi.org/10.21039/irpandp.v1i1.16*.

[15] H. Couchman, *Policing by Machine: Predictive Policing and the Threat to our Rights* (Liberty, 2018).

and Wales. Without a necessity to report on use and deployment, any such attempts to map the real scale and nature of use is at best an approximation.[16]

What is also known is that algorithmic systems in the justice system are diverse and increasingly established. Algorithms, such as risk scoring systems, have long been used by public agencies[17] and considered in regulation.[18] Aside from these risk assessment and prediction tools,[19] there are many other applications of AI, ranging from photographic and video analysis, including facial recognition;[20] to DNA profiling;[21] predictive crime mapping;[22] mobile phone data extraction tools;[23] data mining and social media intelligence (SOCMINT).[24]

The expansion of these tools has been fuelled by a surge in machine learning capabilities; the exponential rise in data and computing power; and a favourable policy context (involving falling public financing for justice and policing, the changing nature of crime towards more digitally enabled practices, and perceptions about the role of policing). Expanding on the latter, although in countries such as the US[25] and the UK crime rates[26] continue to fall, perceptions and rhetoric of **19-007**

---

16 A map of algorithms in the justice system, sourced from the Law Society public website is at *https://www.lawsociety.org.uk/topics/research/mapping-algorithms-in-the-justice-system*.

17 See generally M. Veale and I. Brass, "Administration by Algorithm? Public Management Meets Public Sector Machine Learning" in K. Yeung and M. Lodge (eds), *Algorithmic Regulation* (Oxford: Oxford University Press, 2019), *https://academic.oup.com/book/35048/chapter-abstract/298946506?redirectedFrom=fulltext*; C. Coglianese and D. Lehr, "Regulating by Robot: Administrative Decision Making in the Machine-Learning Era" (2017) 105 Geo. L.J. 1147.

18 They have also been the subject of statutory provisions in the UK. See, e.g. the former Data Protection Act 1998, which contained exemptions for risk assessment systems applied to individuals for the purposes of crime prevention or the levying of taxes. Data Protection Act 1998 s.29(4)(a); a similar provision also exists in the current Data Protection Act 2018 Sch.2 para.3(2).

19 See e.g. R. Moore (ed), *A Compendium of Research and Analysis on the Offender Assessment System* (Ministry of Justice Analytical Series, 2015), *https://perma.cc/W2FT-NFWZ*; J.P. Singh, D.G. Kroener, J.S. Wormith, S.L. Desmarais and Z. Hamilton Z (eds), *Handbook of Recidivism Risk/Needs Assessment Tools* (Wiley Blackwell, 2018).

20 *R. (on the application of Bridges) v Chief Constable of South Wales* [2019] EWHC 2341 (Admin); [2020] 1 Cr. App. R. 3. See *https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf*.

21 A.O. Amankwaa and C. McCartney, "The UK National DNA Database: Implementation of the Protection of Freedoms Act 2012" (2018) 284 *Forensic Science International* 117.

22 S.D. Johnson, D.J. Birks, L. McLaughlin, K.J. Bowers and K. Pease, "*Prospective Crime Mapping in Operational Context: Final Report*" (Home Office Online Report, 19 July 2007); L.W. Perry, B. McInnis, C.C. Price, S. Smith and J.S. Hollywood, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (Santa Monica, CA: RAND Corp, 2013), *https://www.rand.org/pubs/research_reports/RR233.html*.

23 Privacy International, "*Digital stop and search: how the UK police can secretly download everything from your mobile phone*" (March 2018).

24 See generally L. Edwards and L. Urquhart, "Privacy in Public Spaces: What Expectations of Privacy Do We Have in Social Media Intelligence?" (2016) 24 *International Journal of Law and Information Technology* 279.

25 See *https://www.fbi.gov/news/pressrel/press-releases/2019-preliminary-semiannual-crime-statistics-overview*.

26 See the latest crime statistics for England and Wales, which show a stead decrease in crime over the past 20 years, *https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingmarch2020#overall-estimates-of-crime*, the latest ONS release on public perceptions in 2016 showed a marked inconsistency between actual crime at a national level and perceived levels of national crime. There was a more realistic view of crime levels when respondents were asked about their local area, however, see *https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/articles/publicperceptionsofcrimeinenglandandwales/yearendingmarch2016*.

increasing risk support a more invasive form of detection and prevention. These perceptions are not countered by an informed and open debate. This provides for, in the short term, fertile ground for a sense of legitimised deployment of tools, some of which at best can be described as experimental. As found by Alexander Babuta and Marion Oswald:[27]

> "the development of policing algorithms is often not underpinned by a robust empirical evidence base regarding their claimed benefits, scientific validity or cost effectiveness."

Overall, the quasi-secretive nature of deployment raises significant concerns, and it could be argued that it sets a rot in the system as a whole. If we expect our justice systems to command the support and confidence of citizens and society as a whole, it is unwise to press ahead without proper discussion of the trade-offs and the principles used to make decisions about these technologies.

19-008 As is discussed in the next section, each tool comes with a range of challenges (technical and legal—from human rights implications, data protection issues and ethical concerns). Yet, there are no standards, no guidance or transparency to help ensure that these issues are navigated successfully.

IV. The tools deployed

19-009 The range of AI-based tools in the criminal justice system has been steadily growing in recent years, fuelled by the advent of a surge in machine learning capabilities, the exponential rise in data and computing power and set with the policy context described above. This section focuses on the use of predictive tools and facial recognition systems, as two of the most widely used and growing areas of AI deployments and as areas which provide cross-cutting thematic lessons for the deployment of any tools.

(1) Predictive tools

19-010 There are many definitions of predictive policing, although all share a commonality of essentially being based on data analytics in some form. The Rand Corporation defines predictive policing as:

> "The application of analytic techniques—particularly quantitative techniques—to identify likely targets for police intervention and prevent crime or solve past crimes by making statistical predictions."[28]

The real life applications of predictive tools in criminal justice have mostly focused on two applications: predictive crime mapping; and individual risk assessment tools.

---

[27] A. Babuta and M. Oswald, "Data Analytics and Algorithms in Policing in England and Wales: Towards A New Policy Framework" *RUSI Occasional Paper*, February 2020.

[28] Perry, McInnis, Price, Smith and Hollywood, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (2013), *https://www.rand.org/pubs/research_reports/RR233.html*.

(a) Predictive crime mapping

19-011 Predictive crime mapping seeks to forecast where crime is most likely to occur, using statistical methods to identify future risk areas, it is combined with the theory that repeat "victimisation" represents a large proportion of crime. Research from the early 2000s[29] examined aspects such as the "contagion" effect in relation to burglary, in which researchers found an increased risk to nearby properties in the aftermath of an offence. In the case of repeat victimisation, two theories play out—the first being that victims of crime are temporarily at higher risk of crime than non-victims because of returning offenders, or secondly, that it is a result of vulnerable individuals being "flagged" as easier targets.[30] This work is not new, the research in the field of victimisation stems back decades, for instance the Kirkholt project in the 1980s.[31] What is novel in the latest iteration is the data available and the computing power to undertake the analysis.

Ultimately, the objective of such techniques is to move policing from a retrospective mapping of past crimes to a "prospective hot spot"[32] approach. The intention being to inform future police interventions to increase their likelihood of success, such as through pre-emptive police patrols.[33] These technologies are supported by many and are now widespread in both the US and the UK.[34] Given the financial pressures, incentives are aligned to pursue wider and faster uses of such approaches.

However, despite the prima facie benefits of such systems, there are well-founded concerns about the appropriateness of some of these tools, including the lack of transparency from both police departments and private firms regarding how predictive policing models are built; how police forces utilise their data; how reliable the data is, and whether the programs unnecessarily target specific groups more than others. I expand on some of these concerns below.

(i) Bad data

19-012 In basic terms, data is everything in these systems—both the key to success and the Achilles heel. Concerns have been raised about the sorts of data used to train the systems, with some police forces accused of developing tools which integrate information with no proven relationship to crime, such as weather or other socio demographic indicators.[35]

Richardson et al[36] point to, in the US context, deeper flaws in the data being used:

---

29 K.J. Bowers, S.D. Johnson and K. Pease, "Prospective Hot-Spotting: The Future of Crime Mapping?" (2004) 44 *British Journal of Criminology* 641; K.J. Bowers and S.D. Johnson, "Domestic Burglary Repeats and Space–Time Clusters: The Dimensions of Risk" (2004) *European Journal of Criminology* 67–92.

30 K. Pease, "*Repeat Victimisation: Taking Stock*" (Home Office, 1998).

31 See *https://popcenter.asu.edu/sites/default/files/problems/burglary_home/PDFs/Forrester_Chatterton&Pease_1988.pdf*.

32 Bowers, Johnson and Pease, "Prospective Hot-Spotting: The Future of Crime Mapping?" (2004) 44 *British Journal of Criminology* 641.

33 Johnson, Birks, McLaughlin, Bowers and Pease, "*Prospective Crime Mapping in Operational Context: Final Report*" (Home Office Online Report, 19 July 2007).

34 Babuta and Oswald, "Data Analytics and Algorithms in Policing in England and Wales: Towards A New Policy Framework" *RUSI Occasional Paper*, February 2020.

35 *HunchLab: Under the Hood* (Azavea, 2015).

36 R. Richardson, J. M. Schultz and K. Crawford, "Dirty Data, Bad Predictions: How Civil Rights

> "[T]hese systems are built on data produced during documented periods of flawed, racially biased, and sometimes unlawful practices and policies ('dirty policing'). These policing practices and policies shape the environment and the methodology by which data is created, which raises the risk of creating inaccurate, skewed, or systemically biased data ('dirty data')."

A reinforcement of bad practices, by using as a reference point bad data, is likely to see the legacies of the past repeated but with a veil of objectivity attached. In jurisdictions where there is a highly complex legacy concern, this amounts not to a small glitch but a fundamental flaw. In the context of the US,[37] research proves that the police database is neither accurate nor random, and that there is strong evidence of discriminatory decision-making based on race and ethnicity in relation to issues such as detention. Against this backdrop it is no wonder that Richardson et al conclude that:

> "the use of predictive policing must be treated with high levels of caution and mechanisms for the public to know, assess, and reject such systems are imperative."

Notably it is not clear from government actions in the UK or the US that these warnings are being heeded. Supporters of the use of such data have argued that it must be plausible to "clean" the data and remove the bias which exists within it, however, there has been little evidence that the vendors developing and selling these tools have been able to do so.

