RUSI
www.rusi.org

Royal United Services Institute
for Defence and Security Studies

Occasional Paper

Data Analytics and Algorithms in Policing in England and Wales

Towards A New Policy Framework

Alexander Babuta and Marion Oswald

Data Analytics and Algorithms in Policing in England and Wales: Towards A New Policy Framework

Alexander Babuta and Marion Oswald

RUSI Occasional Paper, February 2020

Royal United Services Institute
for Defence and Security Studies

189 years of independent thinking on defence and security

The Royal United Services Institute (RUSI) is the world's oldest and the UK's leading defence and security think tank. Its mission is to inform, influence and enhance public debate on a safer and more stable world. RUSI is a research-led institute, producing independent, practical and innovative analysis to address today's complex challenges.

Since its foundation in 1831, RUSI has relied on its members to support its activities. Together with revenue from research, publications and conferences, RUSI has sustained its political independence for 189 years.

The views expressed in this publication are those of the author, and do not reflect the views of RUSI or any other institution.

Published in 2020 by the Royal United Services Institute for Defence and Security Studies.

RUSI Occasional Paper, February 2020. ISSN 2397-0286 (Online); ISSN 2397-0278 (Print).
Printed in the UK by Stephen Austin and Sons, Ltd.

Royal United Services Institute
for Defence and Security Studies
Whitehall
London SW1A 2ET
United Kingdom
+44 (0)20 7747 2600
www.rusi.org
RUSI is a registered charity (No. 210639)

Contents

Acknowledgements

The authors would like to thank all those who took part in interviews and focus groups for this project, without whom the research would not have been possible. The authors are particularly grateful to the police officers and staff who gave up their valuable time to contribute to this project. A number of individuals have provided especially useful guidance, advice and practical support, including (but by no means limited to) Detective Chief Superintendent Chris Todd, Deputy Chief Constable David Lewis, Chief Superintendent David Powell, Professor Peter Fussey, Dr Gillian Tully, Tom McNeil and Giles Herdale. The authors would also like to thank techUK for their assistance in organising the roundtable event in July 2019, particularly Katherine Mayes and Jessica Russell.

The authors are very grateful to the reviewers who provided invaluable comments on an earlier version of this paper, particularly Johanna Morley, Reuben Binns and Jonathan Dowey, as well as RUSI colleagues (past and present) Malcolm Chalmers, Charlie de Rivaz, Keith Ditcham, Ardi Janjeva and Evangelina Moisi. The authors would also like to thank the RUSI Publications team for their expert editorial input and patience in producing this paper. Finally, the authors are very grateful to the Centre for Data Ethics and Innovation for supporting this work and providing helpful comments and suggestions throughout the project, in particular Lara Macdonald and Mark Durkee.

Executive Summary

RUSI WAS COMMISSIONED by the Centre for Data Ethics and Innovation (CDEI) to conduct an independent study into the use of data analytics by police forces in England and Wales, with a focus on algorithmic bias. The primary purpose of the project is to inform CDEI's review of bias in algorithmic decision-making, which is focusing on four sectors, including policing, and working towards a draft framework for the ethical development and deployment of data analytics tools for policing.

This paper focuses on advanced algorithms used by the police to derive insights, inform operational decision-making or make predictions. Biometric technology, including live facial recognition, DNA analysis and fingerprint matching, are outside the direct scope of this study, as are covert surveillance capabilities and digital forensics technology, such as mobile phone data extraction and computer forensics. However, because many of the policy issues discussed in this paper stem from general underlying data protection and human rights frameworks, these issues will also be relevant to other police technologies, and their use must be considered in parallel to the tools examined in this paper.

The project involved engaging closely with senior police officers, government officials, academics, legal experts, regulatory and oversight bodies and civil society organisations. Sixty-nine participants took part in the research in the form of semi-structured interviews, focus groups and roundtable discussions. The project has revealed widespread concern across the UK law enforcement community regarding the lack of official national guidance for the use of algorithms in policing, with respondents suggesting that this gap should be addressed as a matter of urgency.

In recent years, police use of algorithms has expanded significantly in scale and complexity. This is driven by three closely related factors. First, a significant increase in volume and complexity of digital data has necessitated the use of more sophisticated analysis tools. Second, ongoing austerity measures have resulted in a perceived need to allocate limited resources more efficiently based on a data driven assessment of risk and demand. And third, the police service is increasingly expected to adopt a preventative, rather than reactive posture, with greater emphasis on anticipating potential harm before it occurs.

While new data technologies clearly have the potential to improve police effectiveness and efficiency, concerns were raised regarding their development and implementation. Interviewees highlighted the lack of an evidence base, poor data quality and insufficient skills and expertise as three major barriers to successful implementation. In particular, the development of policing algorithms is often not underpinned by a robust empirical evidence base regarding their claimed benefits, scientific validity or cost effectiveness. A clear business case is therefore often absent. In the context of statistical forecasting, claims of 'predictive accuracy' are often misunderstood

or misinterpreted, making it difficult for the force to assess a tool's real-world benefits. Furthermore, capability development is largely driven by data science, with comparatively little focus on the underlying conceptual framework, criminological theory or legal requirements.

Police use of advanced algorithms, predictive analytics and 'data scoring' tools raises various legal and ethical concerns. The deployment of such technology as a direct response to resourcing constraints prompts significant questions regarding necessity and proportionality: in some cases, it could be argued that the use of such tools would not be 'necessary' if the police force had the resources needed to deploy a non-technological solution to the problem at hand, which may be less intrusive in terms of its use of personal data. In addition to data protection issues, there are a number of human rights considerations, and concerns were raised that these are not always considered at the outset of new projects. To address these concerns, it is recommended that an integrated impact assessment – covering data protection, human rights, discrimination risk, assessment of empirical accuracy and operational effectiveness, as well as any other relevant legal requirements – should be conducted at the outset of any new police analytics project, to assess whether a clear justification for using the tool has been established.

While predictive policing tools have received much criticism for being 'racially biased', with claims that they over-predict individuals from certain minority groups, there is a lack of sufficient evidence to assess the extent to which bias in police use of algorithms actually occurs in practice in England and Wales, and whether this results in unlawful discrimination. Most studies purporting to demonstrate racial bias in police algorithms are based on analysis conducted in the US, and it is unclear whether these concerns are transferable to the UK context. However, there is a legitimate concern that the use of algorithms may replicate or amplify the disparities inherent in police-recorded data, potentially leading to discriminatory outcomes. For this reason, ongoing tracking of discrimination risk is needed at all stages of a police data analytics project, from problem formulation and tool design to testing and operational deployment.

Treating algorithmic insights as a form of 'police intelligence', associated with a level of confidence, would ensure that users of the tool critically assess the validity and relevance of all information when forming their overall judgement, thereby ensuring ultimate accountability of the decision-making process. When police forces procure 'commercial off-the-shelf' analytics tools, appropriate access rights must be granted for the force to be able to audit the underlying statistical models if needed, for instance to assess risk of bias and error rates. Intellectual property rights must not be a restriction on this scrutiny.

Research participants universally recognised a lack of any official national guidelines for police use of algorithms. Furthermore, interviews revealed a lack of clarity regarding the delineation of responsibilities between different organisations for the development of standards and guidelines, regulation and oversight. A new set of nationally approved guidelines appears to be essential to ensure the legitimate development and deployment of statistical algorithms for policing. Establishing these guidelines will require a joint approach between the National Police Chiefs' Council (NPCC) and the Home Office, with input from the College of Policing. Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) should inspect forces' compliance against these standards as part of crime data integrity inspections, drawing on the combined expertise of the Information Commissioner's Office and the Equality and Human Rights Commission as appropriate. Context-specific evaluation methodologies should also be developed to ensure the empirical validity of statistical algorithms used by the police.

To ensure a coordinated approach to the development and deployment of data analytics tools in policing, the NPCC and Association of Police and Crime Commissioners (APCC) should establish a national coordinating group for data analytics. The group should maintain a high-level catalogue of all algorithms used by police forces nationwide to inform operational decision-making, and introduce a mechanism by which consistent specialist expertise can be accessed by forces in the areas of data science, ethics and tool evaluation. The group should also assess the feasibility of establishing a mechanism for police forces to access a centralised team of specialist legal advisers, in the same way that government departments can access specialist legal advice via the Government Legal Department.

There was widespread recognition of the need for meaningful and independent 'ethical oversight' of police data analytics projects, but a lack of clarity on how this should be achieved in practice. In particular, there was no definitive conclusion as to whether this oversight should be delivered at the local force level, or in the form of a centralised national structure (or both). It appears unlikely that existing general police ethics committees could provide meaningful ethical scrutiny of specialist police data analytics projects. However, the resource and funding requirements for the establishment of bespoke digital ethics committees are considerable and could be prohibitive, suggesting the need for forces and Police and Crime Commissioners (PCCs) to consider a regional model of digital ethics committees subject to consistent terms of reference and transparency requirements.

Any future policy framework should be principles-based and complement existing police guidance in a 'tech-agnostic' way. Rather than establishing prescriptive rules and standards for different data technologies, the framework should establish standardised processes to ensure that data analytics projects follow recommended routes for the empirical evaluation of algorithms within their operational context and evaluate the project against legal requirements and ethical standards. The new guidance should focus on ensuring multi-disciplinary legal, ethical and operational input from the outset of a police technology project; a standard process for model development, testing and evaluation; a clear focus on the human–machine interaction and the ultimate interventions a data driven process may inform; and ongoing tracking and mitigation of discrimination risk.

Recommendations

Police Forces

- Before investing in new data analytics software as a full operational capability, an integrated impact assessment should be conducted, to establish a clear legal basis and operational guidelines for use of the tool. This should incorporate the following elements:

  - Data protection impact assessment.
  - Equality impact assessment, describing the potential impact of the proposed project on people with protected characteristics.
  - Human rights impact assessment.
  - Empirical evaluation of accuracy and operational assessment of 'real-world' effectiveness.
  - Assessment of expected level of errors, where this can be established or estimated, and potential consequences of these errors.
  - Assessment of any positive obligations under Article 2 or Article 3 of the European Convention on Human Rights or associated public safeguarding issues.
  - Assessment of any other legal requirements which may be relevant for specific projects (for instance, investigatory powers authorisations, evidential or valid decision-making requirements pursuant to criminal procedure, PACE requirements and investigations legislation, and any limitations on interventions a statistical algorithm may inform).
  - Independent ethical assessment, the format of which will depend on what ethical oversight arrangements are in place.
- Throughout the project lifecycle, the police force should keep under constant review the resources required for the project and anticipated efficiency gains, to ensure the project is meeting its goals as set out in the initial business case and impact assessment.
- A 'senior responsible owner' should be assigned to each police data analytics project, to ensure full accountability to the Chief Constable and PCC, and oversight not just for the performance of the tool but also how it is deployed operationally.
- Statistical forecasting systems based on algorithms should not be described as 'predictive policing' or 'risk assessment' tools, but more accurately as 'classification and prioritisation systems', with the human user maintaining ultimate responsibility for the overall risk assessment.
- The output of statistical algorithms should be classified as a form of police intelligence, alongside a confidence rating indicating the level of uncertainty associated with the prediction. How the confidence rating is established and maintained will depend on the type of algorithm used, and the method for calculating this should be established alongside routes for empirical evaluation. Officers and analysts should be expected to consider the output alongside other forms of relevant police intelligence when arriving at their overall judgement or decision.

Policing Bodies, Regulators and Other Government Departments

- The NPCC, in consultation with the APCC, should continue their ongoing work to develop new national guidelines for police use of data analytics, drawing on the existing 'Algocare' model and the framework currently being developed by the CDEI. The 'integrated impact assessment' detailed above should be a core requirement of these new guidelines. This new guidance should form part of the new 'National Data Ethics Governance Model' proposed in the recent NPCC-APCC 'National Policing Digital Strategy'. The Home Office should ensure that this work is appropriately supported, including by way of input from relevant stakeholders.

- HMICFRS should establish an External Reference Group for police use of data analytics, with a view to incorporating use of data analytics and its effectiveness into future crime data integrity inspections. This should draw upon the combined expertise of the Information Commissioner's Office and the Equality and Human Rights Commission as appropriate.
- The NPCC and APCC should establish a national coordinating committee for data analytics. The group should:
    - Maintain a high-level catalogue of all algorithms used by police forces nationwide to inform operational decision-making, to encourage cooperation between forces, sharing of best practice and avoidance of duplication.
    - Introduce a mechanism by which consistent specialist expertise can be accessed by forces in the areas of data science, ethics and tool evaluation.
    - Explore the feasibility of establishing a mechanism for police forces to access a centralised team of specialist legal advisers, in a similar way that government departments can access specialist legal advice via the Government Legal Department.
- The UK Police Ethics Guidance Group should conduct a comprehensive review of ethics committees, to assess whether existing force ethics committees could be 'upskilled' to provide meaningful ethical review of police technology projects, or whether bespoke digital committees could be established in parallel. This review should also consider the viability, resourcing and funding requirements of a national or regional ethics review process based on standardised terms of reference.
- The Home Office Data and Identity Directorate should clarify roles and responsibilities regarding the development of context-specific evaluation methodologies for statistical algorithms used by police forces in England and Wales. This should include guidance on how confidence levels and error rates should be established, communicated and evaluated. These evaluation methodologies should be developed drawing on the expertise of the Forensic Science Regulator in establishing scientific standards for forensic science.
- Further empirical research is needed to assess the extent to which racial bias in police use of algorithms occurs in practice in the UK, and whether this results in unfair discrimination.

Software Developers

- When developing proprietary 'commercial off-the-shelf' algorithmic software for use by police forces, the provider must ensure appropriate rights of access are granted for the procuring force and national regulators to be able to audit the underlying statistical models if needed (for instance, to assess risk of bias and error rates). Intellectual property rights must not be a restriction on this scrutiny.
- When developing so-called 'data scoring' algorithms related to individuals, the use of unsupervised machine learning methods (such as auto-encoders) to create features

should be avoided at the feature engineering stage. Human-interpretable features are essential to provide sufficient transparency regarding what factors were taken into account during computation, assess whether the process was discriminatory, relevant and proportionate to the decision at hand or had any causal justification.

- To avoid model degradation, machine learning models should not be 'retrained on the fly', but should be reviewed and updated regularly, with particular focus on the suitability of the input data used, and the extent to which this is consistent with the original training data. This should be conducted by specialist data scientists while maintaining sufficient legal and operational input.

Introduction

Research Rationale

RUSI WAS COMMISSIONED by the Centre for Data Ethics and Innovation (CDEI) to conduct an independent study into the use of data analytics by police forces in England and Wales, with a focus on algorithmic bias. The primary purpose of the project is to inform CDEI's review of algorithmic bias in policing, one of four sectors under examination as part of CDEI's wider review into bias in algorithmic decision-making.[1] CDEI's review will produce recommendations on how to manage bias in algorithmic decision-making, drawing on findings from these four sectors. As part of this review CDEI is working towards a draft framework for the ethical development and use of data analytics tools in policing.

