Executive summary

- In the UK policing context, the use of algorithmic decision-making tools could be described as being in a developmental stage with implementation on a force by force basis;
- Such tools may be used in a number of different contexts, including decision-making or risk-assessments relating to individuals;
- ‘Algorithms in Policing –Take ALGO-CARE™’ is a proposed decision-making framework for the deployment of algorithmic assessment tools in the policing context;
- Algo-care aims to translate key public law and human rights principles into practical considerations and guidance that can be addressed by public sector bodies;
- Concerns around transparency and accountability cannot be addressed by a one-size-fits-all approach;
- The factors identified by Algo-care necessitate the careful drafting of procurement contracts with third party software suppliers to require disclosure of algorithmic workings in a way that would facilitate investigation;
- A number of challenges remain to the satisfactory audit and validation of machine learning algorithmic tools.

Introduction

1. This submission summarises the output of a recent collaboration between the authors to reflect upon the deployment of an algorithmic tool within Durham Constabulary and the development of a general decision-making framework to guide users in the deployment of algorithmic tools in the policing context. The submission also includes aspects of the corresponding author’s research into the use of algorithmic tools within policing, including a freedom of information based study.

2. In this submission, we define an algorithm as a mathematical formula implemented by technology: ‘a sequence of instructions that are carried out to transform the input to the output.’ (Al Paydin, 2016) We are concerned with machine learning whereby the computer learns and extracts the algorithm for the task from the given input data. We do not comment on coded rules, programmed logic or database interrogation or linking.

Extent of current and future use of algorithms in decision-making

3. This submission comments on the use of algorithmic decision-making tools within the policing and criminal justice context. In the UK policing context, the use of algorithmic decision-making tools could be described as being in a developmental stage, with decisions on implementation being taken on a force by force basis. As UK policing becomes more aware of the capability of such algorithmic tools in comparison to current practice, these types of tools can often be viewed as no more than modern decision support tools.

4. It has been suggested that 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-making or risk-assessments relating to individuals. (Oswald, Grace, 2016)

5. One UK force has been reported to be making substantive use of a predictive policing tool developed by the private sector (‘PredPol’, implemented by Kent Constabulary) in order to predict areas where offences are likely to take place. A recent freedom of information-based study suggested a relatively small number of UK police forces (14%) were using computational or algorithmic data analysis or decision-making in relation to the analysis of intelligence, with tools stated to be used for all three of main purposes mentioned above. (Oswald, Grace, 2016)

6. Durham Constabulary has developed an algorithmic risk-assessment tool in category iii) above (decision-making or risk-assessments relating to individuals). The tool was developed by statistical experts based at the University of Cambridge in collaboration with the constabulary. It has been developed to aid decision-making by custody officers when assessing the risk of future offending and to enable those suspects forecast as moderate risk to be eligible for the constabulary’s Checkpoint programme. Checkpoint is an intervention currently being tested in the constabulary and is an out of court disposal aimed at reducing future offending.

7. The Durham tool uses the ‘random forest’ method, a machine learning algorithm that uses many independent decision-trees with random selections of data as training and the combined result (the forest) for the forecast. It has thirty-four ‘predictors’, the majority of which relate to the suspect’s previous offending behaviour together with age, gender, residential postcode, and number of intelligence reports. The tool does not currently access data held on the Police National Computer or Police National Database, or indeed other information contained in police systems such as incidents or the content of intelligence reports for example. The model has been designed to overestimate the risk using a precautionary approach and so minimise the numbers of false negatives (where an offender was forecasted low risk but was actually high risk of committing a serious offence) which are regarded as the most dangerous scenarios. (Please also see the parallel submission relating to the Checkpoint programme and the development of the related algorithmic tool).

8. It is understood that other UK forces are considering the development of similar tools, although this may be in connection with different programmes or contexts, with potential for such tools to be implemented to prioritise investigative actions or where the police have to decide whether to supply public protection risk information, based on an actuarial judgement (such as ‘Clare’s Law’). For schemes where difficult risk-based judgements are required, a fair and trustworthy algorithmic decision-making tool may potentially be helpful, provided not used in a determinative way.