(ii) Over policing

19-013 One of the impacts of this dirty data problem, or the "bias in bias out" problem, is the risk of over policing. The effects of new predictive policing in over-policed areas became especially prominent during the Black Lives Matter protests in the US.[38] The Black Lives Matter protests prompted many police departments to expand their use of predictive policing technologies like risk assessment algorithms, in an attempt to demonstrate increased professionalisation and objectivity. However, these new predictive policing tools often relied on historically biased crime data, leading to the over-surveillance of minority communities and peaceful protestors, despite claims that the technology would reduce racial bias in policing.

Liberty has been instrumental in identifying impacts of new technology in UK law enforcement. For example, Liberty's written evidence to the Justice and Home Affairs Committee's inquiry into the use of new technologies in law enforcement highlighted that "the Home Office had quietly rolled out new GPS technology for the electronic monitoring of people on immigration bail".[39] Furthermore, in

---

Violations Impact Police Data, Predictive Policing Systems, and Justice" (2019) *New York University Law Review* 192–233.

37 K. Lum and W. Isaac, "To Predict and Serve" *Significance Magazine*, October 2016, *https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x*.

38 E. Heh and J. Wainwright, "No privacy, no peace: Urban surveillance and the Movement for Black Lives" (2022) 3(2) *Journal of Race, Ethnicity and the City* 121–141, *https://doi.org/10.1080/26884674.2022.2061392*; N. Lally, "It makes almost no difference which algorithm you use: on the modularity of predictive policing" (2022) 43(9) *Urban Geograph* 1437–1455, *https://doi.org/10.1080/02723638.2021.1949142*.

39 "*Liberty*—Written evidence (NTL0020)" *Liberty*, (2021), *https://committees.parliament.uk/writtenevidence/38701/pdf/*.

Liberty's report "Policing by Machine",[40] concerns are raised about what the authors describe as "compound … unfair treatment of marginalised communities". The concern is that using historical crime data presents an unfair and inaccurate picture of crime committed in a particular area, but rather presents a picture of how crime was responded to by the police. Some crimes will not have been reported, other crimes may not have been followed up by the police, and for those areas where the police are regularly patrolling, they may be more likely to encounter more offences.

Liberty also make the important point that statistics might show reports of crimes, and arrests but arrest is not to be confused with guilt:

> "[I]n 2016/2017, black people were over three times more likely to be arrested than white people, and people of with a mixed ethnic background were over two times more likely to be arrested than white people, but without a corresponding rate of conviction."[41]

19-014 The report concludes that "predictive policing programs are better at predicting likely police involvement in a certain community—not potential crime".

A study undertaken in 2017[42] in the US, found that the systems used:

> "have been empirically shown to be susceptible to runaway feedback loops, where police are repeatedly sent back to the same neighbourhoods regardless of the true crime rate."

This is caused by the type of machine learning used in most cases, which the authors identify as "batch mode … where decisions made and observed results supplement the training data for the next batch". This creates an inherent feedback loop problem—for Area A where there is no police patrol there is likely to be less detected crime and therefore less data returned to the system, conversely Area B, with increased police patrols, is likely to return higher figures, reinforcing the decision to patrol Area B. In marginalised communities, which have long raised concerns about being the target of over policing, these tools threaten to exacerbate the problem.[43]

In their review,[44] Lum and Isaac sought to test the accuracy of drug crime data, and the results are a stark illustration of the point. They started by creating a dataset which combined, what they describe as, a "synthetic population … which is a demographically accurate individual-level representation of a real population—in this case, the residents of the city of Oakland" with survey data from the 2011 National Survey on Drug Use and Health. This combined data was then compared to the police records, showing "that the crimes known to the police were not a representative sample". Mapping these onto the map for Oakland—as shown in the figures in this chapter the authors demonstrate how wrong following the data alone could be.

19-015 Figure 1(a) shows the crimes recorded and Figure 1(b) the fuller picture of estimated drug use. For those unfamiliar with the Oakland area, the authors use-

---

40 Couchman, *Policing by Machine: Predictive Policing and the Threat to our Rights* (2018).

41 Couchman, *Policing by Machine: Predictive Policing and the Threat to our Rights* (2018).

42 D. Ensign, S.A. Friedler, S. Neville, C. Scheidegger and S. Venkatasubramanian, "Runaway Feedback Loops in Predictive Policing" 81 *Proceedings of Machine Learning Research* 1–12, 201, Conference on Fairness, Accountability, and Transparency.

43 N. Rossback, "Innocent until Predicted Guilty: How Premature Predictive Policing Can Lead to a Self-Fulfilling Prophecy of Juvenile Delinquency" (2023) *Florida Law Review* 167.

44 K. Lum and W. Isaac, "To Predict and Serve" *Significance Magazine* October 2016, *https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x*.

fully highlight that the areas of concentration in Figure (1) correlate with two areas with largely non-white and low-income populations. The research by the team concluded that these areas experienced approximately 200 times more drug-related arrests than areas outside of these clusters, despite the study showing that drug crimes were more evenly spread across the city. The authors state:

> "this suggests that while drug crimes exist everywhere, drug arrests tend to only occur in very specific locations—the police data appear to disproportionately represent crimes committed in areas with higher populations of non-white and low-income residents."

The knock on effects of using such data to predict and direct policing in the future is obvious and troubling.

19-016 Lum and Isaac go further to test how using a predictive tool, based on the Oakland Police data would perform. Using PredPol[45] as their test, the authors found that, despite its claims to be robust and neutral, the results were to amplify the biases in the existing data. The locations identified by the system were those already identified as over being over-represented in the police data. Because of this bias, the results showed that:

> "using PredPol in Oakland, black people would be targeted by predictive policing at roughly twice the rate of whites. Individuals classified as a race other than white or black would receive targeted policing at a rate 1.5 times that of whites"[46]

despite the data showing that drug use was evenly spread across the population.

Mapping hotspots, directing police resources to the areas of most need remains an easy and appealing message. Even more so, when this is coupled with the evidence that random police allocation is not hugely effective, and that the correct use of predictive policing has been shown to reduce crime—it leads to, for some, a compelling narrative to push for greater expansive use of these tools. However, many of the concerns highlighted point towards troubling flaws which inevitably, if left unchecked will make our communities less safe, not more. The veneer of statistics and technology compound the danger as users' deference to the results weakens the norm response to question and challenge results. Although human decision making is also not free from bias, the "black box" of some of the systems deployed make the ability to question, challenge and re-align to account for bias incredibly difficult.

---

[45] PredPol is a leading Predictive Policing tool used by police forces.

[46] K. Lum and W. Isaac, "To Predict and Serve" *Significance Magazine*, October 2016, *https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x*, p.18.

19-017

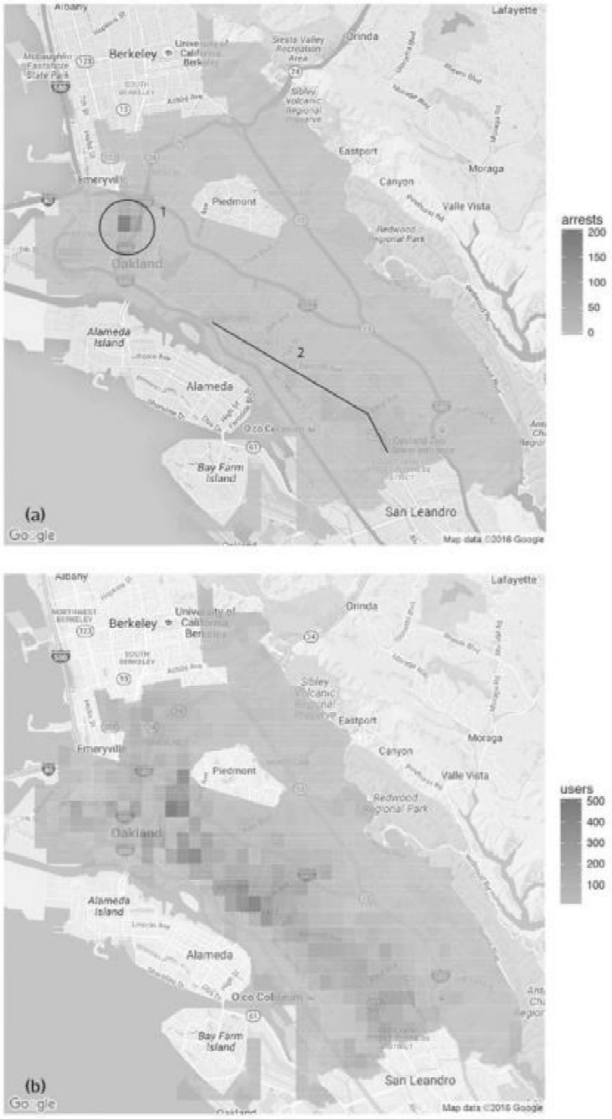

Figure 1: (a) Number of drug arrests made by Oakland police department 2010. (1) West Oakland, (2) International Boulevard. (b) Estimated number of drug users, based on 2011 National Survey on Drug Use and Health

(b) Individual risk assessment

In the toolbox of predictive policing, individual risk assessment is the other growth area. Ranging from systems which focus on the point of arrest, systems **19-018**

which seek to identify suspects, systems for scoring and identifying at risks victims and systems used by the prison and probation services—there are a whole host of tools deployed, some developed in-house, some purchased from private sector vendors.

Much attention has focused on those systems which are designed to be used at the individual level to ascertain the risk of future criminal behaviour. Indicators, identified as being relevant to assessing the degree of risk, are identified and individuals are assessed against these frameworks. Such approaches are not new in policing, widely used by the prison and probation services in the UK to assess risks of re-offending and individual management plans.

Across the UK policing establishment there are many examples of the use of systems which seek to apply a "risk" assessment, based on data analytics, including those listed in a table published by Liberty in 2019 of the known uses of predictive policing programmes in the UK.[47]

Table 1: List of predictive policing programmes in the UK[48]

19-019

POLICE FORCE	PREDICTIVE MAPPING PROGRAMMES	INDIVIDUAL RISK ASSESSMENT PROGRAMMES
Avon and Somerset	X	X
Cheshire	X	
Durham		X
Dyfed Powys	X (in development)	
Greater Manchester Police	X	
Kent	X	
Lancashire	X	
Merseyside	X	
The Met	X	
Norfolk	X	
Northamptonshire	X	
Warwickshire and West Mercia	X (in development)	
West Midlands	X	X
West Yorkshire	X	

19-020 In addition, since the research by Liberty in 2019, many other examples have been found, including:

---

[47] L. Dencik, J. Redden, A. Hintz and H. Warne, "The 'golden view': data-driven governance in the scoring society" (2019) 8(2) *Internet Policy Review*, *https://policyreview.info/articles/analysis/golden-view-data-driven-governance-scoring-society*.

[48] As created by Liberty in 2019 and as referenced by Dencik, Redden, Hintz and Warne, "The 'golden view': data driven governance in the scoring society" (2019) 8(2) *Internet Policy Review*.