Much commentary has highlighted the potential issues regarding the implementation of advanced analytics in policing, particularly relating to the impact on individual rights.[2] The authors' previous research has drawn attention to the limited evidence base on the efficacy and efficiency of different systems, their cost-effectiveness, their impact on individual rights and the extent to which they serve valid policing aims.[3] Despite these concerns, there remains a significant lack of national guidance or standards regarding the use of data analytics tools in policing, with stakeholders from across the law enforcement community suggesting that this should be addressed as a matter of urgency.[4]

This project aims to address this gap. The purpose of the research is to collate and synthesise existing evidence regarding the police's use of data analytics and associated legal and ethical concerns, and to engage closely with practitioners from across the UK law enforcement

[1] For further information, see Centre for Data Ethics and Innovation (CDEI), 'Interim Report: Review into Bias in Algorithmic Decision-Making', 25 July 2019, <https://www.gov.uk/government/publications/interim-reports-from-the-centre-for-data-ethics-and-innovation/interim-report-review-into-bias-in-algorithmic-decision-making>, accessed 19 January 2020.

[2] For example, see Hannah Couchman, 'Policing by Machine: Predictive Policing and the Threat to Our Rights', Liberty, January 2019; Rashida Richardson, Jason M Schultz and Kate Crawford, 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice', *New York University Law Review*, May 2019; Andrew Guthrie Ferguson, *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement* (New York, NY: NYU Press, 2019).

[3] Alexander Babuta, Marion Oswald and Christine Rinik, 'Machine Learning Algorithms and Police Decision-Making: Legal, Ethical and Regulatory Challenges', *Whitehall Report*, 3-18 (September 2018).

[4] Alexander Babuta and Marion Oswald, 'Data Analytics and Algorithmic Bias in Policing', RUSI Briefing Paper, September 2019.

community, to explore in detail how these technologies are developed and deployed in the field, the legal and ethical concerns arising from their use, as well as processes and policies which could be implemented to address these concerns.

This project makes an important and timely contribution to the academic and policy landscape by providing new insights into police use of data analytics, drawing on first-hand experience of senior officers and government officials, as well as academic and legal expertise. It provides an independent evidence base to inform the development of new policy and guidance for police use of data analytics, specifically in relation to advanced algorithms.

Methodology

This project combined a targeted review of literature focused on data analytics and algorithmic bias with semi-structured interviews and focus groups with respondents from across the UK law enforcement and policymaking community. Twenty-six respondents participated in research interviews for the project: 15 representatives of UK law enforcement organisations; six academic and legal experts; three representatives of regulatory and oversight bodies; and two representatives from civil society organisations. Interviews were conducted in London between June and October 2019, with a number of these taking place via telephone (as indicated). In addition, two roundtable events were held in London in July 2019. The first brought together 16 representatives from the commercial police technology sector and was organised in partnership with techUK, while the second brought together 27 participants from police forces, civil society organisations, government departments, academics and legal experts.[5]

A participatory research approach was chosen due to the value that can be gained from the experience of stakeholders in the assessment of the real-world context, diagnosis of the issues and consideration of policy requirements.[6] Interviews and focus groups were conducted in a semi-structured format, enabling the research team to adopt a broadly consistent line of questioning in each interview, but allowing for flexibility to probe specialised areas of knowledge and experience in respondents. Interview request and guideline letters were sent in advance so that respondents had a clear understanding of the purpose of the project and were able to give their informed consent to the interview.

A purposive, selective sampling strategy was used, whereby participants were identified by who could provide detailed information about the issues under investigation. Participants

[5] Throughout this paper, an anonymised coding system is used to refer to interviewee data. The prefix 'L' is used to refer to law enforcement representatives, while 'A' refers to academic and legal experts, 'R' refers to members of regulatory or oversight bodies, and 'C' to representatives of civil society or campaigning organisations.

[6] Andrea Cornwall and Rachel Jewkes, 'What is Participatory Research?', *Social Science & Medicine* (Vol. 41, No. 12, 1995), pp. 1667–76; Fiona de Londras, 'Participatory Research: Some Provocations for Doctoral Students in Law', in Laura Cahillane and Jennifer Schweppe (eds), *Legal Research Methods: Principles and Practicalities* (Dublin: Clarus, 2016), p. 150.

were selected based on their first-hand knowledge and experience of developing, using or researching data analytics tools in the policing context. A snowball sampling strategy was used, whereby initial interviewees suggested subsequent participants for interview. Data saturation was reached, as indicated by the observation that latter interviews did not produce data that led to any new emergent themes.[7] Interview data was analysed following an inductive grounded theory approach. A preliminary coding process allowed recurring themes to be identified, and then a more granular analysis allowed particular trends and patterns within these themes to be explored in further detail.[8]

This study has several limitations. First, although the authors consulted with respondents from UK-wide organisations, the scope of research is limited to the use of technology by police forces in England and Wales. The findings may not be generalisable to other law enforcement agencies, either in the UK or overseas. Second, the research is limited primarily to police technology projects which are in the (semi-) public domain. It is possible that other tools and capabilities are being developed beyond those discussed in this paper, which were not considered as part of this study.

Definitions and Scope

This paper focuses on police use of data analytics, more specifically the use of algorithms. An algorithm can be defined as 'a set of mathematical instructions or rules that, especially if given to a computer, will help to calculate an answer to a problem'.[9] As summarised by the Law Society, 'given a specified computational problem which generally describes a desired input–output relationship, an algorithm describes a computational procedure which achieves this relationship'.[10] Traditional rules-based algorithms rely on pre-programmed instructions ('parameters') specified by a human user. More complex algorithms are often 'non-parametric', meaning that the machine itself derives the relationship between inputs and outputs.

Machine learning (frequently referred to as 'artificial intelligence' or AI) is a specific category of advanced algorithm that is able to improve at a certain task after being exposed to new data. As summarised by Stuart J Russell and Peter Norvig, 'an agent is learning if it improves

[7] Michelle O'Reilly and Nicola Parker, '"Unsatisfactory Saturation": A Critical Exploration of the Notion of Saturated Sample Sizes in Qualitative Research', *Qualitative Research* (Vol. 13, No. 2, 2013), pp. 190–97; Lisa M Given, *100 Questions (and Answers) About Qualitative Research* (Thousand Oaks, CA: SAGE Publications, 2015), p. 135; Michael P Grady, *Qualitative and Action Research: A Practitioner Handbook* (Arlington, VA: PDK International, 1998), p. 26.

[8] Juliet Corbin and Anselm Strauss, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory,* 3rd Edition (Thousand Oaks, CA: SAGE Publications, 2008).

[9] This definition of 'algorithm' comes from Colin McIntosh, *Cambridge Advanced Learner's Dictionary,* 4th Edition (Cambridge: Cambridge University Press, 2013).

[10] Law Society Commission on the Use of Algorithms in the Justice System and Law Society of England and Wales, 'Algorithms in the Criminal Justice System', June 2019, p. 10.

its performance in future tasks after making observations about the world'.[11] There are three main types of learning: supervised; unsupervised; and reinforcement learning. In supervised learning, the algorithm 'observes some example input–output pairs and learns a function that maps from input to output'.[12] In unsupervised learning, 'the agent learns patterns in the input even though no explicit feedback is supplied'.[13] In reinforcement learning, 'the agent learns from a series of reinforcements – rewards or punishments'.[14] A fourth category – semi-supervised learning – involves datasets where some input–output pairs are labelled but a large proportion are unlabelled.[15]

This paper focuses on algorithms used by the police to derive insights, inform operational decision-making or make predictions. Biometric technology, including live facial recognition, DNA analysis and fingerprint matching, are outside the direct scope of this study, as are covert surveillance capabilities and digital forensics technology, such as mobile phone data extraction and computer forensics.[16] It is not the intention of this project to duplicate the ongoing policy development initiatives being conducted in these areas. However, because many of the policy issues discussed in this paper stem from general underlying data protection and human rights frameworks, these issues will also be relevant to other police technologies, and their use must be considered in parallel to the tools examined in this paper.

Throughout this paper, the terms 'algorithms' and 'data analytics' are used interchangeably, although the use of algorithms can be understood as one component of the wider data analytics process. 'Algorithmic bias' refers to a statistical algorithm that systematically and unfairly discriminates against certain individuals or groups of individuals in favour of others.[17]

[11] Stuart J Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach,* 3rd Edition (Harlow: Pearson Education Limited, 2016), p. 706.

[12] *Ibid*., p. 695.

[13] *Ibid*., p. 694.

[14] *Ibid*., p. 695.

[15] *Ibid*., p. 708.

[16] For further discussion on live facial recognition and biometrics, see Peter Fussey and Daragh Murray, 'Independent Report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology', The Human Rights, Big Data and Technology Project, July 2019; and the independent review of the governance of biometric data currently being conducted by the Ada Lovelace Institute, <https://www.adalovelaceinstitute.org/our-work/identities-liberties/independent-review-of-the-governance-of-biometric-data/>, accessed 13 February 2020.

[17] See Batya Friedman and Helen Nissenbaum, 'Bias in Computer Systems', *ACM Transactions on Information Systems* (Vol. 14, No. 3, 1996), p. 332.

I. Police Use of Data Analytics

The Current Landscape

IN RECENT YEARS, police use of data analytics has expanded significantly in both scale and complexity. This is driven to a great extent by resourcing pressures, and a perceived need to allocate limited resources more efficiently based on a data driven assessment of risk and demand. Despite general cuts to police funding since 2010, specific funding for digital transformation has been made available, such as the Police Transformation Fund, 'creating strong incentives for forces to frame the development around digital technology to receive further central support'.[18]

However, as has been discussed at length elsewhere,[19] a lack of national coordination means that these initiatives remain highly localised, resulting in duplication of efforts, lack of knowledge transfer and poor system interoperability. As described by one senior officer interviewed for this project, 'it's a patchwork quilt, uncoordinated, and delivered to different standards in different settings and for different outcomes'.[20] Another commented that:

> there has been major investment in capability development for digital investigation and intelligence in recent years. But in terms of maturity, we're right at the thin end of the wedge … It's a very patchy and immature landscape. There are a range of tools that we've tried to consider the value, costs, benefits.[21]

Interviewees repeatedly mentioned that while public attention tends to focus on the more novel and innovative data capabilities (such as facial recognition or predictive analytics), the greatest benefit of these technologies lies in the ability to automate time-consuming and resource-intensive data matching or investigative functions, for instance to consolidate information across multiple siloed databases:

> The most successful examples don't use ML [machine learning], they are factual data analytics, but nothing that would involve the human emotion or predictive side of things … My personal view is that it shouldn't be about solving anything or giving the answers, it's pointing people in certain directions.[22]

[18] Law Society Commission on the Use of Algorithms in the Justice System and Law Society of England and Wales, 'Algorithms in the Criminal Justice System', p. 13.

[19] For a more detailed discussion, see Alexander Babuta, 'Big Data and Policing: An Assessment of Law Enforcement Requirements, Expectations and Priorities', *RUSI Occasional Papers* (September 2017).

[20] Author telephone interview with L6, senior police officer, 10 July 2019.

[21] Author telephone interview with L8, senior police officer, 23 July 2019.

[22] Author telephone interview with L12 and L13, police respondents, 15 August 2019.

> We are often seduced by the talk of prediction and facial recognition, but a lot of the more important and perhaps mundane uses are in the background, more to harmonise big databases ... The predictive stuff may be a red herring.[23]

> Predictive policing is like the icing on the cake, but it might not be a very good icing.[24]

Although still requiring legal justification, many of these 'mundane' uses may not carry the same legal and ethical complexities of more sophisticated algorithmic tools such as predictive analytics. As such, there is significant value to be gained from investing in systems to improve data quality and management, and automate core investigative activities, freeing up staff to focus on more complex and demanding analysis tasks.

Driving Factors

Research conducted for this study has found that the implementation of algorithms in policing is driven by three closely related factors: a significant increase in the volume and complexity of digital data that police forces are required to process; significantly reduced resources coupled with ever-increasing demand; and an increased focus on the preventative aspects of the police's role in society.

Driver 1: Information Overload

A significant increase in the volume and complexity of digital data was identified as the primary factor driving the development of policing algorithms. Interviewees described an 'information overload', presenting greater challenges for police forces to effectively trawl large volumes of unstructured data and search across multiple, siloed data systems:

> Data collection has exploded so we're trying to leverage the opportunities that can come from data. A tipping point was reached quite some time ago … I don't think people realise the volumes of data that 21st-century police forces are working with.[25]

> Within the datasets that we do have, we are missing things. There are opportunities for us to be more effective that we're not taking because we are unaware of what's in our data.[26]

> There are clear benefits to these systems if appropriate safeguards are in place. They save police officers huge amounts of time in terms of the amounts of data they are able to analyse … That frees up police officers to be doing arguably what is much more important work.[27]

[23] Author telephone interview with A4, academic expert in policing and criminal justice, 11 July 2019.

[24] Author interview with L14 and L15, senior police technologists, London, 19 September 2019.

[25] Author telephone interview with L4, senior police technologist, 9 July 2019.

[26] Author telephone interview with L1, senior police officer, 1 July 2019.

[27] Author interview with A3, academic legal expert, London, 28 June 2019.

Notably, technological advances did not emerge as a main factor in the increased adoption of policing algorithms, suggesting that innovation is primarily driven by limitations in existing data management processes and systems, rather than new tools becoming available. There was a recurring sentiment that a failure to effectively analyse this data to identify risk and vulnerability could equate to a failure to protect the public from harm. This perceived 'obligation to innovate' is directly linked to the police's public protection duty and positive responsibilities under Articles 2 and 3 of the European Convention on Human Rights (ECHR): right to life and the prohibition of torture and inhuman and degrading treatment respectively. One senior officer made the stark observation that – in the context of domestic abuse – 'almost every victim of domestic homicide was already in our system'.[28]

Driver 2: More Demand, Fewer Resources

Ongoing austerity measures, coupled with a continuous increase in demand on police forces, was described by all officers interviewed as a main driver of the development of new data capabilities:

> We live in an age of austerity, we have been for some time … It's an increasingly complex world that we operate in with growing demand and reducing resources. There's going to be opportunities to be more efficient in that world by using machine learning.[29]

> We have less resources, more complex crime types that require complex responses and more intense responses … We always have to look at how we're going to prioritise our resources, there's always going to be a need to make a risk-based decision about what we're going to apply our resource to, because we can't do everything in its totality.[30]

> The main driver is the increased demand on the force and the need to anticipate demand and deploy resources smarter.[31]

As discussed in a recent report from Cardiff University, these developments are part of a wider trend across the UK public sector of the use of algorithms and 'data scoring' to inform decision-making and resource allocation in an age of austerity:

> A recurring theme in the rationale for implementing data systems is the context of austerity, with managers and developers often responding to significant cuts by trying to use data to better target

[28] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[29] Author telephone interview with L1, senior police officer, 1 July 2019.

[30] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[31] Author telephone interview with L5, senior police technologist, 10 July 2019.

