Intelligence, policing and the use of algorithmic analysis: a freedom of information-based study

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

Intelligence, policing and the use of algorithmic analysis: a freedom of information-based study

Marion Oswald (University of Winchester) and Jamie Grace (Sheffield Hallam University)[1]

1. Introduction

This article explores some of the legal and societal issues around algorithmic and computational intelligence analysis used within policing in the United Kingdom; methods which interlink with the use of so-called 'Big Data' and are commonly badged as data science. These techniques are used in modern law enforcement as aides to decision-making and evaluation, mapping crimes, and even predicting where the next crimes are likely to occur, and hence, purportedly assisting the police to prevent crimes occurring. In an era of austerity in which police are under pressure to do more with less, these techniques promise to increase the efficiency and effectiveness of a police force. Conversely, the use of algorithms combined with the rise of Big Data presents society with a number of privacy and ethical concerns.

There has been a call for more predictive policing in the United Kingdom, using 'Big Data' techniques in intelligence analysis, as demonstrated by the National Policing Vision for 2016, which hopes that 'predictive analysis and real-time access to intelligence and tasking in the field will be available on modern mobile devices.'[2] This suggests the increasing importance of predictive policing, allowing forces to 'drive improvements in investigations, proactive patrolling, the protection of vulnerable people and the management of offenders and dangerous people.'[3] This statement indicates the variety of different aims to which algorithmic methods may contribute, both on the macro and the micro (individual) level, each of which involves differing benefits and risks. The use of these technologies in policing is not an area that lends itself to a one-size-fits-all approach in terms of application, regulation or oversight. Indeed, the terms 'algorithm', 'Big Data' and 'intelligence' are fluid ones, with all computer software essentially being of an algorithmic nature in the sense of a set of rules in a calculation or problem-solving operation. In the context of this article however, we are concerned with computational methods used to 'increase our knowledge of individuals, groups, and societies by use of data with an unprecedented breadth, depth and scale.'[4] This may include the analysis of 'data sets that are so copious and complex that they have to be processed by computing power to be useful'[5] or the analysis of information gathered through the investigative process in order to transform it into usable knowledge[6], to indicate links or suggest correlations.

The term 'intelligence' itself has no set meaning: 'Traditionally, 'intelligence' was understood to mean information from criminals about criminal activity by a covert source. Today, intelligence is a systematic approach to collecting information with the purpose of tracking and predicting crime to improve law enforcement.'[7] Intelligence can be seen as a form of information or the product of analysis and evaluation, or both.[8] Intelligence is important for

[1] The authors acknowledge the invaluable work of student research assistants J. Gillman, C. Redshaw and C. Wilde.
[2] College of Policing, National Policing Vision 2016 http://www.college.police.uk/About/Pages/National-policing-vision-2016.aspx (last accessed 19 July 2016).
[3] *Ibid.*
[4] Alex Pentland, *Social Physics: How Good Ideas Spread – The Lessons from a New Science* (The Penguin Press, New York, 2014) 217.
[5] Bernard E. Harcourt, *Exposed: Desire and Disobedience in the Digital Age* (Harvard University Press, 2015) 132.
[6] Petter Gottschalk, 'Information Sources in Police Intelligence' (2009) PJ 82 2 (149).
[7] *Ibid* citing Brown *et al*., 2004.
[8] See Jamie Grace and Marion Oswald, ''Being on our radar does not necessarily mean being under our microscope': The Regulation and Retention of Police Intelligence' EJoCLI (2016) 22(1).

its 'predictive value'; it is often incomplete or fragmented and so it must be critically assessed for its usefulness.[9] Intelligence is not pure 'committed crime' data, although by combining crime data with other information, intelligence may be produced. An example might be the fact that a person lives at a particular address; this information might become intelligence if it becomes known that such address has been the location of criminal activity such as drug dealing.[10] Intelligence is inherently uncertain and subjective (hence the National Intelligence Model 5×5×5 process under which the reliability of information is graded[11]). Intelligence does not necessarily pass any (legal) evidential test or threshold, and might not (yet) relate to a specific crime, threat or person. The rather amorphous and changeable nature of intelligence presents challenges for the regulation of its collection, production, retention and use, all the more so when combined with the abilities of algorithmic analysis to make predictions, generate presumptions and allocate risk (thus generating further intelligence). It is therefore an area, it is suggested, deserving of increased scrutiny.

Using freedom of information requests as a research methodology, this exploratory study had four main aims: (1) to understand the extent to which algorithmic or computational software is used to analyse police intelligence and the nature and purpose of this analysis; (2) to investigate the handling of intelligence across police forces in the UK; (3) to consider (in the context of oversight) any disciplinary action taken by forces in relation to incidents of inappropriate intelligence handling; and (4) to review any risks in current approaches or any differences in approach between forces. A freedom of information request methodology was chosen for three main reasons: first, the practical difficulties of discovering the required information from a search of each force's website or by securing access through negotiation with each force; secondly, the advantages of a standardised request in facilitating a qualitative and quantitative comparison of the responses; and thirdly, the benefits of freedom of information to individual researchers as a 'democratising force.'[12]

This article is structured as follows: in section 2, we explore the technological, legal and operational context to the use of algorithmic methods. Section 3 briefly sets out the method and limitations of the study, and then summarises the results focussing on four areas: the use of algorithmic methods by police forces; the creation of algorithms; intelligence handling processes; and disciplinary offences relating to intelligence. Section 4 expands the discussion of the results and we conclude in Section 5 with a number of recommendations.