- Sussex police use a data analytics dashboard that track reported crimes, incidents, calls and arrests.[49]
- West Yorkshire Police uses an Integrated Offender Management (IOM) software system called Corvus IOM Case. The system draws data from other sources, including STORM and Niche RMS, analysing intelligence, crimes, arrests and substance misuse in order to derive an individualised score aimed at providing an indication of an individual's likelihood to re-offend.[50]
- The digital categorisation service (DCS) is an algorithmic tool used to make, record and justify prison security categorisation decisions. The tool "highlights risk information" and suggests an initial categorisation for the prisoner that is re-viewed and can be changed by the staff.[51]
- Courts can request Pre-Sentence Reports which may include Offender Group Reconviction Score, Risk of Serious Recidivism score and Risk of Serious Harm screening.[52]

One of the most widely used systems in the UK is the Offender Assessment System (OASys), which in some form dates back to the 1990s, and initially began life as a paper-based system. This produces scores for offenders using three statistical indicators,[53] at various points, such as pre-sentencing; at the start of a sentence; and at key decision points and reviews, including parole.

19-021 The more recent tools, and those now in development in the field, have sought to harness the potential of progress made in what many refer to as the second wave in machine learning. This includes:

(a) The use of "random forest" forecasting,[54] used by Durham Constabulary (in collaboration with the University of Cambridge) in the Harm Assessment Risk Tool (HART) system.[55] HART is discussed in greater detail at para.19-029.

(b) Harnessing the power of "big data" by connecting data sources across institutions,[56] as in the case of Avon and Somerset Police which uses statistical models software (including the Qlik Sense platform, and tools SPSS Modeler and ESRI Mapping) related to a suspect's future behaviour and

---

[49] Sussex Police, "*Freedom of Information ref 0162/21*" (2021), *https://perma.cc/8BAH-5VPK*

[50] West Yorkshire Police, "*National Data Analytics Solution (NDAS) Privacy Notice*", *https://www.westyorkshire.police.uk/advice/modern-slavery/nationaldata-analytics-solution-ndas-privacy-notice*

[51] HM Prison and Probation Service, "*Security Categorisation Policy Framework*" (2021), *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1011502/security-categorisation-pf.pdf*.

[52] National Offender Management Service, "*Determining Pre Sentence Reports- Sentencing within the new framework*" PI 04/2016 (2021), *https://www.gov.uk/government/publications/determining-pre-sentence-reports-pi-042016*.

[53] OASys General Predictor Score, which relates to non-sexual, non-violent offences; OASys Violence Predictor Score; and "*Offender Group Reconviction Scale*" (version 3).

[54] This aggregates outcomes over what is potentially a very large set of decision trees: see para.2-042 of this book.

[55] HART was developed by Durham Constabulary and the University of Cambridge. The central goal was to promote consistency in decision-making, enabling targeted interventions to find responses to offending that reduce future harm and recidivism.

[56] Babuta and Oswald, "Data Analytics and Algorithms in Policing in England and Wales: Towards A New Policy Framework" *RUSI Occasional Paper*, February 2020.

risks of perpetrating serious offences of domestic violence offence, sexual violence offence, or a burglary.[57] Around 250,000 potential (re-)offenders have been given a score in the system.[58]

(c) Using social media content, as in London Metropolitan Police knife crime system (which amongst other things draws on this material). South Wales Police are also exploring the use of AI to analyse the social media accounts of offenders, or potential offenders.

(d) Using AI and data science based systems to extract data and combine insights to improve existing risk assessment tools which lack accuracy, as in the case of Hampshire Police who are in the development process of improve their existing Domestic Abuse Stalking and Honour Based Violence (DASH)[59] model by deploying machine learning, see Figure 2 which illustrates the design.

19-022

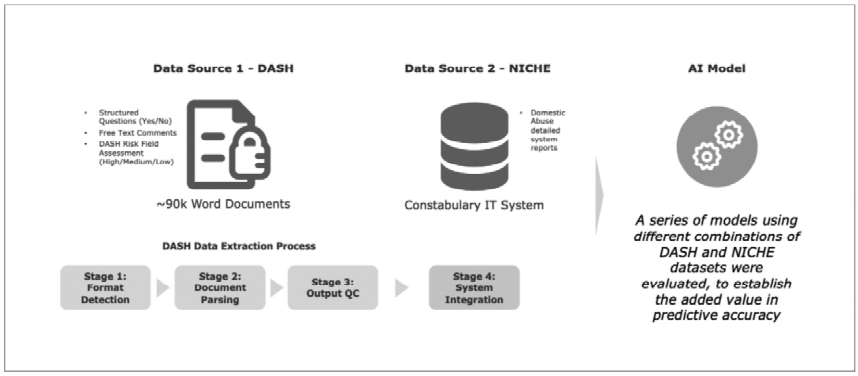

Figure 2—illustration of AI enhanced DASH model[60]

In this example, the AI system being developed is based on two datasets—information extracted data from DASH assessments and additional existing police records. The text from these datasets is translated into numeric format and the AI forecasting model learns to classify the risk of domestic violence re-offending, from multiple data points.[61]

19-023 Although not utilising AI or machine learning, the London Metropolitan Police "Gangs Matrix", is worth noting at this stage. A forecasting tool designed to identify individuals with a propensity for either being a gang member or engaging in gang-related activities. Although the Force states that it was a manual system, deep

---

57 Avon and Somerset Police, "*Avon and Somerset Police—Written evidence (NTL0052)*" (2021), *https://committees.parliament.uk/writtenevidence/40328/pdf/*.

58 L. Dencik, A. Hintz, J. Redden and H. Warne, *Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services* (Data Justice Lab, Cardiff University, 2018), 75.

59 O. Terzig and Rinik, "Shaping the state of machine learning algorithms in policing" *Workshop Report* (June 2019) June 2019.

60 Terzig and Rinik, "Shaping the state of machine learning algorithms in policing" *Workshop Report* (June 2019).

61 Example data points include Offender (age, gender, arrested, suspected, self-harm/suicide warnings); Victim (age, gender); Incident (serious, violence, sexual, offender count, victim count, first domestic occurrence, top 30 postcode—district level); Criminal record (Intimate Occurrences, Murder Offences, Intimate Serious Offence, Intimate Violence Offence); and DASH risk assessment (Incident Rating, Prior DASH High Ratings, 28 Yes/No Questions, Risk Assessment).

concerns were raised about its function. Amnesty International note that 87% of the people in the Gangs Matrix were from black and ethnic minority communities, while 78% were black. In total, 75% were victims of violence themselves, and 35% had never committed a serious offence.[62]

The above all focus on the uses of data collection, and in most cases, AI in systems targeted at citizens and potential offenders. There are, however, a few examples of where research has been conducted to consider the behaviours of law enforcement officers themselves. For instance, Avon and Somerset Police have used algorithmic systems to scan their datasets for potential errors, including at officer level to understand crime misclassifications which might affect their algorithmic systems further downstream.[63] Others in the US have examined the possibility of using predictive analytics to predict police officers at risk of committing misconduct.[64] However, anecdotally, it appears that the large proportion of resources in this field are directed toward the citizen.

As can been seen, these systems are used widely, and draw on wide-ranging information in some circumstances. However, amongst the experts in the field there is a range of opinion as to their validity—debate continues as to the role of statistical results based on *aggregated* models, versus clinical *individual* assessments.[65] Equally, research has also shown[66] that predictive risk-scores often gain legitimacy through claims of "objectivity", tracability to case files, and use by authoritative actors like the Home Office. This "objectivity" may not be real but is thought to be present and creates the risk of false reliance.

In the US, much of the concern has focused on the role of algorithms in sentencing and parole decisions. In the UK, the focus has been seen earlier in the process—at the point of charge, for instance. In both instances, similar issues sit at the heart of the concerns expressed.

(i) Transparency and explainability

Transparency—or in rule of law language "open government" is a trait lacking in many of the systems deployed. As alluded to at the start of this chapter, the private sector development cycle and business models do not, in the absence of any standards or regulatory requirements, place a premium, on transparency and openness. Many of the vendors retain an IP hold over the underlying algorithm and will not allow access to it. From a commercial perspective some have argued that this is simply the cost payable for having these tools developed by the private sec- **19-024**

---

62 *Trapped in the Matrix: Secrecy, stigma, and bias in the Met's Gangs Database* (Amnesty International UK, 2018).

63 Dencik, Hintz, Redden and Warne, *Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services* (Data Justice Lab, Cardiff University, 2018).

64 S. Carton, J. Helsby, K. Joseph, A. Mahmud, Y. Park, J. Walsh, C. Cody, C.E. Patterson, L. Haynes and R. Ghani, "Identifying Police Officers at Risk of Adverse Events" in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (KDD '16, New York, NY, US, ACM 2016).

65 Babuta and Oswald, "Data Analytics and Algorithms in Policing in England and Wales: Towards A New Policy Framework" *RUSI Occasional Paper*, February 2020, see also L.A. Craig and A. Beech, "Best practice in conducting actuarial risk assessments with adult sexual offenders" (2009) 15 *Journal of Sexual Aggression* 193–211, *https://www.researchgate.net/publication/247498451_Best_practice_in_conducting_actuarial_risk_assessments_with_adult_sexual_offenders*.

66 D. Marciniak, "Algorithmic policing: an exploratory study of the algorithmically mediated construction of individual risk in a UK police force" (2023) 33(4) *Policing and Society* 449–463, *https://www.tandfonline.com/doi/full/10.1080/10439463.2022.2144305*.

tor and that without such proprietary protections they would not be able to succeed. However, assuming the creation of a level playing field through some form of standardised regulatory framework, it is hard to see how this would stifle growth or cause longer term harm. Fair competition between vendors, all applying the same standards of care and due diligence, at least to a minimum standard does not disadvantage any particular player and is very much the norm in many flourishing industries.

Even if one were to subscribe to the view that transparency undermines IP and in turn reduces profitability, this seems to inherently place market value above the rule of law and justice, without a public debate to draw such a conclusion. It is not the conclusion which one could argue is undemocratic, but the process—or rather the lack of.

These concerns over transparency and the tension with private profit are not theoretical, as demonstrated in *State of Wisconsin v Loomis*.[67] In this case, the defendant, Eric Loomis, filed an appeal to his sentencing based on due process concerns due to the lack of opportunity to examine the AI tool used to assess his "risk" level. The Wisconsin Supreme Court found against him. The IP of the algorithm was apparently deemed to be more valuable than the defendant's and the court's rights to have visibility of the decision-making and be able to question it.