> resources. This speaks to the contextual duality of data driven technologies as one of data-rich and resource-poor contexts.[32]

However, while algorithmic tools present new opportunities for police forces to better manage demand, it was noted that such analysis would inevitably create additional demand, requiring the police to act on insights which may not have been generated otherwise:

> There may be efficiency gains but this has to be caveated. Technology generates extra resource requirements.[33]

> Many of these tools are demand creation, not demand reduction. How do we deal with the output?[34]

> In our view the efficiency that it's affording is a false economy, because ultimately there are aspects that are making the system less efficient. For instance, there is breakdown of communication and the police become less aware of why certain decisions are made.[35]

Driver 3: Increased Focus on Prevention

The police service is increasingly expected to adopt a preventative, rather than reactive, posture, with greater emphasis on anticipating potential harm before it occurs, identifying vulnerable individuals in need of safeguarding and targeting interventions towards the highest-risk persistent and prolific offenders.[36] The NPCC-APCC 'Policing Vision 2025' affirms that 'by 2025 the police service will have transformed the way it delivers its mission with a keen focus on prevention and vulnerability and the effective management of risk'.[37] The recently published NPCC-APCC 'National Policing Digital Strategy' also includes a commitment to 'translate evolving definitions of threat, harm and risk into digital formats that complement human judgement' and 'use digital tools to rapidly identify harm related behaviours in order to target interventions.'[38] Interviewees described how the use of algorithms can inform a more proactive deployment of interventions towards particular areas or individuals identified as posing the greatest 'risk', or towards the highest-priority victims or potential victims:

[32] Lina Dencik et al., 'Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services', Cardiff University, December 2018, p. 116.

[33] Author telephone interview with A4, academic expert in policing and criminal justice, 11 July 2019.

[34] Author interview with L14 and L15, senior police technologists, London, 19 September 2019.

[35] Author telephone interview with C2, representative of civil society organisation, 17 October 2019.

[36] For further discussion, see Adam Crawford and Karen Evans, 'Crime Prevention and Community Safety', in Alison Liebling, Shadd Maruna and Lesley McAra, *The Oxford Handbook of Criminology,* 6th Edition (Oxford: Oxford University Press, 2017).

[37] National Police Chiefs' Council (NPCC) and Association of Police and Crime Commissioners (APCC), 'Policing Vision 2025', 2017, p. 5.

[38] NPCC and APCC, 'National Policing Digital Strategy: Digital, Data and Technology Strategy 2020–2030', January 2020, p. 7.

> Management of risk is becoming more and more prominent in lots of different areas of policing business … We're much more focused on managing risk rather than simply catching the people who are involved in volume crime.[39]
>
> Even a local beat manager will have thousands of offenders living in their area. Which ones are you going to prioritise? You need help with that, otherwise you're continuously desk bound and doing research on different systems.[40]
>
> There's real room for that sort of tech to better identify high-risk, better screen out high-volume, low-risk where we don't need to prioritise resources, and it enables us to make better decisions and push our resources in the area of greatest need.[41]

However, it was also noted that the police are increasingly required to respond to non-crime problems where previously, another agency would have led the response. One senior police officer described a 'need due to austerity to step into the space vacated by other partners, for instance mental health'.[42] Previous reports have expressed concern that the use of new data analytics technologies risks broadening the police's role into areas of social and public policy with unclear justification.[43] This issue is explored further in Chapter II.

Issues and Limitations

While the research highlighted the potential of new data-driven technologies to improve police effectiveness and efficiency, concerns were raised regarding their development and operational implementation. The lack of a robust empirical evidence base, poor data quality and insufficient skills and expertise were identified as three major barriers to successful implementation.

Evidence Base

Research conducted for this study has found that the development of policing algorithms is often not underpinned by a robust empirical evidence base regarding their claimed benefits, predictive accuracy, scientific validity or cost effectiveness. Furthermore, capability development is largely driven by data science, with comparatively little focus on the underlying criminological theory, legal requirements or conceptual framework on which the technology is based:

[39] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[40] Author telephone interview with L4, senior police technologist, 9 July 2019.

[41] Author telephone interview with L6, senior police officer, 10 July 2019.

[42] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[43] Alan Turing Institute Data Ethics Group and Independent Digital Ethics Panel for Policing, 'Ethics Advisory Report for West Midlands Police', July 2017.

> Often the implementation and adoption of policing technology is done at quite a rapid pace, and what is often missing in those cases is a testing and assessment of the risks and long-term benefits that these systems may provide.[44]

> There is a risk that we procure capabilities based on opinion-based decisions rather than evidence-based decisions.[45]

> One of the reasons why the police want to see these tools work is that they want to save time and resource. But when you're focused on saving time and resource, there isn't going to be enough space to say, well actually, what we need is a 10-year research project to assess how algorithms and humans interact, and how we incorporate these tools into our criminal justice system.[46]

Data science-driven projects tend to be focused on allowing the data to lead to conclusions based on statistical correlations. The importance of causal justification, an underlying criminological or other evidentially informed theory behind the tool, and the legal relevance of the data inputs to the decision informed by the output, can all be overlooked. Although a data-driven project can correctly generate outputs that may suggest causation, without appropriate operational insight, data-driven projects might draw incorrect inference from the data outputs.

Data Quality

Concerns were also raised regarding the quality and integrity of police data, which may not lend itself to algorithmic analysis:

> There needs to be earlier consideration of data quality and data input … data quality is the biggest risk – if you have poor data in the first place, you are in big trouble.[47]

> There's always issues around data quality, what you are and are not able to collect reliably.[48]

Interviewees stressed the importance of context when interpreting the reliability of police-recorded information, and the need for strong processes for data hygiene to ensure timely removal of inaccurate or misleading data.[49] Conversely, interviewees also noted that the use of algorithms could result in the secondary benefit of causing police forces to pay closer attention to data quality:

[44] Author interview with A3, academic legal expert, London, 28 June 2019.

[45] Author telephone interview with L1, senior police officer, 1 July 2019.

[46] Author telephone interview with C2, representative of civil society organisation, 17 October 2019.

[47] Author interview with L14 and L15, senior police technologists, London, 19 September 2019.

[48] Author interview with R3, member of regulatory/oversight body, London, 3 October 2019.

[49] Author telephone interview with A5, academic expert in policing and criminal justice, 11 July 2019.

> If use of algorithmic tools results in greater understanding of the data, this could of itself be a benefit, as it could result in reduction of errors and increase in quality of data.[50]
>
> Even if none of the advantages come from the system itself, having a system like this forces you to keep track of the data you have, hopefully use it in decision-making, hopefully take an interest in data quality. That's aside from whether ML is helping you to predict, classify or prioritise things.[51]

Skills and Expertise

Another common barrier to the successful deployment of data-driven policing tools is lack of access to skills and expertise, not just for the development of systems, but also for ongoing maintenance, validation, review and testing:

> The big issue in policing is not the technology, it's what the military call the 'capability stack', the combination of the technology, the people and the processes that need to be considered … There's still a long way to go because we're not considering all three.[52]
>
> There's resourcing issues. To properly test, validate and revalidate a tool requires continuing resource and expertise, and that's not always in place.[53]

As such, while police forces often develop algorithmic tools with the objective of making more efficient use of limited resources, implementation will inevitably require investing in additional resources in the form of technical, legal and academic expertise.

Use Cases

Recent reports scrutinising police use of data analytics have focused on predictive analytics and algorithmic risk-assessment tools, collectively referred to as 'predictive policing' technology. In recent years, a number of police forces have developed sophisticated algorithmic tools for forecasting demand and assessing risk, which are now used for a diverse range of purposes.

Several interviewees noted that the term 'predictive policing' is potentially problematic. Many uses of advanced analytics focus on categorising and classifying entities into different groups – for instance, 'risk scoring' offenders according to their perceived likelihood of re-offending by comparison with selected characteristics of a specified group. While described as predictive, these algorithms are typically implemented as prioritisation tools – a higher 'risk score' does not necessarily imply that an individual is expected to commit crime; rather, the level of risk management required is judged to be greater than for other individuals within the same cohort. As described by one police respondent interviewed, 'classification and prediction

[50] Author telephone interview with A1, academic policing expert, 1 June 2019.

[51] Author interview with A2, academic legal expert, London, 28 June 2019.

[52] Author telephone interview with L8, senior police officer, 23 July 2019.

[53] Author interview with R3, member of regulatory/oversight body, London, 3 October 2019.

is often the same thing. If you've got a question where you're trying to calculate a probability, you are classifying and then applying that predictively'.[54]

With this in mind, it may be more accurate to describe these systems as classification tools – and not as risk assessments – with the risk-assessment element clearly specified as the human decision, taking into account the algorithmic classification together with the context and other relevant information.

Predictive Mapping

As shown in Figure 1, predictive crime mapping involves the use of statistical forecasting applied to crime data to identify locations where crime may be most likely to happen in the near future. The use of such technology can be traced back to at least 2004,[55] and recent data suggests that as many as 12 (of 43) police forces in England and Wales are currently using or developing such systems, or have done so in recent years.[56] The technique is based on the observation that repeat and 'near-repeat' victimisation accounts for a large proportion of all crime, and that crime is often 'contagious', with the risk of property crimes such as burglary greatly increasing for households near to the burgled property in the aftermath of the initial offence.[57] The empirical basis for predictive mapping has been discussed at length elsewhere, and its use is widely advocated by academic criminologists.

In summary, random foot patrolling has a negligible impact on detecting and preventing crime, because crime is not uniformly distributed in time and space. By contrast, 'hotspot' policing – whereby high-risk locations are identified and patrol resources are concentrated in those areas – has been shown to result in crime suppression not just at the deployment location but also in the surrounding areas. Various randomised control trials have demonstrated that the correct use of predictive mapping software consistently increases the likelihood of detecting future crime events, resulting in net reductions in overall crime rates.[58]

[54] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[55] Kate J Bowers, Shane D Johnson and Ken Pease, 'Prospective Hot-Spotting: The Future of Crime Mapping?', *British Journal of Criminology* (Vol. 44, No. 5, 2004), pp. 641–58; Shane D Johnson et al., 'Prospective Crime Mapping in Operational Context: Final Report', Home Office, Online Report 19/07, 2007.

[56] Couchman, 'Policing by Machine'.

[57] Graham Farrell and Ken Pease (eds), *Repeat Victimization* (Monsey, NY: Criminal Justice Press: 2001).

[58] For example, see Anthony A Braga and Brenda J Bond, 'Policing Crime and Disorder Hot Spots: A Randomized Controlled Trial', *Criminology* (Vol. 46, No. 3, 2008), pp. 577–607; Johnson et al., 'Prospective Crime Mapping in Operational Context'; Rob T Guerette and Kate J Bowers, 'Assessing the Extent of Crime Displacement and Diffusion of Benefits: A Review of Situational Crime Prevention Evaluations', *Criminology* (Vol. 47, No. 4, 2009), pp. 1331–68; College of Policing, 'The Effects of Hot-Spot Policing on Crime: What Works Briefing', September 2013; George Mohler et al., 'Randomized Controlled Field Trials of Predictive Policing', *Journal of the American Statistical Association* (Vol. 110, No. 512, 2015), pp. 1399–411.

One officer interviewed summarised the potential benefits of predictive mapping as follows:

> If you have a capability to identify what kind of event is likely to occur where, you can then decide where your response officers should be at different times. Then you've got a tool that helps your senior officers plan ahead as to where to deploy resources. This could help greatly with resource allocation in real time on the ground.[59]

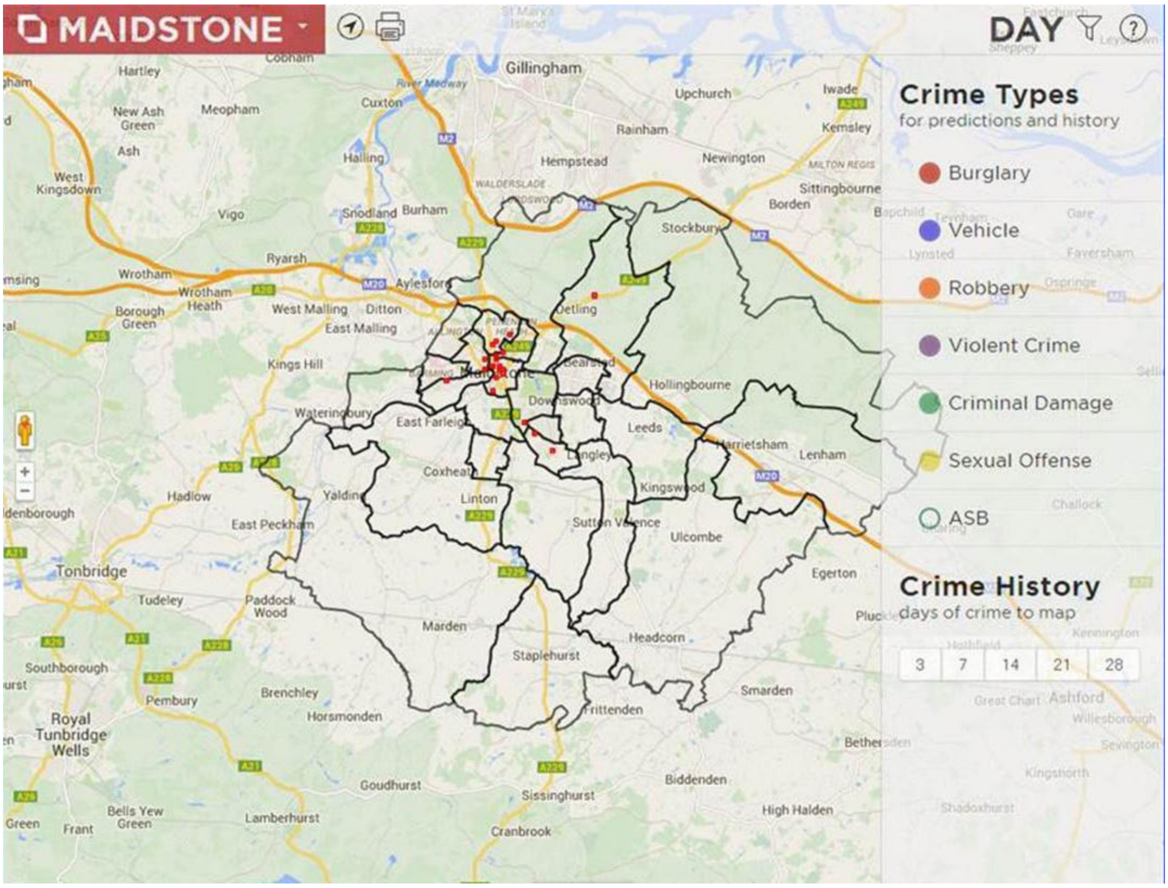
Numerous organisations have urged UK police forces to make better use of predictive mapping to enable more evidence-based deployment of resources.[60] However, predictive mapping has been the subject of considerable criticism, particularly regarding the risk of bias and discrimination. A 2019 report from Liberty recommended that 'police forces in the UK should end their use of predictive policing "mapping" programs, which rely on problematic historical arrest data and encourage the over-policing of marginalised communities'.[61] These concerns are discussed further in Chapter III.

[59] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[60] Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS), 'PEEL: Police Effectiveness 2016 – A National Overview', 2017; Europol, *Serious and Organised Crime Threat Assessment: Crime in the Age of Technology* (The Hague: Europol, 2017), p. 25; London Assembly, Budget and Performance Committee, *Smart Policing: How the Metropolitan Police Service Can Make Better Use of Technology* (London: City Hall, 2013).