2. Context of the study

Operational and legal background

There is increased emphasis in the UK on intelligence and data-led policing and the ability to link information so that concerning patterns of behaviour can be highlighted. Her Majesty's Inspectorate of Constabulary said, in its 2015 report into police information management:

> 'The whole picture may well be greater than the sum of its parts. This is why linking information and building the picture of the crime are so important – and why the

[9] Adrian James, *Understanding Police Intelligence Work* (Policy Press 2016) 24.
[10] 'Building the picture' Her Majesty's Inspectorate of Constabulary, July 2015, 19.
[11] http://library.college.police.uk/docs/APPref/5x5x5-Information-Intelligence-Report.pdf (last accessed 19 July 2016).
[12] Ashley Savage & Richard Hyde, 'Using freedom of information requests to facilitate research' International Journal of Social Research Methodology (2014) 17:3, 303-317, 304.

> consequences of failing to make the right links can have a significant adverse impact on the public.'[13]

HMIC laid emphasis on the availability of up-to-date information where needed through interoperable systems. It concluded that there remained:

> 'a real and pressing need for greater attention to be paid to the management of police information, so that greater consistency is achieved across all forces…local variations in practice carry real risks that the mistakes we identified in our report about police contact with Jimmy Savile could be repeated.'[14]

This present study was conducted against the backdrop of both the requirement for more insightful and powerful management of intelligence and the need for consistency across forces. It also builds upon Grace and Oswald's conclusion that 'criteria in case-law and regulatory guidance for intelligence retention lack clarity and coherence…[furthermore] not enough attention has yet been paid to existing and potential electronic data analysis techniques.'[15]

Intelligence analysis and the algorithmic use of police recorded data is one source of tension between privacy and freedom of expression, and security or public safety. While upholding the police's retention of information about Mr Catt, a non-violent protester, the UK Supreme Court considered in its judgment[16] the potential adverse consequences of overly restricting the availability of intelligence information. Lord Sumption noted (at para. 31) that [emphases added]:

> 'The composition, organisation and leadership of protest groups who are persistently associated with violence and criminality at public demonstrations is a matter of proper interest to the police even if some of the individuals in question are not themselves involved in any criminality. **The longer-term consequences of restricting the availability of this [information] resource to the police would potentially be very serious. It would adversely affect police operations directed against far less benign spirits than Mr Catt. Organised crime, terrorism, drug distribution and football hooliganism are all obvious examples.** One cannot look at an issue of this kind simply in relation to Mr Catt.'

The judgment can be criticised for setting potentially too low a threshold in the common law for information and intelligence appropriate for indefinite retention (particularly where, in Mr Catt's view, the retention associated him with violent extremists) and for paying too little attention to the implications of developments in data science that could change the nature of an apparently nominal or trivial piece of information.

Algorithmic analysis – purposes within policing

We suggest there are currently three main purposes for algorithmic data or intelligence analysis within the policing context: i) predictive policing on a macro level incorporating strategic planning, prioritisation and forecasting; ii) operational intelligence linking and evaluation which may include, for instance, crime reduction activities; and iii) decision-

[13] (n10) 8.
[14] (n10) 14.
[15] (n8).
[16] *R (on the application of Catt) (Respondent) v Commissioner of Police of the Metropolis and another (Appellants)* [2015] UKSC 9.

making or risk-assessments relating to individuals.[17] Each of these categories are explored briefly below and it is recognised that there may be some overlap between them and subsets of activity existing within each category. We have not used the term Big Data in this section as we agree with James that databases for policing purposes, although often large, do not generally fit with the unstructured and unpredictable nature of 'Big Data'.[18]

Macro level predictive policing systems use recorded crime and/or intelligence data to predict areas where offences are likely to take place, with the aim of assisting the police take decisions about where and when to allocate resources so acting as a crime disruption tool. An example is Predpol (developed by a company in Santa Clara and used in the UK by Kent Police) which incorporates recorded crime data and also considers factors such as the demographics of the area, the people, the places and the type of buildings in the locality. Predpol asserts that its software 'does not replace, but requires, the insights of veteran officers and crime analysts.'[19]

Operational intelligence analysis tools can analyse and compare large (bulk) databases, such as telephone or internet connection records, travel data, or databases of people with certain characteristics, for example those with air-side access at an international airport. These tools can identify threats, sometimes from fragments of intelligence, establish links between known subject of interests (and unknowns), and connect sometimes anonymous online personae to real world identities.[20] They also provide additional tools for the analysis of intelligence and other structured data sources in order to proactively identify patterns and connections that may indicate criminal or harmful behaviour (and indeed are so used by intelligence agencies in the protection of national security). In its enquiry into intelligence material concerning Jimmy Savile, Her Majesty's Inspectorate of Constabulary found that:

> 'the failure to connect the various allegations was critical to the eventual outcome of the investigations. There was intelligence available of four separate investigations which was never linked together and, because of that failure to 'join the dots', there was a failure to understand the potential depth of Savile's criminality.'[21]

The Bichard report into intelligence failures in connection with the Soham murders committed by Ian Huntley came to similar conclusions, 10 years before.[22] The undoubted cleverness of algorithmic intelligence analysis creates opportunities to remedy such failures. For instance, network analysis allows the police to 'visualize the density of connections an individual has within a social network'[23] thus identifying those within a group who might not have been otherwise apparent (in the language of Donald Rumsfeld, 'known unknowns'[24]).