19-025 As set out above, there are numerous applications of predictive analytics. But how many of the local citizens in local communities are aware that these systems are being used? There is no duty to be proactively transparent, to consult or engage with local communities and citizens. Even if, for argument's sake, the systems deployed were highly reliable, transparent in their explanation, bias free in their data, they are still shrouded in secrecy, and lack an openness and the legitimacy which flows from positive endorsement and support. There are notable exceptions of course, a number of forces have adopted an open approach, and the Ministry of Justice have published all the model weightings of its recidivism scoring system used in OASys alongside both in-house and peer-reviewed work analysing and explaining disparities in predictive performance by age, gender and ethnicity. But there are no guidelines or rules requiring such openness and transparency of approach.

Noel Sharkey[68] recalled the Metropolitan Police Commissioner Cressida Dick, quoting from the Peelian principles[69] in defence of the use of AI tools in UK policing. Commissioner Dick's defence was one based on the need for the police to have the same tools as society. More relevant in the Peelian principles is that of "policing by consent". Some might argue that this is the cornerstone of policing in the UK. As Sharkey sets out:

---

[67] *State of Wisconsin v Loomis* 881 N.W.2d 749 (Wis. 2016). The defendant, Eric Loomis filed an appeal to his sentencing based on due process concerns due to the lack of opportunity to examine the AI tool used to assess his "risk" level. The court found against Loomis, arguing that it did not believe that the inability to examine the "black box" led to a lack of due process. See *https://www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690*.

[68] N. Sharkey, "AI In Policing: Better Than 'A Knife Through the Chest'?" *Forbes*, March 2020, *https://www.forbes.com/sites/noelsharkey/2020/03/06/ai-in-policing-better-than-a-knife-through-the-chest/#2097d03e548e*.

[69] i.e. the principles proposed by Sir Robert Peel in 1829 and which still guide modern policing.

"The legitimacy of the police is based on a consensus of support that derives from transparency about police powers as well as their integrity and accountability in exercising their powers."

(ii) Bad data

19-026 As seen in the case of hot spot mapping, the quality of the data plays an equally important role in risk assessment tools also. Even in cases where the legal teams and others have gained access to the underlying data and methodologies, the results of those reviews have been stark and worrying. In the much-cited review in 2016, Julia Angwin and colleagues at ProPublica published a review of the COMPAS[70] risk assessment tool used in the US. The system, designed to inform decision makers on aspects such as likelihood to offend, was found to have troubling flaws. For example, only 20% of the people predicted to commit violent crimes actually went on to do so. They also concluded that the algorithm was twice as likely to falsely flag black defendants as future criminals as it was to falsely flag white defendants.[71]

The following case study is taken directly from the ProPublica Report in 2016 as it puts perfectly the case for the reasons to be concerned.

[70] Correctional Offender Management Profile for Alternative Sanctions.

[71] It ought to be noted that the for profit company which developed the tool, Northpointe disputed the results of the report.

19-027

Figure 3: Case study—Angwin, Propublica Report 2016[72]

On a spring afternoon in 2014, Brisha Borden was running late to pick up her god-sister from school when she spotted an unlocked kid's blue Huffy bicycle and a silver Razor scooter. Borden and a friend grabbed the bike and scooter and tried to ride them down the street in the Fort Lauderdale suburb of Coral Springs.

Just as the 18-year-old girls were realising they were too big for the tiny conveyances—which belonged to a six-year-old boy—a woman came running after them saying, "That's my kid's stuff". Borden and her friend immediately dropped the bike and scooter and walked away.

But it was too late—a neighbour who witnessed the heist had already called the police. Borden and her friend were arrested and charged with burglary and petty theft for the items, which were valued at a total of $80.

19-028 Compare their crime with a similar one: The previous summer, 41-year-old Vernon Prater was picked up for shoplifting $86.35 worth of tools from a nearby Home Depot store.

Prater was the more seasoned criminal. He had already been convicted of armed robbery and attempted armed robbery, for which he served five years in prison, in

---

[72] See *https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing*.

addition to another armed robbery charge. Borden had a record, too, but it was for misdemeanours committed when she was a juvenile.

Yet something odd happened when Borden and Prater were booked into jail: A computer program spat out a score predicting the likelihood of each committing a future crime. Borden—who is black—was rated a high risk. Prater—who is white—was rated a low risk.

Two years later, we know the computer algorithm got it exactly backward. Borden has not been charged with any new crimes. Prater is serving an eight-year prison term for subsequently breaking into a warehouse and stealing thousands of dollars' worth of electronics.[73]

19-029 Furthermore, research on the validity of recidivism based risk assessment tools, such as that carried out by Sarah Desmarais, Kiersten Johnson and Jay Singh[74] at a minimum ought to press home the importance of ensuring that tools used are valid, perform well and are trustworthy. Of the 19 that they reviewed, the authors found that the instruments varied widely in the number, type and content of their items. For most instruments, predictive validity had been examined in one or two studies conducted in the US that were published during the reference period. Only two studies reported on interrater reliability. No instrument emerged as producing the "most" reliable and valid risk assessments. Notably, the study also finds that the validity tests were often undertaken by the developers of the instruments under investigation. In summary one might argue that the testing of validity and quality, is patchy: where done it lacks breadth and independence, yet we are using these tools to potentially deprive people of their liberty, and to keep the public safe.

In the UK, most attention has been afforded to systems which are deployed at the point where a decision as to whether and how to charge an individual who has been arrested is made. The most often cited example of this being the HART, which (as introduced above[75]) was developed in-house by Durham Constabulary in collaboration with the University of Cambridge in 2015/16 and deployed across the force at the point of custody decision. HART was derived from the Turning Point programme, which was led by West Midlands Police and the University of Cambridge, a programme which sought to take certain vulnerable groups out of the traditional justice system and offer them alternatives to being charged for minor to moderate crimes—a practice known as "out of court disposal".[76] The intention of the police was to find a tool which assisted in decisions which were otherwise seen and experienced as being hard to make. In the case of the Durham Constabulary, the decision point related to whether an individual would be triaged into the "checkpoint" programme. The checkpoint seeks "to tackle the root causes of offending" by "offering an alternative to prosecution for a very specific sub-set of criminal offenders".[77] It seeks to reduce reoffending by responding to the observation that prosecution for certain types of crime might itself fuel reoffending, and so

---

73 See *https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing*.

74 S. Desmarais, K. Johnson and J. Singh, "Performance of Recidivism Risk Assessment Instruments in US Correctional Settings" (2016) *Psychological Services*, *https://www.researchgate.net/publication/303828802_Performance_of_Recidivism_Risk_Assessment_Instruments_in_US_Correctional_Settings*.

75 See para.13-021.

76 P. Coutts, "Turning Point: The Police's Production and Use of Evidence to Reduce Reoffending" *Alliance for Useful Evidence*, January 2018.

77 M. Oswald, "Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and 'Experimental' Proportionality" (2018) 27 *Information & Communications Technology Law* 223.

for minor crimes such as the possession of drugs, a structured set of interventions based on a pathway model, in collaboration with organisations in sectors such as mental health or alcohol and drug dependency, is offered. Because the police have a six-month window in which they can charge following a crime, and because they wish to keep the threat of charging as a motivator for successful completion of the Checkpoint programme, these programmes are four months in duration. Fewer than 5% of individuals admitted to the Checkpoint programme fail it.[78]

19-030 Although Checkpoint did not necessarily require algorithmic systems, the need for the HART tool arose because of a perception that custody officers found it challenging to identify individuals who have an appropriate level of risk of recidivism, particularly distinguishing moderate risk individuals from low or high ones accurately.[79] The tool was designed to support those custody officers and initially trained on 104,000 custody events that occurred between 2008 and 2012. Each of these events was represented by 34 different predictors most of which concern offenders' criminal behaviour histories.[80]

The criticisms of the HART tool centre on its use of (i) postcode data; and (ii) personal data purchased from data broker Experian (and derived from its Mosaic programme, a "cross-channel consumer classification system"[81]). Location data has been historically linked with issues of clustering and segregation of communities by race and ethnicity.[82] As discussed above, the location focus of hot-spot identification can discriminate against people in certain communities; the same implications arise for this tool. Furthermore, the use and purchase of data about individuals has been criticised as a breach of privacy, not to mention concerns about the lawfulness of the data sets in the first instance.[83]

(iii) Legal frameworks

19-031 A concern raised in relation to some of the tools developed in recent years is the lawful basis on which they are deployed. In the case of OASys, in practice, the system has a robust infrastructure around issues such as transparency, and data validation. However, researchers[84] have found that in this case the norms demonstrated are ones developed out of operational good practice, rather than any requirement to do so. As reported by the Law Society's Commission report, there are gaps in the existing statutory framework in relation to Data Protection. Although there are grounds on which the Government can process data about offenders, as

---

78 Oral evidence to the Law Society, *Commission on the use of AI in the Criminal Justice System*, by Chief Constable of Durham Constabulary, Michael Barton (25 July 2018).

79 There is evidence that custody officers tended to be risk-averse and err on the side of declaring an offender higher risk than data might justify—see the oral evidence to the Commission by Chief Constable of Durham Constabulary, Michael Barton (25 July 2018).

80 See S. Urwin, "*Algorithmic Forecasting of Offender Dangerousness for Police Custody Officers: An Assessment of Accuracy for the Durham Constabulary Model*" (MSc, University of Cambridge 2016), Appendix B for a list of the relevant variables used.

81 See *https://www.experian.co.uk/business/marketing/segmentation-targeting/mosaic/?utm_medium=internalRef&utm_source=Consumer%20Services*.

82 See S. Barocas and A.D. Selbst, "Big Data's Disparate Impact" (2016) 104 *California Law Review* 671, 712.

83 "*Submission to the Information Commissioner—request for an assessment notice of data brokers Experian & Equifax*" (Privacy International (PI) 2018).

84 *Commission on the use of AI in the Justice System*, Law Society of England and Wales, June 2019.

set out in the Data Protection Act 2018 (DPA 2018),[85] concerns have been raised about lack of clear provisions relating to issues such as logging and retention of metadata as it relates to personal data—either in training or deployment. As the Law Society report highlights there is need for further guidance and clarity.

A new draft bill was proposed on 8 March 2023 to amend the DPA 2018 which, if passed, would clarify what automated individual decision-making means in respect of law enforcement processing.[86] The House of Lords Justice and Home Affairs Committee's report "*Technology rules? The advent of new technologies in the justice system*"[87] discussed the risks of predictive policing and proposed many recommendations that touch on issues identified in this paper, including, but not limited to: legislation establishing clear principles and standards for the use of technologies in the justice system (recommendations 9 and 12); establishing a proper governance structure for the use of technologies in the application of the law (recommendation 5); the need for an independent body to support the various entities involved in designing and deploying new technologies (recommendation 6); empowering local specialist ethics committees in the law enforcement community (recommendation 35).[88]

In addition, the Committee concludes that humans should always be the ultimate decision-maker and that there should be more clarity on the possible duty of candour for police to ensure appropriate transparency over their use of AI (recommendation 18).[89] **19-032**

There is a growing body of research,[90] which is identifying concerns of deep-seated bias problems with the data used to train the predictive tools. Some, such as Sandra Mayson,[91] go so far as to argue that the problem is not the AI tools and their quality or data, but the

> "nature of prediction itself. All prediction looks to the past to make guesses about future events. In a racially stratified world, any method of prediction will project the inequalities of the past into the future … Algorithms, in short, shed new light on an old problem."