[61] Couchman, 'Policing by Machine', p. 10.

Figure 1: Screenshot of the Predictive Crime Hotspot Mapping Software Used by Kent Police



Source: Lizzie Dearden, 'How Technology is Allowing Police to Predict Where and When Crime Will Happen', The Independent*, 7 October 2017.*

Individual Risk Assessment

Algorithms can also be applied to individual-level personal data to assess the risk of future offending. An actuarial risk assessment instrument (ARAI) is a statistical model which uses pre-defined 'risk factors' to assign individuals numerical scores corresponding to their predicted likelihood of future offending.[62] ARAIs based on simple rules-based algorithms have been used for offender risk assessment for over two decades.[63] As Leam A Craig and Anthony Beech

[62] For further discussion, see Vernon L Quinsey et al., *Violent Offenders: Appraising and Managing Risk,* 2nd Edition (Worcester, MA: American Psychological Association, 2006).

[63] See Mary Ann Campbell, Sheila French and Paul Gendreau, 'The Prediction of Violence in Adult Offenders: A Meta-Analytic Comparison of Instruments and Methods of Assessment', *Criminal Justice and Behavior* (Vol. 36, No. 6, 2009), p. 569.

describe, 'in North America and the United Kingdom, actuarial risk assessment has permeated the entire criminal justice system'.[64]

In England and Wales, the most commonly used tools are the Offender Assessment System (OASys) and the Offender Group Reconviction Scale (OGRS), which are routinely used by Her Majesty's Prison and Probation Service to measure individuals' likelihood of re-offending and to develop individual risk-management plans.[65] OASys incorporates both static risk factors (such as age and criminal history) and dynamic risk factors (such as accommodation, employment and substance use), allowing progress and changes in offender behaviour to be monitored over time.[66] OGRS includes only a limited range of static risk factors (age, gender and criminal history), and can therefore be used on a wider group of offenders than OASys (for instance, in situations where it is not possible to complete a more thorough assessment that includes socio-economic and personal risk factors).[67] As described by one interviewee who was directly involved in the development of OASys and OGRS:

> There is an established 'what works' evidence base for working with offenders … You should target more of your resources to high-risk offenders, there's lots of research evidence around that … We developed actuarial tools because they outperform structured professional judgement, which we've shown through various empirical studies, certainly for general re-offending.[68]

It is important to note that experts continue to disagree over the statistical validity of individual risk-assessment tools. On the one hand, there is strong empirical evidence that the use of statistical methods improves the overall accuracy of professional ('clinical') judgements in a range of decision-making contexts, including offender risk assessment.[69] On the other hand, it has been argued that aggregated 'predictive accuracy' rates are fundamentally misleading when assessing risk judgements at the individual level.[70] The arguments are statistically complex and it

[64] Leam A Craig and Anthony Beech, 'Best Practice in Conducting Actuarial Risk Assessments with Adult Sexual Offenders', *Journal of Sexual Aggression* (Vol. 15, No. 2, 2009), p. 197.

[65] Robin Moore (ed.), 'A Compendium of Research and Analysis on the Offender Assessment System (OASys), 2009-2013', Ministry of Justice Analytical Series, July 2015.

[66] *Ibid*., p. 3.

[67] *Ibid*., p. 153.

[68] Author interview with R3, member of regulatory/oversight body, London, 3 October 2019.

[69] For example, see William M Grove et al., 'Clinical Versus Mechanical Prediction: A Meta-Analysis', *Psychological Assessment*, (Vol. 12, No. 1, 2000), pp. 19–30; Stefania Ægisdóttir et al., 'The Meta-Analysis of Clinical Judgment Project: Fifty-Six Years of Accumulated Research on Clinical Versus Statistical Prediction', *Counseling Psychologist* (Vol. 34, No. 3, 2006), pp. 341–82; Mary Ann Campbell, Sheila French and Paul Gendreau, 'The Prediction Of Violence In Adult Offenders', *Criminal Justice and Behavior* (Vol. 36, No. 6), p. 569.

[70] Alan A Sutherland et al., 'Sexual Violence Risk Assessment: An Investigation of the Interrater Reliability of Professional Judgments Made Using the Risk for Sexual Violence Protocol', *International Journal of Forensic Mental Health* (Vol. 11, No. 2, 2012), p. 120.

is beyond the scope of this paper to discuss this issue in further detail, but suffice to say that the debate about which approach is more accurate, justified or informative is intense and ongoing.[71]

The College of Policing's Authorised Professional Practice notes that 'by definition, decisions involve uncertainty, i.e., the likelihood and impact of possible outcomes cannot be totally predicted, and no particular outcome can be guaranteed'.[72] The first principle on risk is that 'the willingness to make decisions in conditions of uncertainty (i.e., risk taking) is a core professional requirement of all members of the police service'.[73] As summarised by officers interviewed:

> In so many situations, the police need to make risk-based decisions based on what they're presented with and what they see as part of their core policing role … The police need to be satisfied when they make their initial risk-based decision that they're making that decision with the best possible information they can.[74]

> Policing is about dealing with complexity, ambiguity and inconsistency… Sometimes you make decisions based on the most challenging information, but sometimes that's the only information you've got.[75]

With this in mind, in order to assess whether use of a particular algorithmic tool is justified for offender assessment purposes, the focus should not be on the 'predictive accuracy' of the statistical forecast. Rather, the question should be whether the tool provides useable insights which enhance the officer's ability to make an informed professional judgement in conditions of uncertainty.

As summarised by one interviewee, 'It's easy to criticise an actuarial approach but the question is what's the alternative?'.[76]

Police Use of Data-Scoring Tools

While the use of simple rules-based algorithms for offender risk assessment is now a well-established practice, a more recent development is the use of advanced machine learning

[71] For further discussion, see Stephen D Hart and David J Cooke, 'Another Look at the (Im-)Precision pf Individual Risk Estimates Made Using Actuarial Risk Assessment Instruments', *Behavioral Sciences and the Law* (Vol. 31, No. 1, 2013), pp. 81–102; David J Cooke and Christine Michie, 'The Generalizability of the Risk Matrix 2000: On Model Shrinkage and the Misinterpretation of the Area Under the Curve', *Journal of Threat Assessment and Management* (Vol. 1, No. 1, 2014), p. 42; Royal Statistical Society, 'Algorithms in the Justice System: Some Statistical Issues', evidence submitted to the Law Society Public Policy Commission on use of algorithms in the criminal justice system, November 2018.

[72] College of Policing, 'Authorised Professional Practice: Risk'.

[73] *Ibid*.

[74] Author telephone interview with L6, senior police officer, 10 July 2019.

[75] Author telephone interview with L8, senior police officer, 23 July 2019.

[76] Author interview with R3, member of regulatory/oversight body, London, 3 October 2019.

algorithms applied to police data to generate 'risk' scores of known offenders. Examples include Durham Constabulary's Harm Assessment Risk Tool (HART), Avon and Somerset Constabulary's Qlik Sense data visualisation and risk-modelling system, Hampshire Constabulary's domestic violence risk-forecasting model and West Midlands Police's (WMP) draft Integrated Offender Management model.

Durham's HART algorithm uses random forest forecasting (a form of supervised machine learning) to assign offenders into low-, medium- or high-risk groups corresponding to their predicted likelihood of re-offending over a 24-month period. This score is based on 34 predictor variables, including age, residential location and 29 factors relating to previous criminal history. The system is used to assess offenders' eligibility to participate in Durham's Checkpoint scheme, an out-of-court disposal designed to reduce re-offending by addressing the underlying factors causing individuals to engage in crime. The risk score is intended to be used as one of several factors for the human officer to consider when making their overall risk assessment, thereby enabling more targeted allocation of offender intervention programmes and, in turn, improving their overall success rate.[77]

Other police forces are exploring similar technology. Domestic violence is an area of policing substantially driven by a formal risk-assessment tool, the Domestic Abuse, Stalking and Honour-Based Violence Risk, Identification, Assessment and Management Model ('DASH'). Recent research has suggested that existing risk-assessment methods such as DASH are 'not enabling police forces to identify high-risk revictimization or recidivism cases'.[78] The Hampshire Constabulary tool currently in development aims to use a machine learning model to improve the current risk-assessment process by providing an additional perpetrator-based risk classification.[79]

Avon and Somerset Constabulary's Qlik Sense is a form of self-service analytics software that connects the force's own internal databases and other local authority datasets, and applies predictive modelling to produce individual risk-assessment and intelligence profiles, to assist the force in triaging offenders according to their perceived level of risk (see Figure 2). As described in detail in Cardiff University's report on citizen scoring, the system is used to assess individuals

[77] Author telephone interview with L7, senior police respondent, 18 July 2019; Chris Baraniuk, 'Durham Police AI to Help with Custody Decisions', *BBC*, 10 May 2017; Sheena Urwin, 'Algorithmic Forecasting of Offender Dangerousness for Police Custody Officers: An Assessment of Accuracy for the Durham Constabulary Model', unpublished thesis, University of Cambridge, 2016, <https://www.crim.cam.ac.uk/global/docs/theses/sheena-urwin-thesis-12-12-2016.pdf/at_download/file>, accessed 19 January 2020.

[78] Emily Turner, Juanjo Medina and Gavin Brown, 'Dashing Hopes? The Predictive Accuracy of Domestic Abuse Risk Assessment by Police', *British Journal of Criminology* (Vol. 59, No. 5, 2019), p. 1028.

[79] Petros Terzis, Marion Oswald and Christine Rinik, 'Shaping the State of Machine Learning Algorithms Within Policing', workshop report, University of Winchester, June 2019.

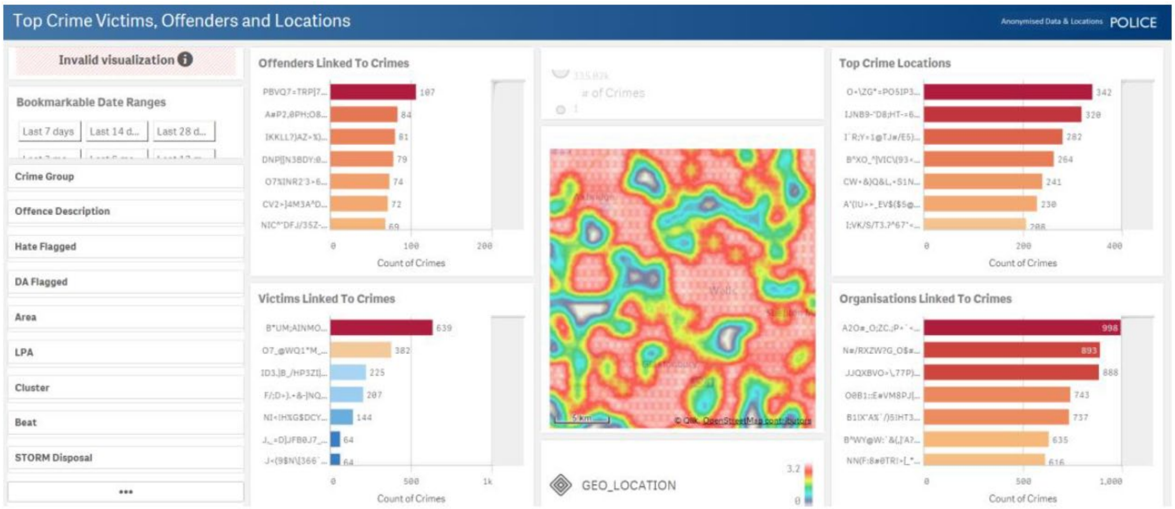
for likelihood of offending, as well as risk of offending escalation, risk of becoming a victim of crime and risk of going missing.[80]

One police interviewee provided further details regarding how Qlik is implemented in practice:

> We score everyone on the system daily (over 250,000 offenders within our crime and intelligence database). They're getting scored every day on their likelihood of re-offending, and we're also introducing the concept of crime harm ... It's a screening tool to help read across massive volumes of data. I see it as being the equivalent of your intelligence analyst saying 'Hey, I've identified these people who might be of interest to you', but doing that en masse across your data. Our intelligence and live-cell colleagues need to know on a daily basis who the high-risk offenders are. The whole point of leveraging an algorithmic take on risk is that we can score 250,000 offenders every single day. You need to draw on an algorithm to make sense of that.[81]

In addition to Qlik's offender scoring functionality, the interviewee also described a 'crime data integrity model' to 'support us with our crime recording and to support our ethical crime recording approach'.[82] The predictive models used for Qlik are subject to ongoing empirical validation: 'We partition the data, we train the model, we test it … We re-visit our models on a quarterly basis to ensure the accuracy is being maintained … We won't deploy a model that's obviously not working, that isn't accurate and adding value.'[83]

Figure 2: Screenshot of the Qlik Sense Data Visualisation Tool Used by Avon and Somerset Police



Source: Dencik et al., 'Data Scores as Governance', p. 78.

[80] Dencik et al., 'Data Scores as Governance', pp. 74–81.

[81] Author telephone interview with L4, senior police technologist, 9 July 2019.

[82] *Ibid*.

[83] *Ibid*.

While not yet implemented operationally, WMP have developed a pilot machine learning algorithm for assessing 'offender escalation', in order to identify individuals who are likely to transition from a state of low- to high-harm offending and deliver interventions to prevent them from doing so. The initial consideration of the first iteration of the model by the WMP data ethics committee included questions around the proposed use of intelligence as input data and how the model would be used operationally – for instance, 'what interventions might be applied to those individuals identified, bearing in mind that potential adverse consequences of inaccurate predictions will be largely dependent on the type of intervention carried out?'.[84]

Beyond the uses described above, advanced algorithms are also used for complex investigative tasks. Two interviewees involved in serious crime investigations described how automation of bulk-data analysis has provided significant efficiency savings when investigating online child sex abuse and exploitation (CSAE). The interviewees described how their agency has developed a machine learning 'prioritisation tool' to identify high-risk users of child abuse forums who may have the potential to harm a child:

> We've got a massive amount of data from closed cases and known offenders that hasn't been explored for other intelligence, offences and connections between individuals … We're systematically re-processing that data to identify new offences and to identify behaviours … We're looking at the entire dataset, identifying further victims, finding links, understanding how offenders communicate.[85]

The algorithm uses semantic keyword matching in combination with machine learning risk assessment to identify high-risk users who should be subject to more detailed, manual examination. The participants described how they have seen a 'quantifiable increase in efficiency', explaining how previously it would take up to two months for an officer to manually risk assess each forum user, by reading each of their forum posts. This manual risk-assessment process was used to train a machine learning algorithm to identify risk factors which could then be used predictively using a support vector machine learning model. Not only has the system led to efficiency improvements in what was previously a 'highly resource-intensive' process, but it has led to new users being identified who were not previously known.[86]

Complex algorithms are also used by the police for numerous other purposes, including to forecast demand in control centres,[87] and triage crimes for investigation according to their predicted 'solveability'.[88] But despite these diverse applications, there are a number of fundamental cross-cutting legal and ethical considerations arising from the application of complex algorithms to police data. This is the focus of the following chapter.