Turning to our third category, there are algorithmic tools, in use in particular in a number of States in the USA, that are used to feed more directly into immediate decisions or judgements

---

[17] We are conscious that these three tiers or categories are similar to the concepts of primary, secondary and tertiary crime prevention measures, as outlined in Paul Brantingham & Frederic L. Faust, 'A conceptual model of crime prevention' Crime & Delinquency 22, no. 3 (1976): 284-296.
[18] (n9) 111.
[19] http://www.predpol.com/how-predpol-works/ (last accessed 19 July 2016).
[20] See UK Home Office 'Operational Case for Bulk Powers' 1 March 2016, 4 and case studies 45-47 https://www.gov.uk/government/publications/investigatory-powers-bill-overarching-documents (last accessed 19 July 2016).
[21] 'Mistakes were made' HMIC March 2013 https://www.justiceinspectorates.gov.uk/hmic/publications/mistakes-were-made/ (last accessed 19 July 2016).
[22] The Bichard Inquiry Report, London: The Stationary Office 2004 http://dera.ioe.ac.uk/6394/1/report.pdf (last accessed 19 July 2016).
[23] Elizabeth E. Joh, 'Policing by Numbers: Big Data and the Fourth Amendment.' (2014) Washington Law Review 89(35) 35-68, 46-47.
[24] Donald Rumsfeld, US Secretary of State for Defence, Defence Department briefing, February 2002.

about individuals, for instance a tool introduced in Chicago for predicting the individuals who are likely to be involved in gun violence[25] and software developed by a company called Northpointe[26] to assess recidivism risk and thus inform parole and sentencing decisions. Northpointe states that its formula includes factors such as whether the defendant has a job and their education levels, but that the specific calculations are proprietary.[27] The adoption of these tools is driven by worthy societal and policy objectives: the proactive reduction of gun crime; the management and reduction of the prison population; the search for 'objective' (and perhaps more cost-effective) ways of decision-making. Indeed, the UK government has for some years now invested in the use of database-driven means of predictive tools for offender management purposes in the prison and probation contexts.[28]

Legality, accountability and transparency

The use of the above mentioned tools raises important issues of legality, accountability and transparency.[29] Operational algorithmic intelligence analysis may require the acquisition of one or more bulk datasets containing information about 'innocent' individuals, thus raising legitimate privacy concerns, although it is beyond the scope of this article to consider these in detail.[30] Operational analysis of intelligence by the police may however be most commonly based on intelligence collected and retained as a result of traditional policing activities. However, algorithmic analysis of such intelligence can reveal or deduce sensitive information about people that they had thought hidden or certainly obscure. Information in the public domain or which an individual may think trivial can, when subjected to algorithmic analysis, reveal much more than the sum of its parts. Legal or regulatory approaches that focus on the *type* of information *collected*, or whether information is in the *public domain*[31] may need to be rethought. Although algorithmically generated information is unlikely to present the full picture and may be inaccurate, it will contribute to the investigatory toolkit and may form the basis for further decisions or actions including arrest. In the context of US protection against unreasonable search, Joh raises the question of the role that artificial intelligence and human judgement should play in Fourth Amendment individualized suspicion required to justify a search or seizure. Will a determination produced by artificial intelligence be accepted provided it was not the sole justification?[32] Similar considerations may well arise in determining reasonable grounds for suspicion under the UK's Police and Criminal Evidence Act 1984.[33] Edwards and Urquhart raise concerns that automated profiling and predictive

[25] Monica Davey 'Chicago police try to predict who will shoot or be shot' The New York Times, May 23 2016 http://www.nytimes.com/2016/05/24/us/armed-with-data-chicago-police-try-to-predict-who-may-shoot-or-be-shot.html?smprod=nytcore-ipad&smid=nytcore-ipad-share&_r=0 (last accessed 19 July 2016).
[26] http://www.northpointeinc.com/ (last accessed 19 July 2016).
[27] Julia Angwin and Jeff Larson, 'Machine Bias' ProPublica, May 23 2016 https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (last accessed 19 July 2016).
[28] See, for example, for details of the electronic Offender Assessment System (OASys) see: National Offender Management Service, *A compendium of research and analsysis on the Offender Assessment System (OASys): 2009-20013*, Ministry of Justice Analytical Series, July 2015, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/449357/research-analysis-offender-assessment-system.pdf (last accessed 19 July 2016).
[29] Lyria Bennett Moses and Janet Chan, 'Using big data for legal and law enforcement decisions; testing the new tools' (2014) UNSW Law Journal 37(2) 643-678, 678.
[30] See the commentary on the Investigatory Powers Bill at http://infolawcentre.blogs.sas.ac.uk/investigatory-powers-draft-bill-materials-and-commentary/ (last accessed 19 July 2016).
[31] (n23) 62.
[32] (n23) 55.
[33] 1984 c.60 http://www.legislation.gov.uk/ukpga/1984/60/contents (last accessed 19 July 2016).