One of the takeaways from this is that in order to meet this challenge it would take a paradigm shift in the way in which risk is conceived and managed in criminal justice.

---

[85] See e.g. Data Protection Act 2018 Sch.7. For a fuller explanation, refer to *Commission on the use of AI in the Justice System*, Law Society of England and Wales, June 2019.

[86] Data Protection and Digital Information (No.2) Bill at *https://bills.parliament.uk/bills/3430*.

[87] House of Lords Justice and Home Affairs Committee, "*Technology rules? The advent of new technologies in the justice system*", *https://publications.parliament.uk/pa/ld5802/ldselect/ldjusthom/180/18002.htm*.

[88] See M. Zilka, H. Sargeant and A. Weller, "Transparency, governance and regulation of algorithmic tools deployed in the criminal justice system: a UK case study" in *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (2022).

[89] House of Lords Justice and Home Affairs Committee, "*Technology rules? The advent of new technologies in the justice system*", *https://publications.parliament.uk/pa/ld5802/ldselect/ldjusthom/180/18002.htm*.

[90] See e.g. Couchman, *Policing by Machine: Predictive Policing and the Threat to our Rights* (2018); R. Richardson, J.M. Schultz and K. Crawford, "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice" [2019] *New York University Law Review*; A. Guthrie Ferguson, "Policing Predictive Policing" (2017) 94 WASH. U. L. REV. 1109, 1146–51; E. Edwards, "Predictive Policing Software i More Accurate at Predicting Policing than Predicting Crime" *HUFFPOST*, 31 August 2016, *http://www.huffingtonpost.com/entry/predictive-policingreform_us_57c6ffe0e4b0e60d31dc9120*; S. Mayson, "Bias In, Bias Out" (2019) 128(8) *Yale Law Journal* 2122–2473.

[91] Mayson, "Bias In, Bias Out" (2019) 128(8) *Yale Law Journal* 2122–2473.

19-033 Such paradigm shift has emerged in other jurisdictions that have made express prohibitions on predictive policing.

In 2023, the German Constitutional Court declared the use of predictive policing AI systems unconstitutional as the relevant provisions authorising police violate the general right of personality and the right to informational self-determination.[92]

At the end of 2023, the White House issued Executive Order, on the "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence" (EO 14110).[93] EO 14110 is the most comprehensive regulatory plan issued to date by the US and includes plans to avoid improper bias and discrimination in the use of AI in the criminal justice system (s.7.1). The US Attorney General, in consultation with the Secretary of Homeland Security and the Director of the Office of Science and Technology Policy have been mandated to investigate and make best practice recommendations for the use of AI systems by law enforcement agencies.

The European Union's (EU) proposed Artificial Intelligence Act (AI Act) has taken various positions on AI in law enforcement over the iterations of the drafting process.[94] A political deal on the requirements of the AI Act was reached on 8 December 2023. Following its entry into force, AI systems for real time biometric surveillance, emotion recognition and predictive policing will be prohibited.[95] AI systems used in law enforcement also considered high-risk AI systems and will be subject to extensive obligations before being put on the market and throughout their lifecycle.[96] There are concerns by some that the proposals do not go far enough, but its progress is notably ahead of many other jurisdictions.

(iv) Predictive tools—conclusion and optimist's view

19-034 The growth in predictive analytics has provided numerous tools for police forces to use at the community and individual level. Some present benefits in resource planning and making more accurate decisions, however many come with concerns which ought to be addressed. As Babuta and Oswald[97] explain, it might be that to assess whether a tool is justified ought to be based on "*whether the tool provides useable insights which enhance the officer's ability to make an informed professional judgement in conditions of uncertainty*", rather than actually predict the

---

92 See *https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2023/bvg23-018.html*; *https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/2023/02/rs20230216_1bvr154719.html*.

93 The White House, "*Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*" 30 October 2023, *https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/*.

94 Negotiations will now begin with EU countries in the Council on the final form of the law. The law is hoped to pass by the end of 2023.

95 Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206–C9-0146/2021–2021/0106(COD))(1), Amendments 217 and 220 (prohibition on biometric surveillance), 224 (prohibition on predictive policing), 226 (prohibition on emotion inference AI systems).

96 Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206–C9-0146/2021–2021/0106(COD))(1), Amendment 234, Amendment 725–731.

97 Babuta and Oswald, "Data Analytics and Algorithms in Policing in England and Wales: Towards A New Policy Framework" *RUSI Occasional Paper*, February 2020.

future. Some argue that, if these technologies are used intelligently, this can be an important corrective to known biases and shortcomings in "human only" decision-making.

However, care needs to be taken with this approach as Mary Cummings from Duke University, states in her paper:

> "Automation bias occurs in decision-making because humans have a tendency to disregard or not search for contradictory information in light of a computer-generated solution that is accepted as correct and can be exacerbated in time critical domains. Automated decision aids are designed to reduce human error but actually can cause new errors in the operation of a system if not designed with human cognitive limitations in mind."[98]

Technology can distort the decision-making if it is not understood and, as discussed above, the idolatry of data can lead to misplaced confidence.

19-035 Transparency and openness with citizens about deployment, ensuring that decisions are explainable, explicitly dealing with the tensions with private vendor requirements which are inconsistent with sound rule of law principles, protecting privacy rights, and ensuring that capability of users is fully understood in addition to their behavioural interactions—whether that be automation bias in the form of over reliance, or indeed in underuse,[99] are all issues which remain outstanding. There are, no doubt, benefits for citizens and society alike, if the right balance and approach were to be developed, but in the gap, there are dangers with deep consequences for the fabric of our justice systems.

Accordingly, in her review of the *Loomis* case, Katherine Freeman argues that courts

> "should not use risk assessment algorithms during the sentencing process without stronger due process protections in place, if courts are to use the algorithms at all."[100]

(2) Facial recognition in policing

19-036 Much press attention is focused on facial recognition technologies and their deployment by public and private bodies. Their use has grown hugely over recent years, but with little by way of regulatory checks or standards of deployment. The concept of the ever-watchful eye of the state, the ability to track down highly dangerous individuals in large crowds, the loss of anonymity and privacy—for some the growth of facial recognition systems in society at large, and specifically by law enforcement, poses serious concerns which strike at the heart of what democracies offer their citizens, and our fundamental human rights.

(a) What is a facial recognition system?

19-037 Facial recognition technologies are designed to detect and identify individuals by comparing digital images against a list or "database" of faces. The systems look

---

98 M.L. Cummings, "*Automation Bias in Intelligent Time Critical Decision Support Systems*" Massachusetts Institute of Technology, Cambridge, MA; K. Freeman, "Algorithmic Injustice: how the Wisconsin Supreme Court failed to protect due process rights in State v Loomis" (2016) 18(5) NC J. Law Technol. 75–106.

99 M. Veale, M. Van Kleek and R. Binns, "Fairness and Accountability Design Needs for Algorithmic Support in High- Stakes Public Sector Decision-Making" in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, CHI 2018 (ACM Press, 2018).

100 K. Freeman, "Algorithmic Injustice: How the Wisconsin Supreme Court failed to protect due process rights in State v Loomis" (2016) 18(5) NC J. Law Technol. 75–106.

for "matches" to those in the database. The aim being to assist law enforcement to detect specific people—for instance suspects—quickly, in a way that is not possible by manned patrols especially in large, crowded places. Importantly, the information gathered in the process of "matching" is deemed to be biometric and therefore afforded a specific definition with the General Data Protection Regulation (GDPR)[101] and the DPA 2018.[102]

There are a number of types of facial recognition technology uses. For example, Live Facial Recognition (LFR) seeks to identify individuals in real time and is sometimes referred to as Automated Facial Recognition Locate (AFR Locate). In addition, there are systems known as Automated Facial Recognition Identify (AFR Identify), which involve retrospective identification—ex poste—from non-live footage, including from static cameras or mobile devices.

In most cases, police forces create a "watchlist"/ gallery of subjects of interest which is drawn from the Police National Database (PND), which stores images taken at the point of custody. In July 2016, a Home Office review reported that there were over 19 million custody images on the PND, over 16 million of which had been enrolled in the facial recognition gallery making them searchable using facial recognition software.[103] This number was later reduced, reportedly to 12.5 million in 2018,[104] possibly as a result of a Home Office review which found that despite the High Court ruling in 2012 in *R. (on the application of C) v Commissioner of Police for the Metropolis*,[105] the police had continued to retain the images of unconvicted individuals.

19-038 In AFR Locate systems, it is the norm for the police to use a mobile patrol—such as a van, which acts as the default command centre. The on-board monitors capture feeds of footage from the cameras in the areas being monitored. As images of individuals are captured by the cameras, the technology isolates facial images, converts them to a biometric template and compares these to the biometric templates of those on the watchlist. If a "match" is detected, an alert is sent and officers on the ground are informed.

The three most well-known police uses of facial recognition systems in the UK are by London Metropolitan Police, the South Wales Police and Leicestershire Police. All three forces have trialled technologies produced by NEC, a Japanese firm.[106] As with many instances of purchasing advanced technologies, there is little room to audit or review the software provided—it is, in many cases, a black box,

---

101 Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46 [2016] OJ L119/1.

102 "*ICO investigation into how the police use facial recognition technology in public places*", 31 October 2019.

103 "*Review of the Use and Retention of Custody Images*", Home Office, February 2017, *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/594463/2017-02-23_Custody_Image_Review.pdf*. It was reported that some of these images may have been duplicates.

104 P. Wiles, "*Annual Report 2017: Commissioner for the Retention and Use of Biometric Material*" (Office of the Biometrics Commissioner 2018), p.88.

105 *R. (on the application of C) v Commissioner of Police for the Metropolis* [2012] EWHC 1681 (Admin); [2012] H.R.L.R. 26.