[84] West Midlands Police and Crime Commissioner, 'Ethics Committee Reports and Minutes', April 2019, <https://www.westmidlands-pcc.gov.uk/ethics-committee/ethics-committee-reports-and-minutes/>, accessed 19 January 2020.

[85] Author telephone interview with L9 and L10, police investigator and analyst, 1 August 2019.

[86] *Ibid*.

[87] Author telephone interview with L4, senior police technologist, 9 July 2019.

[88] See HMICFRS, 'PEEL: Police Effectiveness 2017 – An Inspection of Kent Police', March 2018.

II. Legal Considerations

THIS CHAPTER DISCUSSES legal and ethical concerns identified in the research, organised by reference to the legal framework(s) concerned. For reasons of space, it is not possible to include a full description of all relevant legal frameworks, although the reader is directed to the authors' previous publications for further discussion.[89] Bias and discrimination are discussed in further detail in Chapter III.

A focus on 'data ethics' may distract from fundamental questions regarding the underlying legal basis for police use of algorithms, and whether use of a new analytical tool would constitute lawful exercise of police powers. As summarised by one civil society representative, 'We need to move away from discussions that are about ethics, and move towards thinking about objective human rights standards'.[90] It was noted that there are clearly developed human rights standards that would apply well to the policing context, rather than focusing on nebulous data-ethics principles:

> We hear a lot of talk about things like fairness, accountability and transparency. All of that is important, but they are bypassing a discussion that needs to be had in the first instance, which is 'are these tools rights compliant?' and if they aren't rights compliant, they shouldn't be used in the policing context.[91]

Various legal frameworks and codes of practice are relevant to the use of policing algorithms in England and Wales. However, concerns were raised that these legal requirements are not being considered at the outset of police technology projects:

> While there's an awareness of the relevant legal and governance requirements, there's still a significant delay in bringing the relevant external legal experts in at the early stages of these projects. There needs to be at the very beginning of these projects consultation with whoever has been assigned the competent data protection officer within that police agency.[92]

Data Protection

The main concern raised in relation to data protection was that the Data Protection Act (DPA) 2018 does not provide sufficient protection from automated decision-making. Echoing a previous recommendation from the Law Society,[93] one participant noted the risk that officers

[89] Babuta, Oswald and Rinik, 'Machine Learning Algorithms and Police Decision-Making'; Babuta and Oswald, 'Data Analytics and Algorithmic Bias in Policing'.

[90] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[91] Author telephone interview with C2, representative of civil society organisation, 17 October 2019.

[92] Author interview with A3, academic legal expert, London, 28 June 2019.

[93] The Law Society Commission on the Use of Algorithms in the Justice System and the Law Society of England and Wales, 'Algorithms in the Criminal Justice System', p. 6.

may in practice defer decision-making responsibility to an algorithm and not challenge the outputs, and called for an amendment to Part 3 of the DPA to specify that meaningful human input should be a legal requirement when an algorithmic tool informs a decision-making process which has a legally significant or similarly significant effect on an individual, applicable to many algorithms within policing:

> The Data Protection Act doesn't put safeguards in place for meaningful human intervention. It has to be meaningful, it can't just be a box-tick. Police forces cannot be making these serious decisions on people on an automated basis.[94]

Another concern relates to the police's access to third-party data from organisations such as local authorities and social services, with one roundtable participant describing how:

> Predictive analytics works best when you're able to build them on a wide variety of relevant data … At the moment it's hard to do that because of the data-sharing issue ... Projects that we've really struggled with, it's primarily because we haven't been able to get the data at the right time.[95]

A civil society representative noted that 'there are potentially serious problems with using other data, healthcare, social services, local authorities. The problem here isn't so much around bias and discrimination, but it raises very serious data protection questions'.[96]

Human Rights

Much police activity inevitably involves infringing on citizens' civil liberties and human rights. In cases where there is judged to be some level of intrusion, the authority must be able to demonstrate that such intrusion is in accordance with the law, and necessary and proportionate in the interests of public safety or for the prevention of disorder or crime.

Police interviewees pointed to obligations under Article 2 of the European Convention on Human Rights, which can imply a positive obligation to take preventative operational measures to protect a person whose life is at risk:[97]

> From a human rights point of view, the safety and security of citizens is the first responsibility of the state, and algorithms could contribute to this responsibility.[98]

In relation to the right to fair trial and the presumption of innocence (Article 6 ECHR), one civil society representative suggested that 'there's a serious problem around the reversal of the

[94] Author interview with C1, representative of civil society organisation, London, 7 October 2019.

[95] Roundtable event organised in partnership with techUK, London, 22 July 2019.

[96] Author interview with C1, representative of civil society organisation, London, 7 October 2019.

[97] See Osman v United Kingdom, European Court of Human Rights, 87/1997/871/1083, 1998.

[98] Author telephone interview with A4, academic expert in policing and criminal justice, 11 July 2019.

presumption of innocence'.[99] It was further noted that machine learning algorithms often work by detecting associations and connections, which could engage Article 10 and 11 ECHR (right to freedom of expression, and right to freedom of assembly and association, respectively).[100] One civil society representative suggested that 'things that aren't even criminal in nature, people may feel intimidated out of when they know the police are assessing them in detail. Things like attending protests, expressing political dissent'.[101]

Where there is a potential intrusion into an individual right due to collection and use of personal data or the use of an algorithmic output, this raises complex questions regarding necessity and proportionality. In some cases, it could be argued that the use of such technology would not be 'necessary' if the police force had the necessary resources to deploy a less intrusive, non-technological solution. One senior police officer asked: 'Should we be potentially infringing people's rights to be more efficient, when we arguably wouldn't need to be if we had more resources?'.[102]

In relation to proportionality, the collection of large volumes of data to build a complex algorithm may not be proportionate in relation to the benefits offered by the tool. As summarised by one civil society representative, 'To use these predictive policing programmes, we obviously need huge amounts of data, which runs contrary to the principle of data minimisation'.[103] In some cases, the developer of a system may seek to collect more and more data to correct errors and deficiencies in the system, potentially resulting in indiscriminate or disproportionate data collection.[104]

The Role of the Police

Some have argued that the use of certain new technologies risks broadening the police service's role into other areas of social policy, with unclear justification. In their review of the National Analytics Solution (NAS), a project to trial predictive analytics techniques for policing, the Alan Turing Institute Data Ethics Group and the Independent Digital Ethics Panel for Policing concluded that:

> We see the NAS as moving law enforcement away from its traditional crime-related role and into wider and deeper aspects of social and public policy. This move requires explanation, justification and legitimation, especially where the ethical dimensions and principles of such policing roles are not well established.[105]

[99] Author interview with C1, representative of civil society organisation, London, 7 October 2019.

[100] Author telephone interview with A4, academic expert in policing and criminal justice, 11 July 2019.

[101] Author telephone interview with C2, representative of civil society organisation, 17 October 2019.

[102] Author telephone interview with L1, senior police officer, 1 July 2019.

[103] Author telephone interview with C2, representative of civil society organisation, 17 October 2019.

[104] Author interview with A2, academic legal expert, London, 28 June 2019.

[105] Alan Turing Institute Data Ethics Group and Independent Digital Ethics Panel for Policing, 'Ethics Advisory Report for West Midlands Police'.

The NAS project team responded that they did not believe that this characterisation reflected the nature of the modern police service:

> While the core function of the Police is the prevention and detection of crime there are numerous areas where we invest resources not directly linked to this such as – inter alia – locating missing persons, dealing with people in crisis with mental health issues (often as a joint team with health services), road traffic accidents including fatalities where there is no element of criminal activity, dealing with the homeless, responding to suicide, domestic abuse where there is no recordable crime and dealing with anti-social behaviour issues that do not constitute crime but nevertheless have the ability to have a significant impact on the quality of life of the public affected by them.[106]

In general, interview data affirmed the latter characterisation:

> In reality, detecting and preventing crime is quite a small part of our business … We do many other things, and for a long period of time, a lot of the other stuff that isn't pure prevention and detection of crime has been a holistic approach together with other partners.[107]

> It's really difficult to separate the police role from the role of other agencies like children's services … We need to come together with shared datasets and shared priorities to use this technology better together.[108]

While some may argue that – in an ideal world – the police would not be required to intervene in areas of social policy which do not have a direct criminal element, in reality a large proportion of all police time is spent responding to public safety problems, and this situation is set to persist for the foreseeable future. However, further evidence is needed to demonstrate that the use of algorithms would enable the police to carry out these duties more effectively and proportionately.

Transparency, Explainability and Accountability

Much commentary has highlighted the risks of so-called 'black box' machine learning methods which offer little explanation as to how they arrived at their outputs.[109] As summarised by the Law Society:

[106] West Midlands Police, 'National Analytics Solution Project Team: Response to the Alan Turing Institute and IDEPP', <https://idepp.org/NAS-project-team-response.pdf>, accessed 19 January 2020.

[107] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[108] Author telephone interview with L6, senior police officer, 10 July 2019.

[109] For further discussion, see Lilian Edwards and Michael Veale, 'Slave to the Algorithm? Why a "Right to an Explanation" is Probably Not the Remedy You Are Looking For', *Duke Law & Technology Review* (Vol. 16, No. 1, 2017), p. 18.

> Machine learning systems, in contrast to rule-based systems, are not designed with human interpretability in mind, but optimised instead for connection of input and output data with little regard for the comprehensibility of such connections.[110]

Interviewees recognised a lack of transparency as one of the most significant risks of using machine learning algorithms for policing:

> It's harder for ML, for that we just get a 'medium-risk' score, we can't really identify why that's medium risk.[111]

> You can interrogate the human, but it's much harder to interrogate the technology.[112]

An expert who played a major role in the development and evaluation of OASys explained that machine learning has never been incorporated into OASys or OGRS, primarily due to concerns regarding the 'black box' problem: 'to use machine learning, the predictive validity would have to be significantly higher to justify losing that explainability'.[113]

However, while machine learning is often referred to as 'black box', different methods vary considerably in their transparency, which has implications for what type of algorithm may be suitable in a particular policing context. Supervised machine learning involves explicitly defining input variables and output (target) variables, and training an algorithm to learn how to 'map' a function from the input to the output. As such, the factors taken into account when making the prediction are explicitly defined by the user. Unsupervised learning falls broadly into two categories. The first could be defined as 'exploratory' unsupervised learning, and does not specify an output or target variable, but rather attempts to model the underlying distribution in the data through clustering or association (such as k-nearest neighbours classification). The second category involves the use of unsupervised methods to create features to then be used in a supervised learning task (such as feed-forward auto-encoders built on artificial neural networks). For example, this latter category could include the use of an unsupervised autoencoder to build new features based on complicated combinations of simpler, human-interpretable features (such as age, gender, postcode, and so forth). The resulting features could then be used as input variables for a supervised learning algorithm, such as those designed to assess offending risk. However, in this scenario, the specific factors taken into account when deriving the final 'risk score' may be impossible to discern, due to the use of an unsupervised auto-encoder.

An academic expert suggested that 'essentially, best practice would be to veer away from the use of deep neural networks, in the sense that there is a significant decision-making stage that

[110] The Law Society Commission on the Use of Algorithms in the Justice System and the Law Society of England and Wales, 'Algorithms in the Criminal Justice System', p. 21.

[111] Author telephone interview with L9 and L10, police investigator and analyst, 1 August 2019.

[112] Comment made by a civil society representative at a roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[113] Author interview with R3, member of regulatory/oversight body, London, 3 October 2019.

we can't account for. That is deeply problematic'.[114] Another commented that 'I don't know how you would assess fairness in unsupervised learning. I can't think of a predictive policing application where unsupervised method would improve things because there is always a label on the offence'.[115] For this reason, it would seem inappropriate to employ unsupervised machine learning methods when developing so-called 'data scoring' algorithms related to individuals in a policing context. Human-interpretable features are essential to provide sufficient transparency regarding what factors were taken into account to arrive at a certain outcome, assess whether the process was discriminatory, relevant and proportionate to the decision at hand or had any causal justification.

A closely related issue concerns the accountability of the decision-making process. As summarised by one police representative interviewed, 'Who gets in trouble when it goes wrong? You need to make sure people are still accountable for the machine and what it does'.[116] To address this issue, it is essential to build audit logs into the systems, ensuring there is a clear process to record the reasons why officers take certain action on the basis of an algorithmic output.[117] The NPCC-APCC 'National Policing Digital Strategy' also includes a commitment to 'provide clear lines of accountability on data and algorithm use at the top of all policing organisations, including accessible complaints and redress processes. This could be achieved by extending the Data Protection Officer role and updating Chief Officer responsibilities.'[118] As part of 'updating Chief Officer responsibilities', it appears necessary to assign a 'senior responsible owner' for each policing algorithm to ensure ultimate accountability for the performance of the tool and how it is deployed.[119]

In addition to technical transparency, legal issues could arise if the decision-making process of which the algorithm is part is opaque to the data subject or to the decision-maker. As summarised by the Law Society, 'when an individual is faced with a decision or a measure in a criminal justice context, it is critical they can assess it was legitimate, justified, and ultimately, legal'.[120] As such, when an algorithmic system is used in a criminal justice context, or when the output may eventually form part of an evidential process or trigger an intervention, the details of the system must be sufficiently intelligible both to the decision-maker and to the subject, in particular what factors were taken into account during analysis.

[114] Author interview with A3, academic legal expert, London, 28 June 2019.

[115] Author interview with A2, academic legal expert, London, 28 June 2019.

[116] Author telephone interview with L12 and L13, police respondents, 15 August 2019.

[117] Author telephone interview with L5, senior police technologist, 10 July 2019.

[118] NPCC and APCC, 'National Policing Digital Strategy', p. 15.

[119] *Ibid*.

[120] The Law Society Commission on the Use of Algorithms in the Justice System and the Law Society of England and Wales, 'Algorithms in the Criminal Justice System', p. 21.

Commercial and Procurement Considerations

The research highlighted significant risks associated with police procurement of commercial algorithms, particularly relating to the force's ability to scrutinise the system and the data used to build it. In many cases, algorithms used by the police are developed for different purposes and trained on non-police data:

> The use of the term 'machine learning' is an issue. The majority of tools used by the police are machine learned, they are not by and large true machine learning. Tools come to you as learned versions, they are not continuing to learn.[121]

> One of the problems with buying ML tools is if you don't have access to the training data. You're probably better off data testing it yourself beforehand … For instance, if we use tools that have been developed based on crime data from the US and then apply it to the UK, it's a very different policing context.[122]

It may be difficult to submit algorithms for independent evaluation if the intellectual property is confidential and owned by a third-party developer,[123] and the procurer of the system may need to inspect the algorithm in detail, which almost 'undermines the point of outsourcing'.[124] When the police procure pre-trained algorithms, the data on which the model was trained is unlikely to be representative of operational police data, necessitating re-testing with representative operational data. In these circumstances, transfer learning – whereby a pre-trained model is downloaded and re-trained on new data – could be a suitable approach, as this does not necessarily require access to the original training data.[125] However, the transfer learning process requires specialist data science expertise, meaning the force would still need to build its own data science capability.