policing based on social media and open source intelligence may 'lead to bad arrests, bad prosecutions and possibly even bad convictions.'[34]

Macro level predictive policing, although on the face of it the least contentious of the three categories, still gives rise to a number of significant considerations.[35] Moses and Chan argue that 'The fact that big data will have successes does not mean, however, that big data analysis is always successful'[36] and therefore claims of accuracy or utility should be treated with caution. Harkness argues that in some areas claims for big data are 'wildly overblown' reflecting a 'reductive view of humanity'.[37] The apparent success of systems such as PredPol is open to considerable dispute.[38] [39] [40] Greengard warns that 'some criminals may react to what the machines have learned and alter their behaviour and patterns to evade detection.'[41] In addition, will the prediction of a crime hotspot be self-fulfilling as the focus by police on the area will mean more arrests are made?[42] For those who live in the area, the designation as a hotspot may well become personal data for them,[43] personal data generated by a third party and most likely without the knowledge of the individual resident. For these reasons, Crawford and Schultz argue that, as part of a 'data due process right', governments should have to notify citizens about the use of predictive analysis and citizens should be able to determine if they are living in a 'hotspot'.[44]

In a different context, in terms of health data linkage, Grace and Taylor have argued that the interrelation of UK and European legal frameworks entail that a right to be notified of data linkage, and a potential right to object to that linkage, might well subsist in such scenarios.[45] Regardless of the precise legalities, however, and given that crime mapping and predictive crime data analysis is not specifically regulated by the law in the UK, public confidence in state accountability and transparency will be of equal if not more importance, as has been demonstrated in relation to the success or failure of health records linkage projects in the UK.[46]

Finally, this section will consider tools that directly inform (or perhaps even dictate) judgements about individuals, such as the algorithms mentioned above to assess reoffending risk. It is a well-established principle of natural justice that a person should be able to understand the process by which significant decisions are made that affect his freedoms, and be able to challenge those decisions. The increasing use of algorithmic decision-making presents a considerable challenge to these fundamental principles. Assuming, as has been

[34] Lillian Edwards and Lachlan Urquhart, 'Privacy in public spaces: what expectations of privacy do we have in social media intelligence?' International Journal of Law and Information Technology (2016) 24 279-310, 291.
[35] Kate Crawford and Jason Schultz, 'Big data and due process: Toward a framework to redress predictive privacy harms' (2014) BCL Rev. 55 93.
[36] (n29) 666.
[37] Timandra Harkness, *Big Data: Does Size Matter?* (Bloomsbury 2016) 284.
[38] *Ibid*, 239-240.
[39] Samuel Greengard, 'Policing the Future' (2012) Communications of the ACM 55: 3, 19-21.
[40] (n9) 116.
[41] (n39).
[42] (n35) 103.
[43] (n35) 104.
[44] (n35) 126.
[45] Jamie Grace and Mark Taylor, 'Disclosure of confidential patient information and the duty to consult: The role of the Health and Social Care Information Centre' Medical Law Review (2013) 21(3), 415-447.
[46] Sigrid Sterckx, et al. "'You hoped we would sleep walk into accepting the collection of our data": controversies surrounding the UK care. data scheme and their wider relevance for biomedical research' Medicine, Health Care and Philosophy (2016) 1-14.

claimed,[47] that the factors, values and biases incorporated into the software are often unknown to the police forces using the tool (with control over methods thus effectively being abdicated to another discipline), then decisions risk not only being wrong but also procedurally unfair. Challenges could arise under Article 6 of the European Convention on Human Rights (right to a fair trial) in relation in particular to equality of arms and how rules relating to access to evidence and questioning of witnesses might apply.

Northpointe's software has been accused of producing a racist result; others, including the company itself, deny that this is the case.[48] Without access to the underlying data and workings, in this situation proprietary to a commercial organisation, determining the truth will be impossible. Long term accountability is crucial; unlike human judges an algorithm 'is accountable to nobody.'[49] Moses and Chan query how conclusions from data analysis can be rebutted and raise the issue of the lack of a normative basis for future decision-making: 'A person can be told that they are not going to be released on parole because of their history of violent crime, but correlative statements do not necessarily provide a similarly sufficient explanation unless the causal link can be intuitively grasped.'[50]

3. The Freedom of Information requests and the responses received

Method

In February 2016, all police forces in the UK were contacted as part of the study, pursuant to the legal framework (the Freedom of Information Act 2000[51]) that requires a public body to respond to a freedom of information (FOI) request within a specified timescale. Forty-three responses were received, representing forty-five forces, due to four forces providing joint responses.[52] These responses were received during the period 3 March to 15 June 2016. The forces contacted included less-traditionally recognised forces such as the Port of Dover Police, British Transport Police and Civil Nuclear Constabulary as well as those from outside the jurisdiction of England and Wales; namely Police Scotland and the Police Service of Northern Ireland.