106 See B. Davies, M. Innes and A. Dawson, *An Evaluation of South Wales Police's Use of Automated Facial Recognition* (Universities' Police Science Institute and Crime and Security Research Institute, Cardiff University, 2018), p.11; Met Police, *https://www.met.police.uk/advice/advice-and-information/facial-recognition/live-facial-recognition-trial/*; Leicestershire Police, *https://www.bbc.co.uk/news/uk-england-leicestershire-28307938*.

although there have been instances where changes have been requested and the technology updated.[107]

19-039 There has been much coverage in the press about the success rates—or otherwise—of facial recognition systems. The software involved is able to offer "matches" within bands of *probability*, these are not definitive "positive matches", although in common usage it is often described as such. In the review of the trial by South Wales police,[108] it was reported that the true positive at the first deployment (during football's Champions League fixtures) was 3%, which increased to 46% at the Six Nations rugby events—caused by improvements in the underlying algorithms, and increased operator familiarity. The report found that of

> "2,900 possible matches generated by the AFR system; a total of 144 confirmed true positives by operators; and a total of 2,755 categorised as 'false positives.'"

In the ARF Identify deployment, 73% generated possible suspect matches.

While the technology is still far from the science fiction accuracy of the film and novella "Minority Report", it is improving and no doubt will continue to do so, but for now it falls short of the hype. The current state of the technology has implications for standards of policing, and calls into question the trade-offs being made between personal privacy and liberty and safer streets.

---

[107] P. Nilsson, "How UK Police Are Using Facial Recognition Software" *Financial Times*, 12 October 2018.

[108] Davies, Innes and Dawson, *An Evaluation of South Wales Police's Use of Automated Facial Recognition* (2018).

19-040

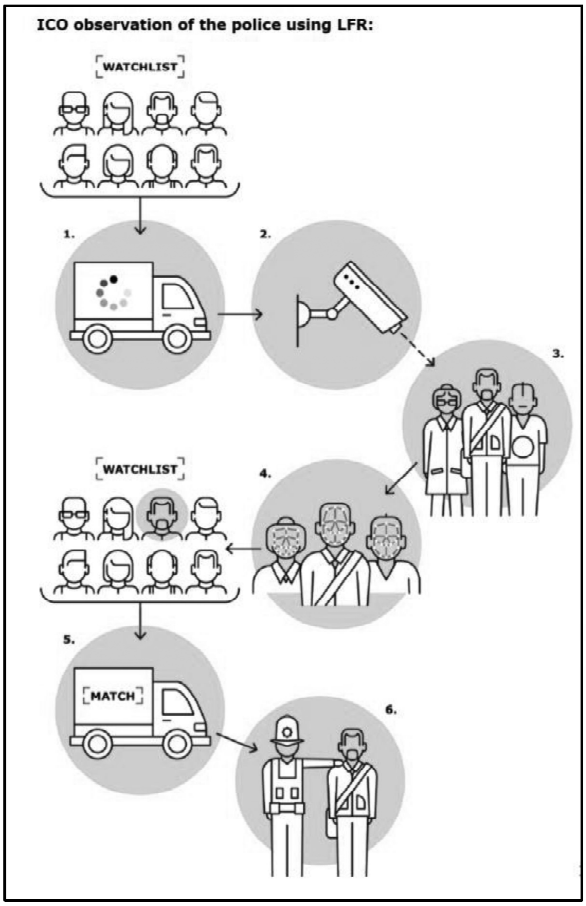

Figure 4: Infographic on the use of LRF[109]

(b) Deployment of facial recognition

19-041 Leicestershire Police was one of the first police forces to trial the live Locate facial recognition technology from April 2014, notably using the tool to look for approximately 90,000 "known offenders" at the Download Festival in June 2015.[110]

South Wales Police received funding from the Home Office in 2017 to deploy automated facial recognition in the context(s) of counter-terrorism; major events; body worn video; mobile phone app(s); automated number plate recognition; and

---

[109] "*ICO investigation into how the police use facial recognition technology in public places*", 31 October 2019.

[110] J. Purshouse and L. Campbell, "Privacy, Crime Control and Police Use of Automated Facial Recognition Technology" (2019) 3 *Criminal Law Review* 188, 190.

child sexual exploitation.[111] Funding was conditional on an evaluation of the technology being undertaken, which was carried out and published by Cardiff University.[112]

The London Metropolitan Police undertook 10 deployments of live facial recognition technology between August 2016 at Notting Hill Carnival and February 2019 in Romford Town Centre.[113] Similarly to South Wales Police, a subset of individuals from the Metropolitan Police Service's databases of photographs were extracted, primarily drawn from photos taken while individuals were in custody but also, controversially, from other sources.[114] A review of the trials, published in July 2019 by the University of Essex Human Rights, Big Data & Technology Project, identified significant flaws in the deployment.[115] Despite this report, the Met Police pressed on and started deploying LFR operationally across the capital in early 2020. Over the course of six deployments between February and July 2022, the Met Police scanned 144,366 people's biometric information, resulting in eight arrests for offences, with more deployments in 2023.[116] The Met Police and South Wales Police have commissioned research that found improved accuracy in the LFR algorithms under certain conditions and will continue their use of LFR.[117] The Home Office has also been a part of new plans to roll-out LFR in high street shops and supermarkets.[118] Criticisms of LFR has not slowed its deployment across the UK.

19-042 Internationally, a similar picture emerges with broad deployment often without standards or real public debate. Other European countries have engaged in testing and made plans for using facial recognition technology. For example, in Hungary, a project called "Szitakötő" (dragonfly) plans to deploy 35,000 cameras with facial recognition capabilities in Budapest and across the country. The cameras will capture drivers' licence plates and facial images for maintaining public order, including road safety. The Czech Government has approved a plan to expand the use of facial recognition cameras—from 100 to 145—at the Prague International Airport. Police in Germany and France have carried out extensive testing. Sweden's data protection authority has recently authorised the use of facial recognition technology by the police to help identify criminal suspects, which allows the police to compare facial images from CCTV footage to a watchlist containing over 40,000

---

111 Davies, Innes and Dawson, *An Evaluation of South Wales Police's Use of Automated Facial Recognition* (2018), 9, 12; Big Brother Watch report slightly different figures, reporting that South Wales "was awarded a total of £2.6m by the Government to carry out automated facial recognition—£1.2m in 2016/2017 and £0.8m for 2017/18 by the Home Office, as well as £0.6m from Home Office Biometrics. South Wales has additionally contributed £100,000". See *Big Brother Watch, Face Off: The Lawless Growth of Facial Recognition in UK Policing* (Big Brother Watch, 2018).

112 Davies, Innes and Dawson, *An Evaluation of South Wales Police's Use of Automated Facial Recognition* (2018).

113 London Metropolitan Police, "*Facial Recognition to Take Place in Romford*" 13 February 2019, *http://news.met.police.uk/news/facial-recognition-to-take-place-in-romford-358589*; "*Interim Report on Live Facial Recognition*" (London Policing Ethics Panel (LPEP), 2018), 7.

114 "*Interim Report on Live Facial Recognition*" (2018), 14.

115 Fussey and Murray, "*Independent report on the London Metropolitan Police service's trail of live facial recognition technology*" (University of Essex Human Rights Centre, July 2019).

116 See *https://www.met.police.uk/SysSiteAssets/media/downloads/force-content/met/advice/lfr/new/lfr-deployment-grid.pdf*; *https://www.met.police.uk/SysSiteAssets/media/downloads/force-content/met/advice/lfr/deployment-records/lfr-deployment-grid.pdf*; *https://bigbrotherwatch.org.uk/2023/05/understanding-live-facial-recognition-statistics/*.

117 See *https://science.police.uk/site/assets/files/3396/frt-equitability-study_mar2023.pdf*; *https://bigbrotherwatch.org.uk/2023/04/bbw-responds-to-lfr-report/*.

118 See *https://www.nytimes.com/2023/06/28/technology/facial-recognition-shoplifters-britain.html*; *https://www.nytimes.com/2023/06/28/technology/facial-recognition-shoplifters-britain.html*.

pictures.[119] The most high profile user of facial recognition is China, which harnesses a network of an estimated 200 million[120] cameras and, in late 2019, introduced measures requiring all mobile phone users registering new SIM cards to submit to facial recognition scans.

More reports are emerging of wrongful arrests made due to false facial recognition identification.[121] For example, there are six known cases in the US, three of which arose from the Detroit police department.[122]

Some jurisdictions are now taking a stronger approach to prevent such cases. In Scotland, parliamentary debates supported a view that there was no justification for LFR given the human rights implications and the Scottish police has said it has put plans for facial recognition on hold.[123] In the US, San Francisco was the first city to ban the use of facial recognition software by the police. Many other states and cities have followed suit.[124]

(c) Concerns

19-043 There are many concerns about both the use of facial recognition systems and the technology itself, leading to calls for a pause in its use. MPs,[125] civil society groups,[126] individual campaigners, and international neutral agencies such as the UN, have all raised serious concerns. In his report, the United Nations Special Rapporteur on Freedom of Opinion and Expression called[127] for an

> "immediate moratorium on the export, sale, transfer, use or servicing of privately developed surveillance tools until a human rights-compliant safeguards regime is in place."

Emerging concerns relate to the recent developments in emotion sensing and more personal biometric surveillance such as gait, fingerprints, DNA, and voice.[128] Emotion sensing technologies that attempt to detect the emotional state of a person raise serious concerns about the accuracy of such technologies, the methods for

---

[119] "*Facial recognition technology: fundamental rights considerations in the context of law enforcement*" EU Agency for Fundamental Rights (2019), *https://fra.europa.eu/sites/default/files/fra_uploads/fra-2019-facial-recognition-technology-focus-paper-1_en.pdf*.

[120] See *https://www.npr.org/2019/12/16/788597818/how-china-is-using-facial-recognition-technology?t=1585786936442*.

[121] T. Benedict, "The Computer Got It Wrong: Facial Recognition Technology and Establishing Probable Cause to Arrest" (2022) Wash. & Lee L. Review 849.

[122] See *https://www.theguardian.com/newsletters/2023/aug/15/techscape-facial-recognition-software-detroit-porcha-woodruff-black-people-ai*; *https://www.theguardian.com/us-news/2023/apr/27/california-police-facial-recognition-software#:~:text=Williams'%20arrest%20was%20the%20first,government%20agencies%20across%20the%20US*.

[123] M. Murgia, "Met police try to calm tensions as live facial recognition hits London" *Financial Times*, 12 February 2020.

[124] See *https://www.banfacialrecognition.com/map/*.

[125] See *https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/1970/1970.pdf*.

[126] See *https://bigbrotherwatch.org.uk/#list-met*.

[127] See *https://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session41/Documents/A_HRC_41_35.docx*.

[128] See *Motion for a European Parliament Resolution on Artificial Intelligence in Criminal Law and its Use by the Police and Judicial Authorities in Criminal Matters* (European Parliament, 2021), *https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_EN.html*.

identifying emotions, and cross-cultural dimensions of facial expression.[129] Progress has been made under EU regulation (discussed below), but issues remain if such technologies are to be operationalised by law enforcement agencies.

Broadly, these concerns can be categorised as concerns relating to data and privacy, bias and discrimination, transparency and accountability.