Partly for these reasons, several police forces have focused efforts on building in-house data science expertise rather than outsourcing to third-party developers. As summarised by one police interviewee:

> To have an external supplier-contract setup is quite sub-optimal for a few reasons … Ideally all this expertise needs to be in-house. We wouldn't outsource our arrests, we wouldn't outsource our intelligence functions. Why are we outsourcing analytics? We need to build these capabilities ourselves. Analytics is seen as an expensive luxury but actually it can be really affordable.[126]

However, while some forces (such as Avon and Somerset Constabulary) have been successful in building in-house data science expertise, the majority remain reliant on commercial arrangements with third-party providers, raising issues concerning intellectual property and commercial confidentiality.

[121] Author interview with L14 and L15, senior police technologists, London, 19 September 2019.

[122] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[123] Author telephone interview with A1, academic policing expert, 1 June 2019.

[124] Author interview with A2, academic legal expert, London, 28 June 2019.

[125] *Ibid*.

[126] Author telephone interview with L1, senior police officer, 1 July 2019.

III. Bias and Discrimination

'PREDICTIVE POLICING' TOOLS have been labelled as 'racially biased', with claims that they over-predict individuals from certain racial groups or particular neighbourhoods where postcode functions as a 'proxy variable' for race.[127] However, it is important to note that most studies purporting to demonstrate racial bias in policing algorithms are based on analysis conducted in the US, and there is insufficient evidence to assess the extent to which these concerns are transferable to the UK context.

In relation to predictive mapping, concerns have been raised that racially biased police practices may become encoded into a statistical algorithm, leading to certain areas receiving a disproportionately large police presence.[128] However, very few empirical studies have actually examined the issue of bias in predictive policing algorithms, and the only randomised controlled trial that has been conducted found no significant differences in the proportion of arrests by racial-ethnic group between locations where mapping software was and was not deployed.[129] The authors note, however, that 'whether the same outcomes would hold given changes in implementation is uncertain', and that 'continued empirical scrutiny along with careful policy development will be needed to guard against bias in predictive policing and ensure fairness in outcomes'.[130]

Based on a review of the academic literature, it appears that there is currently a lack of sufficient empirical evidence to assess the extent to which bias in police use of algorithms actually occurs in practice in England and Wales, and whether this results in unlawful discrimination. Having said this, research has consistently demonstrated racial bias in policing and criminal justice outcomes more generally (though the extent to which this is due to unlawful discrimination is unclear).[131] These disparities are inevitably encoded into police-recorded data, which could lead to algorithmic outputs which replicate or even amplify these biases inherent in the data.

[127] For example, see Solon Barocas and Andrew D Selbst, 'Big Data's Disparate Impact', *California Law Review* (Vol. 104, No. 3, 2016), pp. 671–732; Richardson, Schultz and Crawford, 'Dirty Data, Bad Predictions'; Danielle Ensign et al., 'Runaway Feedback Loops in Predictive Policing', *Proceedings of Machine Learning Research* (Vol. 81, 2018), pp. 1–12.

[128] Sarah Brayne, 'Big Data Surveillance: The Case of Policing', *American Sociological Review* (Vol. 82, No. 5, 2017), pp. 977–1008.

[129] P Jeffrey Brantingham, Matthew Valasik and George O Mohler, 'Does Predictive Policing Lead to Biased Arrests? Results from a Randomized Controlled Trial', *Statistics and Public Policy* (Vol. 5, No. 1, 2018), pp. 1–6.

[130] *Ibid*.

[131] For further discussion, see Niamh Eastwood, Michael Shiner and Daniel Bear, *The Numbers in Black and White: Ethnic Disparities in the Policing and Prosecution of Drug Offences in England and Wales* (London: Release Publications, 2013); David Lammy, 'The Lammy Review: An Independent

Officers recognised the risk of bias in police decision-making, with comments such as ‘there is bias across everything we do’,[132] ‘whenever we have to decide an outcome there’s always an opportunity for bias’,[133] and ‘everyone has an element of bias that is unbeknown to them … The algorithms don’t come with human biases themselves, but they do potentially reflect underlying biases’.[134]

Research for this study has found that bias could be introduced at various stages in the project lifecycle, necessitating ongoing monitoring and tracking of discrimination risk. There are four distinct phases in a police analytics project where bias may be introduced: problem formulation; design; testing; and deployment.

Problem Formulation

At the early ‘problem formulation’ phase, predictive technological solutions have been criticised for focusing on low-level ‘nuisance’ crime, or on areas with high crime levels and thus poor neighbourhoods.[135] The choice of which crime types to focus on may itself be biased, leading to a disproportionate deployment of resources: ‘there’s another level of bias which is around certain types of crime being more heavily prioritised, and unclear reasons for prioritising certain things over others’.[136]

Bias may be introduced in the way the police ‘frame a problem’ (for instance, the labels used to record certain crime types).[137] The choice of which crimes to analyse may create false impressions regarding the scale of such criminal activity, with one interviewee explaining that ‘we are “flashing our flashlight” at this area and so will find offending’.[138] As well as reflecting structural inequalities in society, police data is a reflection of police activity, much of which is proactive and ‘intelligence led’, meaning that the data may reflect operational priorities which may themselves be biased.[139]

In addition, many problems may not lend themselves to a technological solution, and the strong financial incentives on police forces to develop ‘innovative’ digital solutions may create a bias in favour of digital solutions, without equal consideration of non-technological measures. Reliance

Review into the Treatment of, and Outcomes for, Black, Asian and Minority Ethnic Individuals in the Criminal Justice System’, September 2017.

[132] Author telephone interview with L7, senior police respondent, 18 July 2019.

[133] Author telephone interview with L1, senior police officer, 1 July 2019.

[134] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[135] Cathy O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York, NY: Penguin Random House, 2016).

[136] Author interview with A2, academic legal expert, London, 28 June 2019.

[137] Author telephone interview with A5, academic expert in policing and criminal justice, 11 July 2019.

[138] Author telephone interview with L9 and L10, police investigator and analyst, 1 August 2019.

[139] Author telephone interview with A5, academic expert in policing and criminal justice, 11 July 2019.

on algorithmic tools may create a 'risk of modelling very complex social issues in an over-simplistic way', leading to a 'selective understanding of criminal behaviour', disproportionately affecting individuals who are more likely to come into contact with the police (for instance, those with mental-health issues).[140]

Design

When designing a police algorithm, the choice of dataset(s) is crucial to minimise risk of bias. All datasets will be skewed to some degree, but this does not necessarily imply unfair discrimination. As one senior police technologist noted: 'From a statistical point of view, we recognise that all datasets are biased'.[141] This recognition was further highlighted by an academic legal expert, who stated that:

> For each of the negative grounds, bias can come in at various levels. It can come in when people are reporting crime, it could come in at stop and search, it could come in at judicial decision-making. Unequal base rates could be another source of bias, not necessarily due to real-world discrimination.[142]

There was recognition that police-recorded data is rarely an accurate statistical record of crime:

> A fundamental issue is that these police algorithms are built using police data. Crime records, other police-recorded information, which is of itself not an accurate record of crime, but merely a record of policing. That will inevitably lead to certain areas, certain communities, certain people being over-represented on those databases. It's in no way a statistically accurate record of actual crime.[143]

> If you're interested in where crime is happening, it's not just arrest data you need to use, you need to look at everything.[144]

Removing demographic variables from the statistical model is unlikely to be sufficient to eliminate bias, as these demographic features could be implicitly encoded in other variables in ways that are not always immediately apparent:

> Just because you take out explicit or sensitive data such as someone's race or sex, machine learning algorithms can make inferences and connect the dots between all the different data sources they have.[145]

In many cases, it may be desirable to keep certain demographic variables within the statistical model to assess differences in offending patterns between demographic groups:

[140] Author telephone interview with A4, academic expert in policing and criminal justice, 11 July 2019.
[141] Author interview with L14 and L15, senior police technologists, London, 19 September 2019.
[142] Author interview with A2, academic legal expert, London, 28 June 2019.
[143] Author interview with C1, representative of civil society organisation, London, 7 October 2019.
[144] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.
[145] Author interview with A3, academic legal expert, London, 28 June 2019.

> Interestingly, there are some characteristics of offenders that are based on clear criminological theory. Understanding those issues is also crucial when you're looking at your build and the outputs you're getting … We're really clear about every piece of data that's in our tool. We've made decisions early on that we will put ethnicity and gender in there, because we want to see what impact that will have on prediction.[146]

However, problems may arise when attempting to determine whether an imbalance in the dataset is indicative of unfair discrimination or reflective of real-world differences in offending distribution across demographics. As one academic interviewee pointed out, 'this is not just a job for the data scientist alone. This requires domain expertise and input from other experts'.[147]

Testing

Concerns were raised regarding the criteria against which police technology projects are evaluated, with comments such as 'police forces are very keen to try something out but typically very bad at evaluation'.[148] In the context of crime prevention, it can be impossible to assess the 'predictive accuracy' of statistical tools in the field, because there is no way of knowing what may have happened if the police did not intervene to prevent the predicted outcome from happening. As summarised by one academic expert, 'you can't just release someone to see if they're going to be dangerous'.[149]

For this reason, policing algorithms are typically validated retrospectively using a sample for which the ultimate outcomes are known. This does not take into account model shrinkage, overfitting and concept drift as a result of applying the tool to new, unfamiliar data.[150] As summarised by Douglas Mossman, 'By their very design, [statistical risk assessment tools] depend on relationships established in specific populations at specific times in the past, and these relationships may not apply, or may not apply in exactly the same way, to future populations living in different social contexts and circumstances'.[151] Interviewees recognised that issues may arise when procuring commercial off-the-shelf tools which are trained in a particular geographical area but then deployed in another,[152] and that 'whenever you're moving from one area to another, then clearly you have to do further validation and calibration'.[153] As

[146] Author telephone interview with L6, senior police officer, 10 July 2019.

[147] Author interview with A2, academic legal expert, London, 28 June 2019.

[148] Author interview with A6, academic and policy expert, London, 13 August 2019.

[149] Author interview with A2, academic legal expert, London, 28 June 2019.

[150] Christopher D Webster, Quazi Haque and Stephen J Hucker, *Violence Risk: Assessment and Management – Advances Through Structured Professional Judgement and Sequential Redirections,* 2nd Edition (Hoboken, NJ: Wiley & Sons, 2013), p. xiii.

[151] Douglas Mossman, 'Evaluating Risk Assessments Using Receiver Operating Characteristic Analysis: Rationale, Advantages, Insights, And Limitations', *Behavioral Sciences & the Law* (Vol. 31, No. 1, 2013), pp. 23–39.

[152] Author telephone interview with A1, academic policing expert, 1 June 2019.

[153] Author interview with R3, member of regulatory/oversight body, London, 3 October 2019.

noted by one police representative, 'You'll always have some attrition between your high level of accuracy that an ML output can give you, and how it is deployed in the hands of the user. High-level accuracy results will inevitably be much lower in the field than [what] they claim to be'.[154]

When assessing the operational benefits of a policing algorithm, a focus on 'predictive accuracy' may distract attention away from whether the tool is having any real-world impact in preventing, detecting or reducing crime. This requires a testing process that looks beyond the algorithm to also assess the interventions made on the basis of the prediction:

> If you start producing predictions and then those predictions are acted upon, how is that happening and how does that feed back into the process? You need to incorporate feedback loops to assess how these predictions are being acted upon.[155]

Deployment

It has been argued that the deployment of predictive policing algorithms may lead to certain minority groups being disproportionately targeted, creating 'feedback loops' of discrimination:

> We pile loads of resources into a certain area and it becomes a self-fulfilling prophecy. Purely because there's more policing going into that area, not necessarily because of discrimination on the part of officers.[156]

> Typically, we put our resources in deprived areas where there are high crime levels, and generally those areas tend to be where there is a large ethnic minority group … If we simply make that decision based on algorithms, a particular ethnic community may be getting a disproportionately large police response.[157]

There is a risk that the use of algorithmic predictions could lead to human users 'making causal inferences that do not hold up to scrutiny', which will 'operationalise the biases' present in the data.[158] Recent research carried out by innovation foundation Nesta discovered that 'when practitioners artifice they will, in the vast majority of cases, draw on elements of professional intuition, together with unconscious bias, to inform the intuitive element of the decision-making process'.[159] Automation bias and confirmation bias are particular concerns in this regard:

[154] Author telephone interview with L6, senior police officer, 10 July 2019.

[155] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[156] *Ibid*.

[157] *Ibid*.

[158] Author telephone interview with A4, academic expert in policing and criminal justice, 11 July 2019.

[159] Thea Snow, 'Decision-Making in the Age of the Algorithm', Nesta, November 2019, <https://media.nesta.org.uk/documents/Decision-making_in_the_age_of_the_algorithm.pdf>, accessed 19 January 2020.

> The more the machine points the human in a particular direction, the more likely the human is to follow this without question.[160]

> There's a real danger on becoming over-reliant on poor-quality algorithmic systems at the expense of informed human decision-making.[161]

> Confirmation bias is the greatest concern, feedback loops around our high-priority offenders and potentially missing other pockets of criminality or offending.[162]

Nesta's research highlights that there is a delicate balance to be struck, 'discouraging practitioner deference to the tool, but not to the point that they feel nervous to use it at all'.[163] The College of Policing's APP advises that 'RI [risk identification], RA [risk assessment] and RM [risk management] tools should be regarded as an excellent but limited, means of improving the likelihood of identifying and preventing future offending or victimisation. They can enhance professional judgement but not replace it'.[164]

Interviewees stressed that the risk-assessment tools used by police forces are being used in this advisory capacity, indicated by comments such as:

> None of our algorithmic outputs feed an automated decision system. The whole point of this is to help make sense of large volumes of data to help practitioners make judgements. It's the professional judgement that determines what will happen … It's purely a decision-support tool.[165]

Nevertheless, concerns were raised that over-reliance on machine filtering could lead to other relevant factors being ignored.[166] As actuarial tools rely on identifying statistically significant correlations in historic data, a fully automated approach could mean that practitioners may fail to identify relevant risk factors because they were not found to be statistically significant in historic data. As one police respondent explained, 'the algorithm is only as good as the data they put in, and often police officers have data in their head that doesn't get into our systems'.[167] For example, a 'low-risk' label could be interpreted to mean that an individual requires no further monitoring or intervention. Such 'low-risk' individuals may have specific needs and vulnerabilities that should be addressed as part of a bespoke risk-management plan; needs and vulnerabilities which may not be detected by a statistical algorithm. Such individuals may then fail to receive the necessary support to prevent them from returning to problematic behaviour.

[160] Author telephone interview with L12 and L13, police respondents, 15 August 2019.

[161] Author interview with C1, representative of civil society organisation, London, 7 October 2019.

[162] Author telephone interview with L5, senior police technologist, 10 July 2019.

[163] Snow, 'Decision-Making in the Age of the Algorithm'.

[164] College of Policing, 'Authorised Professional Practice: Risk'.

[165] Author telephone interview with L4, senior police technologist, 9 July 2019.

[166] Author telephone interview with L12 and L13, police respondents, 15 August 2019.

[167] Author telephone interview with L5, senior police technologist, 10 July 2019.