It was considered that taking a smaller sample with a view to generalising the findings across the UK would be an error as each police force operates separately and to ignore these individual remits would cause our findings to lack validity.

In gathering the data for the study itself, a standardised FOI request was created so that responses from separate forces could be compared. This allowed for an acceptable level of validity as well facilitating a level of replicability.

Limitations

Despite the legal obligation to reply to FOI requests, not all police forces in fact replied and so the study does not represent a complete perspective. Despite this, the response rate was 90% and includes responses from key (in terms of size and remit) police forces including

[47] Jathan Sadowski, 'Police Data could be labelling 'suspects' for crimes they have not committed' The Guardian 4 February. 2016 https://www.theguardian.com/technology/2016/feb/04/us-police-data-analytics-smart-cities-crime-likelihood-fresno-chicago-heat-list (last accessed 19 July 2016).
[48] (n27).
[49] (n37) 256.
[50] (n29) 675-676.
[51] 2000 c. 36 http://www.legislation.gov.uk/ukpga/2000/36/contents (last accessed 19 July 2016).
[52] Norfolk and Suffolk, and Warwickshire and West Mercia.

Metropolitan Police, Greater Manchester, West Midlands, Thames Valley, West Yorkshire, Police Service of Northern Ireland and Police Scotland.

While the FOI request attempted to include explanation of the context of the request and clarification of key terms, this is an area that is inevitably open to subjective interpretation by the responder and therefore the consistency of responses may have been affected.[53] Indeed a number of the FOI responses referred to crime data rather than intelligence. In addition, a number of forces engaged exemptions under the Act to justify the withholding of certain information on policing or national security grounds and therefore the responses received are not likely to represent the full picture.

Results

Algorithmic intelligence analysis

Forces were asked whether they used any sort of computational or algorithmic data analysis or decision-making in relation to the analysis of intelligence, and to confirm the nature and purpose of any such algorithms. Six responses (14.0%)[54] indicated that they did use some form of computational or algorithmic intelligence analysis or decision making software. This meant the majority (86%) who responded answered that they did not.

In terms of the nature of the algorithmic/computational software, although there was a certain commonality in that all were used in some form or another as 'analytical tools', each of the six responses that replied positively cited different further uses including:

- 'profiling individual risks through to broader forecasts';
- 'assessing risk and crime patterns';
- to 'link entities known within our crime and intelligence system';
- the use of two systems with one system that 'cleans raw communication information into a format that can be used evidentially' and another system that encompasses a 'search facility that links other computer programmes for the ease of analysis';
- the use of 'macros to better identify series and themes and key words to abstract data' from their systems.

One force serving an area with particular challenges in terms of counter-terrorism responsibilities and community policing, stated that it uses a number of algorithmic tools with the principal aims of evaluating intelligence and providing an accurate assessment in respect of risk and harm associated with crime and the organised crime groups engaged in this criminality. This would lead to better prioritisation of tasks and consequently lead to a 'more informed collection of intelligence'.

In the six responses that stated that their force(s) did use some form of algorithmic or computational software for the purpose of intelligence analysis or decision making, there were a number of identifiable common themes concerned with the purpose of the software. Three forces mentioned the use of the software as an aid to traditional police intelligence analysis, as a ‘business insight tool’ and for assessing crime patterns. A further theme was use as a tool in

[53] For an excellent discussion of the challenges around drafting the request, see Ashley Savage & Richard Hyde, ‘Using freedom of information requests to facilitate research’ International Journal of Social Research Methodology (2014) 17:3, 303-317, 307-308.

[54] Avon and Somerset Constabulary, Warwickshire Police and West Mercia Police (a joint response), Devon and Cornwall Police, Humberside Police, North Yorkshire Police and the Police Service of Northern Ireland.

'prioritising' work for the intelligence analysis team to carry out, mentioned by three forces. Similarly, four forces wrote that the purpose of the software was to assess and rank risk.

Creation of algorithms

If forces used any sort of computational or algorithmic data analysis in relation to the analysis of intelligence, they were asked to confirm who created those algorithms. Of the six responses that indicated that they used some form of algorithmic analysis, three did not provide an answer or wrote 'n/a'. One force reported that their algorithmic and computational methods were created by an in-house trained member of staff. Another force stated that their systems were created by private companies and then modified for use within the force. The third response implied that the algorithms were created internally.

Intelligence handling processes

All forces that responded said that they used some form of nationally-recognised database and framework for the handling and analysis of intelligence, particularly the Police National Database (PND), which handles intelligence; and/or the National Intelligence Model (NIM), the Management of Police Information guidance (MOPI), or the Management of Risk in Law Enforcement (MoRiLE) and 5x5x5 intelligence analysis models. However, there were a few forces that said they used some other software or databases.