(i) Data and privacy

19-044 In its recent White Paper,[130] the European Commission stopped short of the suggestion that it might impose a five-year ban on facial recognition systems, but it did state that

> "the gathering and use of biometric data[131] for remote identification purposes, for instance through deployment of facial recognition in public places, carries specific risks for fundamental rights for example on people's dignity. Relatedly, the rights to respect for private life and protection of personal data are at the core of fundamental rights concerns when using facial recognition technology."

Given the absence of regulatory frameworks in the UK, these words ought to be stark warning.

As noted above in para.13-036, there are currently approximately millions of biometric, searchable faces on the PND in the UK. Police forces have powers to take certain photos with or without consent, and to disclose or retain them for purposes relating to the prevention or detection of crime, the investigation of an offence or the conduct of a prosecution or to the enforcement of a sentence.[132] There are no powers under this section of the Police and Criminal Evidence Act 1984 (PACE) to photograph and retain or repurpose images of individuals in public spaces more generally, and the main legal framework for these remains the DPA 2018.

19-045 The issue of retention of photographs has become an area of contention. In 2009, the Court of Appeal found that there had been an *unjustified* inference with the claimant's right to respect for private life (European Convention on Human Rights

---

[129] L. Urquhart and D, Miranda, "Policing faces: the present and future of intelligent facial surveillance" (2022) 31(2) *Information & Communications Technology Law* 194–219.

[130] White Paper, *Artificial Intelligence—A European approach to excellence and trust* (European Commission, 2020) COM(2020) 65 final.

[131] Biometric data is defined as "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique authentification or identification of that natural person, such as facial images or dactyloscopic [fingerprint] data." (Law Enforcement Directive 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA [2016] OJ L119/89 art.3(13); Regulation 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46 [2016] OJ L119/1 (GDPR) art.4(14); Regulation 2018/1725 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation 45/2001 and Decision 1247/2002 [2018] OJ L295/39 art.3(18).

[132] Police and Criminal Evidence Act 1984 s.64A(4).

art.8), when his photo was taken and retained by police, despite him not committing any offence.[133]

In the 2012 case of *RMC and FJ*, the High Court ruled in favour of two individuals who challenged the Metropolitan Police Service for retaining custody photographs taken under PACE.[134] This ruling stated that the "existing policy concerning the retention of custody photographs … is unlawful", and the police were given a "reasonable further period" for revising this policy. The lack of action flowing from this ruling has been a subject of concern for those in the field. The Biometrics Commissioner, established in 2012, commented in his annual reports[135] on the lack of action on the matter. This was in turn raised as a matter of concern by the House of Commons Science and Technology Committee.[136] The Home Office review in 2016, led to the right for individuals to ask for their photos to be removed from the custody database if they qualified, although very low numbers are reported to have been successful in these applications.[137]

In the UK, a 2019 report by the Information Commissioner's Office (ICO) makes very clear that the use of LFR in public spaces by organisations, in both the public and private sectors involving the processing of personal data and requires compliance with GDPR and the DPA 2018. It goes on to highlight why it feels it is an important area to ensure standards are met:

> "Current and future use of facial recognition technology is a regulatory priority for the [ICO]. This is based on the following:
>
> 1) scale of privacy intrusion, with the potential to affect large numbers of people, in many cases without their knowledge, as they go about their daily lives;
> 2) the potential for facial recognition technology to enable surveillance on a mass scale, and the impact this has on individuals' human rights and information rights;
> 3) technological bias or inaccurate data that could lead to detriment, e.g. an individual being misidentified and potentially apprehended, thus undermining the integrity and legitimacy of the use of the technology;
> 4) uncertainty about the effectiveness of the technology in meeting the expected law enforcement or public interest aims; and
> 5) the potential for poor compliance to undermine public confidence in the police and trust in the technology." [138]

19-046 Although the ICO did not find that there was reason for regulatory action, it did conclude that there needed to be improvements in how police authorised and deployed the technology if it was to retain public confidence and address privacy

---

[133] *R. (on the application of Wood) v Commissioner of Police of the Metropolis* [2009] EWCA Civ 414; [2009] H.R.L.R. 25. See generally J. Grace and M. Oswald, "'Being on Our Radar Does Not Necessarily Mean Being under Our Microscope': The Regulation and Retention of Police Intelligence" (2016) 22(1) *European Journal of Current Legal Issues* 1–17.

[134] *R. (on the application of C) v Metropolitan Police Service* [2012] EWHC 1681 (Admin); [2012] H.R.L.R. 26.

[135] D. Marshall and T. Thomas, *Privacy and Criminal Justice* (Palgrave Macmillan, 2017), p.129.

[136] House of Commons, Science and Technology Select Committee, *Current and future uses of biometric data and technologies*, Sixth Report of Session 2014–15 (Stationery Office, 2015), HC Paper No.734 (Session 2014/15), p.3.

[137] Press Association, "'Custody Image' Deletion Request Figures Revealed" *Mail Online*, 12 February 2018, *http://www.dailymail.co.uk/wires/pa/article-5379353/Custody-image-deletion-request-figures-revealed.html* (figures from 37/43 police forces in England and Wales, based on Freedom of Information requests).

[138] "*ICO investigation into how the police use facial recognition technology in public places*", 31 October 2019.

concerns. It highlighted evidence of data processing good practice, but also areas where data protection compliance could be improved. It identified what it calls "missed opportunities" to improve public awareness and confidence in the technologies. Most notably, the ICO concludes that "[t]he absence of a statutory code of practice and national guidelines contributes to inconsistent practice, increases the risk of compliance failures". The ICO also found that more specific objectives and watchlists were necessary to meet the high bar set within the data protection legislation, and that there was room to be clearer as to the definition of "effective" in deployments of the technology in order to make assessments of proportionality, in addition concerns of technology bias and false matches were raised as an area which required progress.

Also in 2019, the case of *R. (on the application of Bridges) v Chief Constable of South Wales*[139] was heard. This judicial review centred on the South Wales Police's use of AFR which the claimant argued was a breach to his right to privacy, data protection laws and anti-discrimination laws.

In relation to data protection, the High Court confirmed as many believed, that AFR does involve the processing of personal data going so far as to state "[l]ike fingerprints and DNA … it is information of an 'intrinsically private' character", that the core information collected is indeed biometric, and as such the DPA 2018 is the primary legislation regulating the use of AFR. The court also found that in the instances used the South Wales Police used the AFR lawfully and did not consider the legal framework at present to be insufficient but highlighted that this would need to be revisited. The court did, however, endorse that:

> "(a) steps could, and perhaps should, be taken further to codify the relevant legal standards; and (b) the future development of AFR technology is likely to require periodic re-evaluation of the sufficiency of the legal regime. We respectfully endorse both sentiments, in particular the latter. For the reasons we have set out already, we do not consider that the legal framework is at present out of kilter; yet this will inevitably have to be a matter that is subject to periodic review in the future."

19-047 This decision was subsequently successfully appealed in 2020. The Court of Appeal[140] found that there were fundamental deficiencies in the legal framework, which in turn led to a breach of the appellants rights.

The judgment sets out the rationale for its decision as hinging on the following aspects:

> "First, AFR is a novel technology
>
> Secondly, it involves the capturing of the images and processing of digital information of a large number of members of the public, in circumstances in which it is accepted that the vast majority of them will be of no interest whatsoever to the police.
>
> Thirdly, it is acknowledged by all concerned that this is 'sensitive' personal data, within the meaning of the DPA 2018. That Act in turn reflects EU legislation. This represents an

---

[139] *R. (on the application of Bridges) v Chief Constable of South Wales* [2019] EWHC 2341 (Admin); [2020] 1 Cr. App. R. 3. See *https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf*.

[140] See *Bridges* [2019] EWHC 2341 (Admin) at [86]–[89].

institutional recognition of the sensitivity of the data concerned, a feature which is not present for example for ordinary photographs.

Fourthly, the data is processed in an automated way."[141]

19-048 Importantly, although the Court of Appeal accepted large aspects of the analysis of the Divisional Court, it concluded that

> "the legal framework which the Divisional Court regarded as being sufficient to constitute the 'law' for the purposes of Article 8(2) is on further analysis insufficient."[142]

The court described its concerns related to the legal framework as the "who question" and the "where question"—in both cases concluding that too much discretion is currently given to individual police officers:

> "It is not clear who can be placed on the watchlist nor is it clear that there are any criteria for determining where AFR can be deployed."[143]

The court also stated that:

> "the fact remains, however, that [South Wales Police] SWP have never sought to satisfy themselves, either directly or by way of independent verification, that the software program in this case does not have an unacceptable bias on grounds of race or sex … because, for reasons of commercial confidentiality, the manufacturer is not prepared to divulge the details so that it could be tested. That may be understandable but, in our view, it does not enable a public authority to discharge its own, non-delegable, duty under section 149."[144]

19-049 South Wales Police have not appealed the ruling, and as has been pointed out, although the case was specifically related to the use by South Wales Police of AFR, "the outcome is much more far reaching with regard to the use of this technology by policing more generally".[145] However, some scholars suggest that the judgment leaves considerable room for police AFR to continue with only minor, piecemeal amendment to the legal framework. Purshouse and Campbell argue "that the relatively unfettered rise of police facial recognition in England and Wales illuminates deeper flaws in the domestic framework for fundamental human rights protection and adjudication, which create the conditions for authoritarian policing and surveillance to expand".[146]

Since the ruling, the UK Police have issued new professional practice guidelines on the use of AFR technology. The guidance focuses on ensuring live facial recognition technology is used in a "responsible, transparent, fair and ethical way and only when other, less intrusive methods would not achieve the same results", that it is "targeted, based on intelligence and have a set time for use to start and end", that other proportionality and necessity criteria are met before deployment, and chief

---

[141] See *Bridges* [2019] EWHC 2341 (Admin) at [105].
[142] See *Bridges* [2019] EWHC 2341 (Admin) at [90].
[143] See *Bridges* [2019] EWHC 2341 (Admin) at [96].
[144] See *Bridges* [2019] EWHC 2341 (Admin) at [199].
[145] T. Porter, "*Surveillance Camera Commissioner*", *https://videosurveillance.blog.gov.uk/2020/08/11/what-next-for-automated-facial-recognition/*.
[146] J. Purshouse and L. Campbell, "Automated facial recognition and policing: A Bridge too far?" (2022) 42(2) *Legal Studies* 209–227.

officers "involve their elected police and crime commissioner to provide oversight".[147]

The UK framework may still be problematic in relation to watchlists given concerns around the role of social media intelligence (SOCMINT) by police and the legalities of sourcing content for investigations from these channels, as well as the use of lateral surveillance practices.[148] While the guidance could be said to satisfy the Court of Appeal's concerns about the lack of criteria, the methods are still questionable because they remain very broad. The real friction in *Bridges* was the "interplay between private vendors and police" meaning that the commerciality rules ran in the face of Public Sector Equality Duty. Despite good intentions, concerns remain that the guidance will not fully address the public sector equality concerns and the limitation will not ultimately restrict the use of LFR in the UK.