Addressing Bias

The research identified various approaches to addressing risk of bias in police algorithms, including using alternative datasets, statistical methods for detecting and eliminating bias, and treating the algorithmic output as a form of 'intelligence'.

In relation to the use of non-police data, one police participant suggested that 'to address bias, we need to be better at information sharing, using data from third parties that may not be picked up by the police'.[168] However, others questioned whether this would provide sufficient reassurance, as 'the reasons that data becomes biased [aren't] entirely dependent on how one institution behaves, but rather how we as a society behave … It's still problematic regardless of what data is being used'.[169] Other participants suggested more sophisticated statistical methods for identifying bias and then adjusting the dataset accordingly:

> There are ways in which you can do it ranging from excluding attributes, including attributes but swapping the labels around, you mess around with the data to some degree so some of the relationships are broken on purpose … There are technical ways in a model-building point of view that you can try to overcome it.[170]

One academic expert suggested that 'there are different stages at which you could intervene to try to reduce bias, but changing the data is probably going to be easiest and most similar to what they've done before'.[171] Indeed, several police respondents were aware of computer-science methods for reducing bias, but acknowledged that these had not yet been implemented in existing police technology projects:

> We could [look for bias in our analytics] a bit more overtly ... There's probably a duty to do that.[172]

> I know there are pieces of software that can scan and detect for bias. We haven't used those, but we might in future … We do look at the outputs of the forecasting tool and compare those to our population to see if it's representative.[173]

However, there may be significant consequences of altering police data such that it no longer reflects the data that was recorded – for instance, in relation to the GDPR requirement that personal data is, where possible, kept accurate and up to date.

[168] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[169] Author telephone interview with C2, representative of civil society organisation, 17 October 2019.

[170] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[171] Author interview with A2, academic legal expert, London, 28 June 2019. For further discussion on the issue of fairness in machine learning algorithms, see Reuben Binns, 'Fairness in Machine Learning: Lessons from Political Philosophy', *Proceedings of Machine Learning Research* (Vol. 81, 2018), pp. 149–59.

[172] Author telephone interview with L5, senior police technologist, 10 July 2019.

[173] Author telephone interview with L6, senior police officer, 10 July 2019.

Bearing in mind the difficulties associated with eliminating bias from a dataset, roundtable participants suggested that – rather than attempting to 'fix' the data, it may be more constructive to focus on ensuring fairness in the outcomes of the overall analytical process.[174] In particular, several respondents suggested that the output of an algorithmic prediction should be treated as a form of intelligence, and therefore associated with levels of confidence for the officer to consider when taking the output into account:

> Should you treat this type of information as a form of intelligence? My personal view is that this is essential for the future. There is guidance out there about handling probabilistic data in the operational community in terms of the corroborative value – the need to look at continuity, validity and integrity.[175]

> Officers are quite comfortable with intelligence feeding into decision-making, what we need to work out is how the output of an algorithm feeds into that process.[176]

Treating algorithmic insights as a form of intelligence would place the burden of responsibility on the human user to demonstrate that they had critically assessed the validity and relevance of such information when forming their overall judgement or decision, thereby ensuring ultimate accountability of the overall decision-making process. It would be incumbent on the user to ensure that any subsequent action taken on the basis of the algorithmic prediction was not unfairly discriminatory.

[174] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[175] Author telephone interview with L8, senior police officer, 23 July 2019.

[176] Roundtable event organised in partnership with techUK, London, 22 July 2019.

IV. Regulation, Governance and Oversight

Current Framework

INTERVIEWEES UNIVERSALLY RECOGNISED a lack of any official national guidelines or standards for police use of algorithms:

> There are many statutes relevant to operational work of the police and other agencies but it's not clear how these relate to data analytics.[177]
>
> There is no strong guidance, no regulation. What governs our behaviour and conduct is the Code of Ethics. Within the Code of Ethics, which standards of professional behaviour would apply to the use of an algorithm? I'm not sure … there's not much governing the use of algorithms specifically.[178]
>
> There are no professional standards that I'm aware of … You can call upon various pieces of law, but I think it's important that we have more specificity … There are lots of codes of practice around specific data capabilities, but none that govern the use of algorithms specifically.[179]

In the area of predictive algorithms, the NPPC's Business Change Council has adopted the 'Algocare' model, and together with additional explanatory documentation, recommends its use to Chief Constables. Numerous interviewees mentioned Algocare as the only de-facto national guidance.[180] Several participants suggested that primary legislation may be needed to regulate these new technologies, and that guidance alone was unlikely to provide sufficient safeguards. However, one police participant suggested that 'the ideal legislative framework will be a long time coming. We can't afford to wait for that', and that 'if there's no code of practice, the technology will just develop without us'.[181] One academic legal expert noted that 'practitioners often struggle with translating and interpreting complex legal frameworks, particularly human rights legislation … You can say, we'll take the safeguards from various different frameworks and apply them to predictive policing, but they aren't specific to predictive policing'.[182]

[177] Author telephone interview with L5, senior police technologist, 10 July 2019.

[178] Author telephone interview with L1, senior police officer, 1 July 2019.

[179] Author telephone interview with L8, senior police officer, 23 July 2019.

[180] See Marion Oswald et al., 'Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and "Experimental" Proportionality', *Information & Communications Technology Law* (Vol. 27, No. 2, 2018), pp. 223–50.

[181] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[182] Author interview with A3, academic legal expert, London, 28 June 2019.

The importance of clear guidance and adherence to it has been emphasised in other contexts. For instance, in R Bridges vs. Chief Constable of the South Wales Police, the High Court of England and Wales dismissed the claim that the generic legal framework (consisting of the common law, data protection and human rights legislation, codes of practice issued under legislation and the police's own local policies) was insufficient to govern the police's trial of live facial recognition (LFR), and noted that the police force's own policies were part of the scope of existing regulation.[183] In light of this judgment, the MPS recently announced that the force will be deploying LFR technology operationally, and published a series of documents on their website including an LFR guidance document and standard operating procedure detailing the authorisation process that must be followed, operational parameters of when, where and how it will be deployed, and guidance regarding the generation and management of LFR watchlists.[184]

Similarly, a new set of nationally approved guidelines appears essential to ensure the lawful and ethical development and deployment of statistical algorithms for policing. Recognising this requirement, the new 'National Policing Digital Strategy' outlines a commitment to 'develop a National Data Ethics Governance model, which will outline standards and guidelines to be adhered to and embedded in our decision making processes'.[185] The following sections provide recommendations on what should be included in these new guidelines, based on the findings of the research.

Ethical Oversight

There was widespread recognition among research participants of the need for meaningful independent 'ethical oversight' of police data analytics projects, but a lack of clarity on how this should be achieved in practice. In particular, it remains unclear whether this oversight should be delivered at the local force level, or in the form of a centralised national structure. While various groups currently provide ethical oversight to policing, there remains a lack of consistency in approach between forces, and unclear delineation of local, regional and national responsibilities.

While all police forces in England and Wales have now established local ethics committees,[186] these are not focused on digital technology, and several interviewees suggested that they lack the necessary technical expertise to meaningfully scrutinise police technology projects. A senior police officer noted that:

[183] R Bridges vs. Chief Constable of the South Wales Police, High Court of Justice, Queen's Bench Division, CO/4085/2018, 4 September 2019, para. 84. This judgment is now subject to appeal, and so must be treated with a degree of caution.

[184] Metropolitan Police Service, 'Live Facial Recognition', <https://www.met.police.uk/live-facial-recognition-trial/>, accessed 28 January 2020.

[185] NPCC and APCC, 'National Policing Digital Strategy', p. 15.

[186] Some of these are joint force initiatives, meaning that one committee provides oversight to several regional forces.

> Our ethics committee are nowhere near this sort of subject matter and wouldn't be qualified to provide appropriate oversight and direction on this. Potentially if you had the right people around the table it could work well … In my experience of how ethics committees are currently set up, it would need something different and bespoke in order to meet this need.[187]

Conversely, another police participant disagreed, suggesting that existing ethics committees could be adapted to incorporate scrutiny of data analytics:

> We have our ethics committee already. Ethics committees nationally need to pivot into the modern world. There's no reason why you can't have local data scientists on those committees. We've taken the use of predictive analytics to our ethics committees previously. I don't think we need another committee to deal with that.[188]

Recognising a requirement for a 'bespoke' option, WMP has created its own data ethics committee, the first of its kind within UK policing, to advise the Chief Constable and Police and Crime Commissioner (PCC) on the force's data analytics projects. A WMP roundtable participant explained that 'we want extensive public consultation before we roll out anything', expressing a desire to 'create a culture of ethics by design' and ensure that ethical issues are reviewed at regular intervals throughout the research process.[189]

Interviewees praised the WMP data ethics committee and tasking process for providing a robust and unique level of scrutiny not currently replicated across other areas of policing.[190] However, concerns were raised regarding its resourcing, scaleability and long-term sustainability, as well as its influencing power. One civil society representative commented that 'I think the WMP ethics committee is a good model. The problem is that's almost entirely dependent on goodwill and a good relationship between the PCC and the force. They don't have any actual power other than the fact that WMP have agreed to listen to the decisions of the ethics committee'.[191] The new 'National Policing Digital Strategy' also commits to establishing 'a core principle that the public's views on data analytics are pro-actively built into an ethical assessment at the design stage of any digitally-enabled service improvement'.[192] However, it is unclear how such a process would be implemented in practice across all 43 forces, and whether it will be accompanied by a commitment to halt or review projects if this were the suggested outcome of the consultation.

There was also concern that force-level ethics committees may not be scaleable to the national level, and may lead to a culture of 'postcode ethics',[193] with some interviewees arguing that there is a need for a centralised structure, which could also include other law enforcement agencies

[187] Author telephone interview with L6, senior police officer, 10 July 2019.

[188] Author telephone interview with L4, senior police technologist, 9 July 2019.

[189] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[190] *Ibid*.

[191] Author interview with C1, representative of civil society organisation, London, 7 October 2019.

[192] NPCC and APCC, 'National Policing Digital Strategy', p. 15.

[193] Author telephone interview with L8, senior police officer, 23 July 2019.

such as the National Crime Agency.[194] The Independent Digital Ethics Panel for Policing was mentioned by several interviewees as an existing structure which could provide this national-level oversight,[195] but concerns were raised regarding its long-term resourcing and access to relevant expertise. Participants also noted the risk of 'creating layers and layers of oversight' and introducing resource-demanding review processes which may stifle innovation.[196]

It was also suggested by one civil society representative that ethics committees could 'act as a fig leaf over wider discussions' that the police should be having directly with the general public.[197] Nevertheless, police participants recognised that 'transparency is the most important thing. The moment we hide anything, we lose public trust'.[198] Interviewees recognised the importance of including independent academic institutions in the ethical review process. As summarised by one police interviewee, 'there's a big case for links to academics, the local university links. There's a natural opportunity there to use them as part of external scrutiny and ethics'.[199]

In summary, it is crucial to ensure independent review and scrutiny of police data analytics projects, at both the local and national level. Whether existing ethics committees are sufficient to provide this oversight, or whether bespoke 'digital ethics committees' need to be established in parallel, is a matter of debate. The UK Police Ethics Guidance Group should conduct a comprehensive review to assess whether existing ethics committees could be 'upskilled' to provide meaningful ethical review of police technology projects, or whether bespoke digital ethics committees should be established in parallel. This review should also assess the viability of a national or regional ethics review process based on consistent terms of reference.

Roles and Responsibilities

There are several participants in the 'regulatory space'[200] relevant to police use of data analytics. Some play a major role, including Chief Constables, Police and Crime Commissioners, the College of Policing, NPCC, HMICFRS and the Home Office. Others have a more limited influence or one confined to specific issues, including the Information Commissioner's Office (ICO), Investigatory Powers Commissioner's Office, Surveillance Camera Commissioner and Forensic Science Regulator (FSR). Furthermore, several bodies are engaging in advisory or investigatory activities, including the Centre for Data Ethics and Innovation, the Office for Artificial Intelligence, parliamentary committees, independent committees, bodies with sector expertise or policymaking functions, and campaigning organisations.

[194] *Ibid*.

[195] Author telephone interview with L1, senior police officer, 1 July 2019.

[196] Author telephone interview with L12 and L13, police respondents, 15 August 2019.

[197] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[198] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[199] Author telephone interview with L4, senior police technologist, 9 July 2019.

[200] Leigh Hancher and Michael Moran, 'Organizing Regulatory Space', in Robert Baldwin, Colin Scott and Christopher Hood (eds), *A Reader on Regulation* (Oxford: Oxford University Press, 1998).

Despite this crowded regulatory space, there is a lack of clarity regarding the delineation of responsibilities between these organisations for the development of guidance and standards, regulation and oversight of police use of data analytics:

> It all overlaps and intertwines … The ICO, FSR, all the other regulators, they all have different codes of conduct.[201]

> For data protection, we go to the ICO. But for ethics, it's more emerging, it's progressing, it's a bit harder to pin down.[202]

> I don't see the present regulatory environment as being comfortable about the increasing complexities of data. Where does the ICO finish and the FSR start?[203]

Several interviewees proposed a 'co-regulatory approach' for police use of data analytics, involving the establishment of new multi-agency teams to develop new guidance, set minimum standards, and oversee and inspect against those new standards:

> Before we allocate responsibility around this, the regulatory roles and responsibilities need to be determined, and there might need to be a joint regulatory team drawing on expertise from various regulators.[204]

> There is a great deal of focus in new legislation on co-regulatory approaches, but these aren't being implemented in practice as yet from what I've seen. That needs to be avoided if we are to avoid the mistakes of the past.[205]

As police use of data analytics engages various legal frameworks, a collaborative approach seems appropriate in this context. Interviewees suggested that the development of new guidelines and standards would require a joint approach between the NPCC, the Home Office and the College of Policing:

> The Home Office have a crucial part to play in setting those standards for policing, whether those are cross-government standards that a department like CDEI might come up with … The College of Policing is a national body that should develop the service, so they need to be engaged. The NPCC together with the College probably need to be at the heart of this for the sake of policing, so we're not leaving all 43 forces to do their own thing.[206]

[201] Author telephone interview with L11, police technologist and analyst, 1 August 2019.

[202] Roundtable event organised in partnership with techUK, London, 22 July 2019.

[203] Author telephone interview with L8, senior police officer, 23 July 2019.

[204] *Ibid*.

[205] Author interview with A3, academic legal expert, London, 28 June 2019.

[206] Author telephone interview with L1, senior police officer, 1 July 2019.