One force mentioned that it was collaborating with five other forces on a 'niche system'. This is likely to be the records management system manufactured by Niche Technology Inc.[55] and which was mentioned in four other responses. The stated purpose of this system is to provide uniformity in process and procedure when dealing with intelligence. The Chief Executive of the recently established Police ICT Company, commenting on the 'Niche' records management system, stated 'they were able to solve a crime that they could have not done before because of that [system].'[56]

Another interesting finding was that some of the police forces outside of England were more likely to describe the use of localised or activity-specific software or databases which may reflect the jurisdictional separations in the UK. For example, Police Scotland cited the use of the Scotland Intelligence Database (SID) which they described as a database that collates and manages intelligence, follows NIM guidelines and uploads information to the PND daily.[57] Other forces outside England mentioned 'internal policy directives', Organised Crime Group Mapping, Child Sexual Exploitation Monitoring, 'Acquisitive Crime Trends' and an internally created 'Intelligence Triage Model' and it seems likely that these methods would also be used in other areas.

A further theme was the disparity of detail provided by different forces. Certain forces provided richly detailed and useful answers that gave a much clearer response than other forces that merely provided a few sentences. The answers from two forces, consisting of 636 words and 1,068 words respectively, contrast with the response from a 'lead' force of just 47 words with no actual mention made of the use of PND within the response! Two forces used

[55] http://nicherms.com/ (last accessed 19 July 2016).
[56] http://police.governmentcomputing.com/news/police-ict-company-chief-worries-about-body-worn-video-procurement-4929388 (last accessed 19 July 2016).
[57] See SID case study from commercial developer ABM http://www.abm-uk.com/pdf/ABM_CaseStudy_SID_PDF.pdf (last accessed 19 July 2016).

the phrase 'crime data' in place of intelligence, suggesting that there is no single or agreed definition of 'intelligence' or clear differentiation with crime data in this context.[58]

Disciplinary action relating to intelligence processes

Properly managing police-recorded data and intelligence during collection, recording, retention and analysis is a careful and vital business. Allan Brimicombe has noted that:

> 'The personal details present in police-recorded data means that they are subject to the provisions of the 1998 Data Protection Act and disclosure must be prevented. Breaches reported to the Information Commissioner's Office (ICO), such as loss or theft of sensitive personal data will attract a financial penalty of up to £500,000. Not surprisingly high security IT systems operating within a well-founded information security framework are required for the storage, analysis and eventual archiving or destruction of such data in order to prevent breaches. Individuals with access to the data also usually need to be vetted by the relevant police force(s) and may be asked to sign the Official Secrets Act.'[59]

As part of our study, police forces were asked to indicate the number of officers and staff disciplined for not following the force's intelligence processes for years 2013-14 and 2014-15 and the nature of the failure. Twenty-six respondents said that they were able to provide information relating to the question. However, seventeen said they could not provide information for cost reasons or because they did not hold information relevant to the request. Where cost was cited, this was commonly due to the lack of an appropriate categorisation for disciplinary actions relating to intelligence. Thus it was argued by the relevant forces that the cost of the manual search necessary to locate the relevant information would exceed the limit under section 12 of the Freedom of Information Act.

Of the twenty-six responses that said they held information relevant to the request, fifteen reported that officers had been disciplined for intelligence process related wrongdoings. There were two common themes for such disciplinary action. The first theme regarded the misuse of police systems, which in the responses was characterised by the disciplining of individuals for 'unauthorised checks' and/or 'accessing information for non-policing purposes'. The second theme regarded the failure to comply with usual practice which was characterised by responses that highlighted individuals' 'failure to submit intelligence' and/or 'neglect of duty'.

One finding that became evident in the analysis of this question was the lack of clarity and consistency between the forces in both their answers and their categorisation of what exactly they viewed as a failure to follow intelligence processes. For example, one force was able to identify the number of people disciplined, the year of the misconduct and the disciplinary outcomes, whereas other forces lacked clarity and focus in their responses.

4. Discussion

The indication from the study that 86% of UK police forces do not currently use any algorithmic or computational intelligence analysis may seem surprising, when viewed in the context of the ongoing policy objective to improve police information management and intelligence linking. Of course, this percentage may not represent the complete picture; for

[58] Supporting conclusions in (n8).
[59] See Allan Brimicombe, 'Analysing Police-Recorded Data' Legal Information Management (2016) 16.2 71-77, 73.

understandable reasons, the exemptions engaged may disguise other uses of such tools, or indeed further gaps in operational capability.

The six responses that did acknowledge the use of algorithmic methods suggested that tools were used for all three of the purposes set out in section 2 of this article i.e. predictive policing on a macro level; operational intelligence linking and evaluation; and decision-making or risk-assessments relating to individuals. However, the reasons for using the technology were expressed in general terms – assessing risk, linking entities and so on. There was no clear indication as to the specific activities, crimes, schemes or laws that were the focus of these tools.