There is also an interesting juxtaposition between the UK policy a framework for LFR use that addresses public concerns, and the shifts in the EU AI Act, which seeks to prohibit LFR by default and only allow use in specific circumstances.[149]

19-050 As discussed above, the discussion of facial recognition technology has also been actively debated within Europe. The European Parliament has been opposed to mass surveillance and called for a ban on private facial recognition databases such as Clearview.[150] In April 2023, the European Data Protection Board adopted "*Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement*".[151] Guidelines 05/2022 call on prohibiting high-risk uses of AFR in law enforcement that poses unacceptable threats to individual and collective rights and freedoms, such as when it is used in mass surveillance, without individuals' knowledge, inferring emotions or categorising based on sensitive characteristics. Furthermore, it explains the indiscriminate collection of personal data to populate facial recognition databases does not comply with necessity and proportionality requirements.[152] The current text of the AI Act reflects the specific concerns in these guidelines and will legislate to control, and in some circumstances, ban on biometric surveillance.[153]

(ii) Concerns and legislation around bias

19-051 As in the case of predictive policing tools, bias in training sets, and data, as well as bias in deployment are a concern for many. As the non-profit campaigning

---

147 See *https://www.college.police.uk/article/live-facial-recognition-technology-guidance-published*.

148 L. Urquhart and D. Miranda, "Policing faces: the present and future of intelligent facial surveillance" (2022) 31(2) *Information & Communications Technology Law* 194–219.

149 Urquhart and Miranda, "Policing faces: the present and future of intelligent facial surveillance" (2022) 31(2) *Information & Communications Technology Law* 194–219.

150 See "Use of artificial intelligence by the police: MEPs oppose mass surveillance" *European Parliament News*, 6 October 2021, *www.europarl.europa.eu/news/en/press-room/20210930IPR13925/use-of-artificial-intelligence-by-the-police-meps-oppose-mass-surveillance*.

151 European Data Protection Board, "*Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement*", *https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-052022-use-facial-recognition_en*.

152 European Data Protection Board, "*Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement*".

153 Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206–C9-0146/2021–2021/0106(COD))(1), Amendments 217 and 220 (prohibition on biometric surveillance).

organisation Big Brother Watch reported, the software used by the UK police forces "has not been tested for demographic accuracy biases".[154] Systems which run the risk of performing poorly on specific communities, whether that be by race of gender.[155] In these two cases, both run the risk of discrimination and would be in breach of equality duties, as well as GDPR concerns.[156]

As with other criticisms associated with data (e.g. poorly constructed training data sets and software, as well as focused deployment in areas already subject to heavy policing activity), there is a heightened risk of bias reinforcing bias.

(iii) Concerns and legislation around accountability, transparency and oversight

19-052 As in any other sector of society or the economy a sound system requires a set of checks and balances—accountability, transparency and oversight being core to this.

However, in the use of facial recognition systems, there appears to be a significant gap. These concerns have been reported and commented on by many, for instance the Biometrics and Forensics Ethics Group, an advisory non-departmental public body sponsored by the Home Office,[157] the House of Commons Science and Technology Select Committee,[158] the Information Commissioner,[159] the Biometrics Commissioner,[160] in addition to civil society groups and commentators, and in the recent Court of Appeal case: *R. (on the application of Bridges) v Chief Constable of South Wales*.[161]

Despite the high figures of general levels of trust among the public in the use of these technologies, as they become more pervasive and woven into the standard approaches to policing, the greater the attention and the greater the need to ensure not only that there is a lawful basis, but also that there is consent to use it. For those communities who feel marginalised, within whom trust in the systems is reported

---

154 "*Face Off: The Lawless Growth of Facial Recognition in UK Policing*" (Big Brother Watch, 2018), 16.

155 J. Buolamwini and T. Gebru, "*Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*" in Conference on Fairness, Accountability and Transparency (2018).

156 See "*General Data Protection Regulation 2016*" Commission on AI in Criminal Justice, Law Society, June 2019, recital 71 ("the controller should use appropriate mathematical or statistical procedures for the profiling, implement technical and organisational measures appropriate … that prevent, inter alia, discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or that result in measures having such an effect").

157 Biometrics and Forensics Ethics Group Facial Recognition Working Group, *Ethical Issues Arising from the Police Use of Live Facial Recognition Technology* (Biometrics and Forensics Ethics Group 2019), 3.

158 House of Commons, Science and Technology Select Committee, *Current and future uses of biometric data and technologies*, Sixth Report of Session 2014–15 (Stationery Office, 2015), HC Paper No.734 (Session 2014/15), p.3.

159 R. Hill, "*ICO to probe facial recog amid concerns UK cops can't shake their love for unregulated creepy tech*" theregister.com, 3 December 2018, *https://www.theregister.com/2018/12/03/ico_investigation_facial_recognition/*.

160 D. Marshall and T. Thomas, *Privacy and Criminal Justice* (Palgrave Macmillan, 2017), p.129.

161 *R. (on the application of Bridges) v Chief Constable of South Wales* [2020] EWCA Civ 1058; [2020] 1 W.L.R. 5037. The judgment highlighted both gaps in the legislative framework, but also the concerns around the lack of transparency of the algorithms used in the system.

to be lower than in other groups,[162] this matters significantly if they are to trust that these systems are designed to protect the community as a whole and not to victimise it.

19-053 Accountability is also a concern, especially in relation to a public private division of roles. As noted previously in this chapter, many of these systems are purchased from third party private sector vendors. There is evidence to suggest that the police do not have the requisite in-house skills and capacity to scrutinise and test the technology they are deploying, and to fully understand its operations and limitations. Given the issues raised previously around automation bias, it is highly important that those using these tools understand how they operate and when to trust / not to trust the results. As with any supply chain, there will be some products which are well developed, have good in-house governance checks and quality assurance mechanisms and those which do not. As with any area of the economy, these issues are further exacerbated by the desire to deliver products quickly, and continue their development in situ, linked closely to ideas of quickly and cheaply developing a minimal viable product, and improving it through agile development.[163]

Determining which human rights are engaged, and how, is a complex exercise it takes a wide ranging in depth analysis to fully understand the issues engaged and the appropriate responses. It does not seem either feasible, or good public policy, to ask individual forces to navigate these issues in an ad hoc manner without central support, leadership and coordination.

The user is not the only party with responsibility; it rests across the supply chain. In the broader economic debate, there is an increasing awareness of the duties which businesses have to society. Profit with a purpose,[164] or at least conscious governance and standards of ethics are finding a voice around the Board table and investment decisions. As more and more businesses recognise that they too have a duty to the delivery of the UN's Sustainable Development Goals,[165] so too should those developing the technologies which are being deployed in society. There should be an expectation that they would consider their impact on human rights and data protection. "Privacy by design" was a popular mantra in the early days of GDPR adoption, "human rights by design" ought to be as equally important. As set out in the Law Society's report,[166] there ought to be a clear obligation, for those planning to sell their products to the public sector, in this field, to have completed a human rights impact assessment at the concept stage, design stage and at the "go to market" stage. Utilising a standard approach would provide certainty and a level playing

---

[162] Research was carried out on behalf of the London Policing Ethics Panel, the (London) Mayor's Office for Policing and Crime and the University College London Institute for Global City Policing by Opinion Research Services 11. Unlike the ICO's research, the survey was not national, but weighted to provide a representative sample of London's population. This survey again found broad support for the use of LFR for policing purposes, with 57% of all those surveyed agreeing that it was acceptable for the MPS to use LFR. But majorities of Asian (56%) and black (63%) people surveyed were opposed. Support is also lower amongst young people in London, with 55% of 16–24 and 52% of 25–39-year-olds opposed to the police use of LFR. See "*Interim Report on Live Facial Recognition*" (LPEP, 2018).

[163] See generally S. Gürses and J. van Hoboken, "Privacy after the Agile Turn" in E. Selinger, J. Polonetsky and O. Tene (eds), *The Cambridge Handbook of Consumer Privacy* (Cambridge: Cambridge University Press, 2018).

[164] See *https://www.thebritishacademy.ac.uk/programmes/future-of-the-corporation*.

[165] See *https://sustainabledevelopment.un.org/sdgs*.

[166] *Commission on the use of AI in the Justice System*, Law Society of England and Wales, June 2019, Recommendation 4.

field for those developing these products and a degree of quality assurance for those purchasing them. Ex ante frameworks are a suitable approach where the harm if unchecked is likely to be high, and in this case the evidence is overwhelming. Of course, the responsibility in the supply chain does not rest with the developers alone—this is a duty which is shared across the supply/demand equation. As a procuring party there is a responsibility to undertake reasonable due diligence. Procurement teams have a hugely powerful position, they are able to flex their buying power to demand certain standards from suppliers—and undertaken across a market, in a coordinated way would be an easy and effective approach.

V. Conclusion

19-054 As the ICO points out, there are some significant gains to be made through the well governed use of AFR systems, and the same applies to aspects of the predictive policing tools. But these ought not to come at the cost of fundamental rights. If the data is clean, bias free and processed in compliance with the DPA and GDPR requirement—is that the only test? Or does a citizen's right to convene, anonymously, without surveillance have a role in the way we wish to construct the social licence of our societies? Does the right to feel that you are able to walk through your neighbourhood and not feel that your community is being "targeted" have a bearing in these decisions? It is of some concern that the majority of the debates on these technologies focuses more on correcting the technology flaws, than the trade-offs we would be making in our core principles. This has continued at pace in response to the developments in generative AI capabilities. In both the cases discussed in this chapter, there is little by way of structured, open and informed debate with society and citizens, there remain, despite the developments in some regions, a lack of standards to maintain even minimum protections, there is a heavy reliance on private enterprise which is not aligned with either privacy concerns or rule of law principles, there remain huge gaps in capability and capacity, and all the while there is a headlong sprint for faster, wider and "better" deployment.

As one eminent member of judiciary said to me on the night of the launch of the Law Society Commission's report—"it seems to me, with a heavy heart, that your work has come too late and that the horse has long bolted". As I said then, and I repeat now, I would argue the horse is kicking hard at the door and is not yet fully free, and even if I am wrong, I do not believe we ought to sit back and resign ourselves to never catching up and grasping the reins. If we have a sense of the importance of our criminal justice system, if we believe that police act with consent, if we believe in the virtue of open and balanced debate to strike the right tone between often complex trade-offs, then we will at least try and find that horse, and tame it.