Regarding regulation, oversight and enforcement of these new standards, there was agreement that it would be counter-productive to attempt to create a new regulatory body. Rather, roundtable participants suggested that the ICO is the most appropriate body to scrutinise forces' use of data analytics, in conjunction with the Equality and Human Rights Commission (EHCR).[207] The ICO was recognised as having the necessary statutory backing to enforce against national standards, but was described by one academic expert as being 'woefully underfunded', with all of its funding coming from fees levied on data controllers, and no funding provided by central government.[208]

There was widespread recognition of the crucial role of HMICFRS in inspecting forces against a new set of national standards:

> HMIC assessments directly affect how policing learns and cuts its costs. The PEEL inspection process now is so rigorous and robust, it's a real opportunity to get things into policing through our industry watchdog … In five years' time, I would expect the deployment of algorithms to be a standard thing that HMIC are looking at, alongside other points of delivery in policing.[209]
>
> HMIC are good at inspecting against an agreed set of standards perhaps set by another body. If, for example, CDEI were able to say these were the agreed standards, HMIC would be a good body to inspect against those … It would have to be undertaken by somebody who has some level of understanding of how these projects are undertaken.[210]

The organisational framework is already in place to incorporate police use of data analytics into HMICFRS inspections, as there is a dedicated team focused specifically on crime data integrity inspections.[211] In 2020/21, HMICFRS should establish an External Reference Group for police use of data analytics, with a view to updating the crime data integrity inspection framework to include inspection against the proposed new guidance. This should draw upon the combined expertise of the ICO and the EHRC, and engage external subject matter expertise (for instance, from data scientists and specialist legal experts).

In addition to a new framework for guidance, inspection and oversight, a new centralised committee should be established to coordinate the use of algorithms across policing in England and Wales.[212] This group should maintain a high-level catalogue of all algorithms used by police forces nationwide to inform operational decision-making, to encourage cooperation between forces, sharing of best practices and avoidance of duplication. The group could also introduce a mechanism by which consistent specialist expertise can be accessed by police forces in the areas of data science, ethics and tool

[207] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

[208] *Ibid*.

[209] Author telephone interview with L4, senior police technologist, 9 July 2019.

[210] Author telephone interview with L1, senior police officer, 1 July 2019.

[211] Author interview with R1 and R2, members of regulatory/oversight body, London, 14 August 2019.

[212] Author interview with L14 and L15, senior police technologists, London, 19 September 2019.

evaluation, and explore the feasibility of establishing a mechanism for police forces to access a centralised team of specialist legal advisers, in a similar way that government departments can access specialist legal advice via the Government Legal Department.[213]

Towards a New Policy Framework

Any future policy or guidance should be 'tech-agnostic' and principles based. Rather than establishing prescriptive rules and standards for different data technologies, the framework should instead establish standardised processes to ensure that data analytics projects follow recommended routes for empirical evaluation of algorithms within their operational context and evaluate the project against legal requirements and ethical standards:

> It needs to be tech-agnostic, it can't become out of date this time next year. It needs to be generic enough to stand the test of time, but it needs to give us clear direction so we can apply a consistent model.[214]
>
> Because ultimately everything becomes very problem specific, it would be difficult to have a one size fits all and be prescriptive about what methods should be used in specific circumstances. What I've not seen is guidance that talks about what the process should involve. The broad questions there about avoiding discrimination, ensuring data quality, that kind of thing.[215]

The guidance may need to include separate sections for different technologies, as the legal and ethical considerations will vary depending on the tool and application.[216] Existing surveillance codes and related inspections were suggested as a potential model, as there are numerous existing codes of practice for surveillance, each governing specific activities or technologies.[217]

In terms of establishing a standardised, 'tech-agnostic' process for police data analytics projects, the cross-industry standard process for data mining (CRISP-DM) is a useful starting point, though various other models exist.[218] An overview of the CRISP-DM process is replicated in Figure 3.

[213] Author telephone interview with A4, academic expert in policing and criminal justice, 11 July 2019.

[214] Author telephone interview with L1, senior police officer, 1 July 2019.

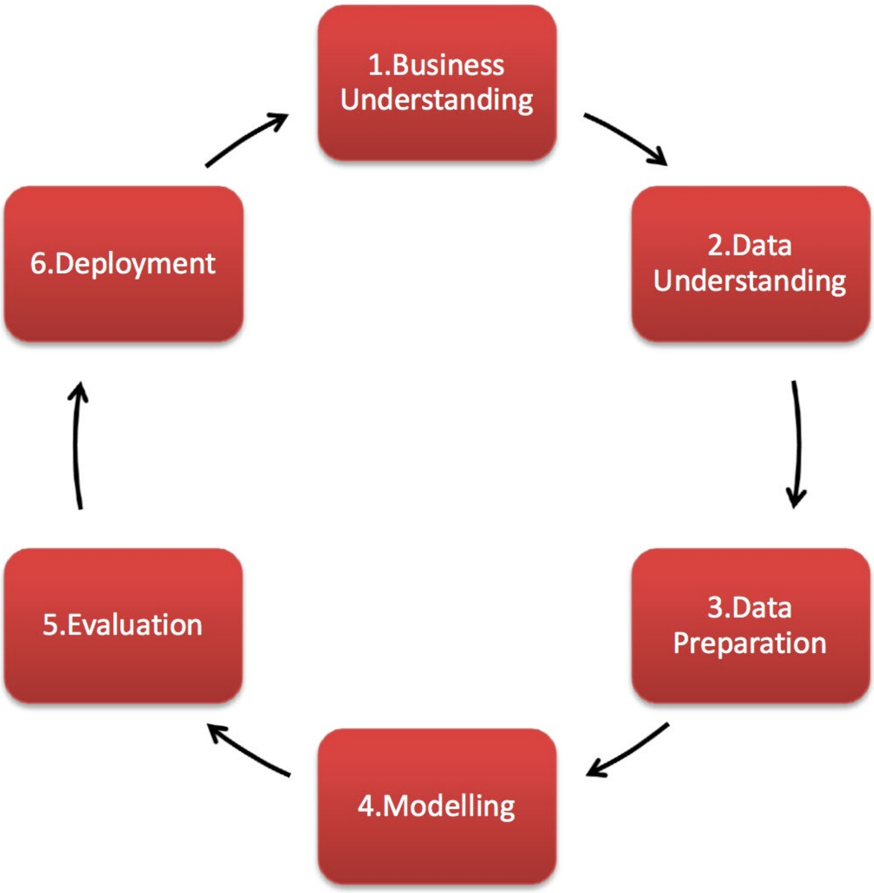
[215] Author telephone interview with L2 and L3, senior police officer and police technologist, 4 July 2019.

[216] Author interview with A2, academic legal expert, London, 28 June 2019.

[217] Author telephone interview with L6, senior police officer, 10 July 2019.

[218] Roundtable event organised by RUSI and CDEI, London, 25 July 2019.

Figure 3: Simplified Overview of the Cross-Industry Standardised Process for Data Mining (CRISP-DM)



Source: SmartVision, 'What is the CRISP-DM Methodology?', <https://www.sv-europe.com/crisp-dm-methodology/>, accessed 28 January 2020.

Of particular importance is the need to start with a clear 'business understanding' before moving into the 'data analytics' phase. In the policing context, this means starting with a clear business case documenting the problem the force is hoping to address, the reasons for selecting that problem, and why it is believed that this problem may lend itself to a technological solution. This contrasts with the 'exploratory' approach, which involves starting with the data analysis and not establishing a clear purpose for analysis until after the insights have been generated.

Alongside the CRISP-DM approach for data mining, the Agile project delivery cycle, as defined in the UK government's service manual, provides a useful and useable framework for overall project management of police data analytics projects.[219]

[219] UK Government, 'Agile Delivery', <https://www.gov.uk/service-manual/agile-delivery>, accessed 28 January 2020.

The following specific areas of focus are particularly relevant in the policing context, and must be explicitly included in any new guidance.

Legal, Ethical and Operational Input

Multi-disciplinary input at the start of a project is essential to assess the evidence on the particular problem under investigation and whether it lends itself to algorithmic analysis, the subjective nature of particular data, the risk of proxies for protected characteristics, the implications of errors for interventions, the legal basis for the use of a tool, and data protection and human rights concerns. In establishing a clear business case, the force should critically assess whether any additional resource investment is outweighed by the potential efficiency savings provided by the technology. Beyond specialist data science expertise, efforts should be made to receive ongoing feedback from academic criminologists and legal experts.

Integrated Impact Assessment

There was a strong sentiment among interviewees that there should be a mandatory requirement for an integrated impact assessment, incorporating:

- Data protection impact assessment.
- Equality impact assessment, describing the potential impact of the proposed project on people with protected characteristics.
- Human rights impact assessment.
- Empirical evaluation of accuracy and operational assessment of 'real-world' effectiveness.
- Assessment of expected level of errors where this can be established or estimated and potential consequences of these errors.
- Assessment of any positive obligations under Article 2 or Article 3 of the European Convention on Human Rights or associated public safeguarding issues.
- Assessment of any other legal requirements which may be relevant for specific projects (for instance, investigatory powers authorisations, evidential or valid decision-making requirements pursuant to criminal procedure, PACE requirements and investigations legislation, and any limitations on interventions a statistical algorithm may inform).
- Independent ethical assessment, the format of which will depend on what ethical oversight arrangements are in place.

The impact assessment should be considered a 'live' document and should be subject to regular review should the scope, context or purpose of the project change over time.

Standard Process for Model Development

A standard process for model development is essential to create consistency in approach and ensure relevant legal and ethical issues are taken into account in the tool design phase. This process should include an assessment of data requirements, both in terms of what data is needed to build the tool, how the input data will be analysed and how the resulting output will be stored and shared. Many projects will not require personal data, and when assessing whether a particular intrusion of privacy is necessary to achieve a particular policing aim, the force should assess whether the same insights could be derived by using an aggregated or anonymised dataset, rather than personal data.[220]

Procurement Considerations

To ensure meaningful scrutiny of data analytics software, the guidance should advise on appropriate rights as part of contractual agreements with any third-party supplier, such as:

- Knowledge of, and confidence in, the training datasets used to build the model.
- Appropriate access to algorithmic workings to facilitate third-party investigation and questioning in an adversarial context.
- Rights to use, amend and disclose the tool, its workings and input datasets where required for legitimate public sector decision-making.
- Regular rights of audit, testing and validation.
- Rights to require updating of the model, removing or adding input factors.
- Rights to request use of a data science expert as a witness to explain the algorithmic tool (for instance, in a criminal justice context).

Testing and Evaluation

Robust, empirical testing should be a core focus of any new framework. Participants highlighted a lack of guidance on how trials should be conducted and evaluated, describing how police forces are ‘being actively encouraged to develop this technology’ but ‘experimenting without any cloak of protection’.[221] As described by one police interviewee, ‘it is essential to have a new framework for evaluation, and a standardised methodology for testing … We have to be allowed to test and to be allowed to fail … How do we generate an empirical evidence base if we are not allowed to test?’.[222]

The interviewee suggested that the testing and evaluation process should include three stages:

> Bench test (to assess whether the tool performs the required analysis as expected in a controlled setting); Scenario test (empirical validation against operationally relevant data, to establish whether the police’s data lends itself to such analysis); Operational test (including how the human responds to the analytical output).[223]

Effective testing requires establishing specific, measurable and achievable evaluation criteria at the outset. This testing should also include an assessment of whether the algorithmic outputs

[220] Author interview with A6, academic and policy expert, London, 13 August 2019.

[221] Author interview with L14 and L15, senior police technologists, London, 19 September 2019.

[222] *Ibid*.

[223] *Ibid*.

are systematically skewed or biased towards a particular group or individuals within that group. In the context of machine learning, statistical models require ongoing, iterative review and validation: ‘It is about the continuing testing, validation, calibration. The danger is you design a robust tool and then that’s it. But it’s got to be looked at again over time’.[224]

Evaluation Standards

Universal evaluation standards may be needed to ensure the quality and empirical validity of algorithms used for policing (as is already the case for forensic science), ‘mapped against a rigorous scientific standard that is subject to inspection’.[225] It appears to be essential to establish context-specific evaluation methodologies to ensure the scientific validity of statistical algorithms used by the police, but it remains unclear where this responsibility should lie. Guidance is also needed on how confidence levels and error rates should be established, communicated and evaluated. Clarity is needed from the Home Office regarding who should be responsible for establishing these evaluation methodologies.

Human Interaction with the Tool and Classification of the Output

The decision-making process informed by the algorithm requires as much attention as the tool itself. Consideration should be given to how the output is presented to the officer and whether presenting a conclusive classification without revealing uncertainties could unduly influence officer judgement. Algorithmic outputs are by their nature uncertain. As summarised by one academic, ‘there needs to be a lot of training in terms of letting police officers know these systems are not infallible, they need to be made aware of the types of errors that can be made … What should be the focus is not just on the data that gets inputted into the system but on acknowledging the risks in interpreting these datasets’.[226]

Consideration must be given to how outputs are classified in police information systems (for instance, as a form of ‘intelligence’ alongside a confidence rating), thus limiting the extent to which they can be used or disclosed. Specific guidance may be needed to ensure officers are able to interpret and use the insights in conjunction with their own professional judgement. This would need to integrate with both the National Decision Model and Management of Police Information guidelines.

Data-Driven Interventions

It is crucial throughout all stages of the project to be mindful of what interventions the algorithmic insight may feed into and to be specific about the potential end use(s) of the tool. Many deployments of police algorithms and automation will not impact significantly on citizens’

[224] Author interview with R3, member of regulatory/oversight body, London, 3 October 2019.

[225] Author telephone interview with L6, senior police officer, 10 July 2019.

[226] Author interview with A3, academic legal expert, London, 28 June 2019.

human rights or civil liberties. However, in a criminal justice context, the use of an analytical tool can have profound implications for human rights:

> When we're talking about predictive analytics, it's worth talking about the context in which you're making that prediction … The sharper end might be very focused on an individual, the softer end might be that you identify a specific problem evolving in a specific community, then you go to a school and have a focused assembly … You don't want to be excluding certain tools from the start before you know what the insights are … It's about at what stage in the process you're excluding certain interventions.[227]

In particular, the force should regularly review whether the algorithm is still being used for the purposes and interventions for which it was developed, or whether it has been 're-purposed' and used to inform a different decision-making process.

Ongoing Tracking and Mitigating of Discrimination Risk

It follows from the public sector equality duty and prohibition of unlawful discrimination that careful tracking of inputs, outputs, influences on a decision and resultant actions are required.[228] The predictive weight given to the protected characteristic or proxy is likely to be particularly significant in this regard. The generation of a risk score about an individual, or their inclusion on a database with a particular label, is likely to engage Article 8 of the ECHR,[229] as would subsequent intrusive action. Being categorised in a way that the individual regards as inaccurate, incomplete or offensive may of itself constitute 'less favourable treatment' and therefore be discriminatory if demonstrated to be due to a protected characteristic. Ongoing monitoring and oversight is needed to actively scan for discrimination and assess how statistical models may engage protected characteristics.

[227] Roundtable event organised by RUSI in partnership with techUK, London, 22 July 2019.

[228] 'Equality Act 2010 (UK)', Sections 29(6) and 149.

[229] Catt vs. UK, European Court of Human Rights, 43514/15, 2019.

About the Authors

Alexander Babuta is a Research Fellow in National Security Studies at RUSI, where his research focuses on the use of emerging technologies for security purposes.

Marion Oswald is the Vice-Chancellor's Senior Fellow in Law at the University of Northumbria, an Associate Fellow of RUSI and a solicitor (non-practising). She is Chair of the West Midlands Police Data Ethics Committee, a member of the National Statistician's Data Ethics Advisory Committee, a member of the Ada Lovelace Institute Advisory Group on biometrics and an executive member of the British and Irish Law, Education and Technology Association.