Are, for instance, algorithmic tools currently or planned to be used in the application of the preventive, predictive Domestic Violence Disclosure Scheme (known as 'Clare's Law')? This is a scheme under which the police choose whether to supply public protection risk information, based on an actuarial judgment with regard to the particular risk of the requesting member of the public coming to harm through domestic violence.[60] For this Scheme, and others such as the Child Sex Offender Disclosure Scheme ('Sarah's Law'), as well as the widespread use in employment vetting via Enhanced Criminal Record Certificates,[61] where difficult risk-based judgements are required, it is easy to see why a powerful and *reliable* algorithmic decision-making tool would be attractive and potentially helpful provided used as a complement to human judgement. However, any 'hard' implementation of such a tool for these Schemes, resulting in less human judgment and more machine-driven decisions about information sharing for public protection, may well raise questions as to whether the police were breaching principles of natural justice in UK public law terms, or disproportionally interfering with the right to private and family life, in the language of European human rights law.[62] How would use of a tool designed by a commercial corporation be regarded, too, given that 'companies designing software for law enforcement recognise that their systems do not necessarily conform to practical or ethical standards'?[63] Similar tools in the US have been criticised for entrenching racial and other discriminatory profiling, thus overestimating the risk presented by some individuals and underestimating the risk posed by others.[64] As Zliobaite and Custers argue, avoidance of such a result may not be as simple as merely removing the sensitive characteristics from the algorithm; automated decision-making models could still produce discriminatory results, because data mining methods are based upon assumptions that are not satisfied in reality.[65]

Our findings exposed differing approaches within police forces towards transparency about intelligence, with some forces giving helpful, expansive answers, others single line responses. Some forces were able to respond without issue to the question concerning disciplinary actions relating to intelligence. Others could not, suggesting the need for an agreed disciplinary offence category for the misuse of police intelligence and databases/algorithms. We have seen in the area of investigatory powers exercised by intelligence agencies that lack

[60] See Jamie Grace, 'Clare's Law, or the national Domestic Violence Disclosure Scheme: the contested legalities of criminality information sharing' The Journal of Criminal Law (2015) 79 (1), 36-45.
[61] See (n8).
[62] Claiming that data linkage or extraction and analysis for public protection purposes in criminal justice settings is suitably *necessary*, in the legal sense of the term, is not always straightforward. The judgment of the European Court of Human Rights in *Avilkina v Russia* (application no. 1585/09) (2013) demonstrated that the targeted extraction of health data, for example, from hospital records cannot lawfully be used to obtain intelligence with the aim of mounting politically-motivated prosecutions.
[63] (n35) 105.
[64] (n27).
[65] Indre Zliobaite and Bart Custers, 'Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models' Artificial Intelligence and Law (2016) 24(2) 183-201.

of *appropriate* transparency around methods combined with vague legal underpinning can result in damage to public trust as well as legal challenge. As Sir David Omand remarked 'The shock of discovering what was happening, for very good reason—to defend the public and our security—was all the greater.'[66]

In addition, and just as importantly we would argue, algorithms will only be as good as the data available: 'Extensive data cleaning is required to maximise the analytical use of the data', as Brimicombe has observed in relation to police data analysis.[67] The never-ending development of technology (for instance the roll-out of body-worn cameras[68]), as well as the legal turmoil expected in the post-Brexit UK (with the trend toward stronger data protection regulation possibly being reversed), will mean that more data will be collected by the police in the future, not less. How this data is used will be crucial. The eight precursor Scottish forces that are now part of Police Scotland use one intelligence management system (SID), and a number of forces in England are working together to implement a new records management system. There can be benefits in a local or regional approach to data management: flexibility; the ability to focus on local priorities; quicker procurement processes. Our study raises questions however about the extent of joined-up thinking, in England and Wales in particular, around operational level intelligence linking and records management. As well as increasing the risk of legal challenge, any significant variations in functional approaches must, we believe, risk holding back progress on operational imperatives. As James has argued, '[t]here is just as much of a need to improve the extrapolation of meaningful data from existing structured databases as there is for the development of big data analytics.'[69]

5. Conclusions and recommendations

Algorithmic tools offer operational opportunities for policing, in particular to remedy the intelligence linking failures of the past. Learning lessons from the recent Investigatory Powers Bill debate in the UK however, it is arguable that the legal, ethical and policy issues surrounding the use of such tools by the police need to be tackled transparently and without delay. Statutory guidance on police algorithmic analysis of intelligence would be welcome, but should create regulatory nuance and precision. A one-size-fits-all approach should be avoided as each category of algorithmic usage raises particular issues and concerns.

Legality: Necessity and proportionality

The necessity and proportionality of intelligence collection, retention, use and disclosure should not now be judged merely on the basis of the potentially intrusive nature of the information itself, or on whether it represents something in the public domain, but rather as to what the information might represent, indicate or disclose when incorporated into algorithmic analysis. This then suggests that we need additional regulation of intended police *uses* of information, not only regulation of how information is *acquired*. It should be possible to craft the legal and other regulation of intelligence analysis and the algorithmic use of police recorded data, more clearly and in a more nuanced way over time - principally because not all

[66] Professor Sir David Omand, Oral Evidence to the Joint Committee on the Draft Investigatory Powers Bill, 7 December 2015 https://www.parliament.uk/documents/joint-committees/draft-investigatory-powers-bill/oral-evidence-draft-investigatory-powers-committee.pdf (last accessed 19 July 2016).
[67] See (n59) 74.
[68] The Chief Executive of the recently established Police ICT Company, commenting on the procurement of body-worn video equipment, stated that there was a wide variation in procurement practices and 'little thought given to data, its purpose, and its use in sharing and storage, which gives real issues in terms of how we interface with the criminal justice system.' http://police.governmentcomputing.com/news/police-ict-company-chief-worries-about-body-worn-video-procurement-4929388 (last accessed 19 July 2016).
[69] (n9) 145.

intelligence analysis, or the algorithmic use of police recorded data, takes place in the same manner or for the same purpose(s). By way of a parallel example, at the time of writing the College of Policing in the UK have put forward for consultation a draft Code of Practice (an 'Authorised Professional Practice' document, in College parlance) for Undercover Policing.[70] This is in direct response to sustained and heavy criticism in recent years of the way in which highly-intrusive, even reckless undercover policing in the UK, particularly in certain sensitive cases, has been under-regulated for decades.[71]

Accountability: Oversight of algorithmic methods

Crawford and Schultz argue for the introduction of a third party arbiter who 'could examine the relationship between those who designed the analytics and those who run the individual processes to make sure that their roles are appropriate and distinct.'[72] We support this view and would go further to suggest that such an arbiter could have a role in auditing algorithmic methods, factors and underlying calculations, together with the reliability of the data assessed, so that they can be challenged if necessary. Consideration could be given to an advance approval process for algorithms used in the public sector for risk-assessment and decision-making purposes, in a similar way to medical devices.[73] This would mean that public sector procurement contracts with third party suppliers must require a certain level of transparency from the commercial party, including disclosure of the algorithmic workings in a way that would facilitate investigation and onward disclosure. Accountability and transparency requirements might even lead towards the conclusion that open source algorithms should be used by default. Commercial confidentiality cannot be permitted to be a barrier to appropriate scrutiny, particularly in a democratic polity that has suffered a number of recent crises in the public consciousness, over police accountability, of late.

Transparency

Finally, we return to transparency. We appreciate and agree that transparency can only go so far in the area of policing in order that operational details, methods and gaps in capabilities are protected. Yet balanced, carefully calibrated, transparency is vital, and is possible in the UK context. As Dawson and Stanko have highlighted, any 'lack of transparency (in the form of basic descriptions of what these [police intelligence analysis] systems are and what information they hold) hinders analytic output, academic cooperation and perhaps even fuels wider public scepticism around the police and accountability'; importantly good practice exists in pockets, for example in Canada, where, while a detailed understanding of the Violent Crime Linkage System (ViCLAS) 'is quite rightly not in the public domain due to investigative sensitivities - but official, clear, easy to find information on what the system is, who can use it, how to go about access and related research *is available*'.[74]

We would suggest that those responsible for overarching policy developments in this area might wish to give additional consideration to how appropriate transparency can be improved, thus giving the public an opportunity to engage in an informed debate over current

---

[70] See http://www.college.police.uk/News/College-news/Pages/undercover-policing-guide.aspx and http://www.college.police.uk/News/College-news/Documents/Undercover_policing_guidance-for_consultation.pdf (last accessed 19 July 2016).
[71] https://www.theguardian.com/uk/undercover-police-and-policing (last accessed 19 July 2016).
[72] (n35) 127.
[73] In the UK, approval is given by the Medicines and Healthcare products Regulatory Agency https://www.gov.uk/topic/medicines-medical-devices-blood/medical-devices-regulation-safety (last accessed 19 July 2016).
[74] Paul Dawson and Elizabeth A. Stanko, 'The best-kept secret(s) of evidence based policing' Legal Information Management (2016) 16.2 64-71, 67.

and potential future uses of algorithmic methods. Recent research around the UK Government's Framework for Data Science Ethics indicated that public awareness of data science was limited; opportunities for data science were seen as greatest where 'there was a clear remit for government, where the current status quo was seen as inadequate, where the outcome was appropriate, and where data science could complement other methods.'[75] People had doubts about whether computers could make better decisions than humans, and were cautious about techniques that clustered individuals or used correlations between datasets that initially appeared unrelated.[76] The research recommended that the risk and proportionality of 'outcomes' be more clearly recognised, and further clarity achieved around the standards governing the government's data science models.[77]

Expanding such research engagement into the policing context will be crucial, ultimately, since as Brimicombe concludes, with words that could be as much about predictive, algorithmic policing intelligence analyses as they are about data-driven police research:

> 'The more in-depth analyses undertaken using police-recorded data, including using new and novel techniques, the greater will be the recognition of their value beyond routine operations and worth the effort by front-line officers to invest in data quality, which would in turn further enhance the value of the data. Police-recorded data can then become a reliable source of evidence in experimental and quasi-experimental evaluation of what works in policing and crime prevention.'[78]

[75] Ipsos Mori Social Research Institute 'Public dialogue on the ethics of data science in government' May 2016, 3-4 https://www.ipsos-mori.com/Assets/Docs/Publications/data-science-ethics-in-government.pdf (last accessed 19 July 2016).
[76] *Ibid*, 4.
[77] *Ibid*, 6.
[78] (n59) 77.