9. In the USA, algorithmic tools have been introduced that are used to feed more directly into immediate decisions or judgements about individuals. One such tool was introduced in Chicago to predict those individuals who are likely to be involved in gun violence, and software developed by a company called Northpointe is being used 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. In the USA, algorithmic risk assessment tools were initially used only by probation and parole departments but have now expanded to bail hearings and sentencing (Barnes et al., 2012).
10. The Supreme Court of Wisconsin has recently ruled on the challenge by convicted criminal Eric Loomis to the use of Northpoint’s algorithmic risk assessment tool called ‘COMPAS’ during his sentencing. Although Loomis ultimately failed in his challenge, the court held that the tool should be subject to a number of cautions that a court should take into account as follows:

- the proprietary nature of COMPAS prevents disclosure of information relating to how factors are weighed or risk scores are determined;
- the scores are based on group data, and so are not able to identify a particular high-risk individual;
- concerns have been raised about disproportionate classification of ethnic minority offenders as high risk;
- the scores are based on a national sample; there had been no cross-validation for a regional/State sample;
- COMPAS was not developed for use at sentencing.

11. The above factors identified by the Wisconsin Supreme Courts usefully highlight a number of the risks and concerns generated by the use of algorithmic tools in the policing context, particularly where individuals may be impacted on a micro level. The following section sets out proposals for a framework by which key concerns can be addressed both before and during the deployment of an algorithmic decision-making tool.

‘Good practice’ in algorithmic decision-making

12. This section sets out a proposed decision-making framework for the deployment of algorithmic assessment tools in the policing context developed by the corresponding author in collaboration with Durham Constabulary. The framework – ‘Algorithms in Policing –Take ALGO-CARE™’ - reflects the experience of Durham Constabulary in developing and rolling out its algorithm associated with the Checkpoint programme. It also aims to translate key public law and human rights principles into practical considerations and guidance that can be addressed by public sector bodies.

13. While the authors note that a number of organisations are developing, or advocate developing, other high level principles in respect of algorithms and A.I. (which can be helpful to represent ethical norms and in setting a general direction of travel), we would submit that they often do not provide enough practical certainty for the development of administrative and assessment frameworks (or for practitioners to refer to in their day-to-day work). Algo-care aims to address these concerns, and to provide a decision-making framework that could work in different policing contexts, and potentially more widely across the public sector.

14. The current working version of ‘Algorithms in Policing –Take ALGO-CARE™’ is set out in Fig. 1 below, together with additional explanatory notes (Fig. 2). It is also set out in JPEG format in Appendix 1. Each word in the mnemonic – Advisory; Lawful; Granularity; Ownership; Challengeable; Accuracy; Responsible; Explainable – is supplemented by questions and considerations representing key legal considerations (such as necessity and proportionality, natural justice and procedural fairness), as well as practical concerns such as intellectual property ownership and the availability of an ‘expert witness’ to the tool’s functionality. The framework could be used in parallel with privacy/equality impact assessments.
**Algorithms in Policing – Take ALGO-CARE™**

A proposed decision-making framework for the deployment of algorithmic assessment tools in the policing context

<table>
<thead>
<tr>
<th>A</th>
<th>Advisory</th>
<th>Is the assessment made by the algorithm used in an advisory capacity? Does a human officer retain decision-making discretion? What other decision-making by human officers will add objectivity to the decisions (partly) based on the algorithm?</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>Lawful</td>
<td>On a case-by-case basis, what is the policing purpose justifying the use of algorithm, both its means and ends? Is the potential interference with the privacy of individuals necessary and proportionate for legitimate policing purposes? In what way will the tool improve the current system and is this demonstrable? Are the data processed by the algorithm lawfully obtained, processed and retained, according to a genuine necessity with a rational connection to a policing aim? Is the operation of the tool compliant with national guidance?</td>
</tr>
<tr>
<td>G</td>
<td>Granularity</td>
<td>Does the algorithm make suggestions at the right level of detail/granularity, given the purpose of the algorithm and the nature of the data processed? Do the benefits outweigh any technological or data quality uncertainties or gaps? Is the provenance and quality of the data sufficiently sound? Consider how often the data should be refreshed. If the tool takes a precautionary approach towards false negatives, consider the justifications for this.</td>
</tr>
<tr>
<td>O</td>
<td>Ownership</td>
<td>Who owns the algorithm and the data analysed? Does the force need rights to access, use and amend the source code and data analysed? How will the tool be maintained and updated? Are there any contractual or other restrictions which might limit accountability or evaluation? How is the operation of the algorithm kept secure?</td>
</tr>
<tr>
<td>C</td>
<td>Challengeable</td>
<td>What are the post-implementation oversight and audit mechanisms e.g. to identify any bias? Where an algorithmic tool informs criminal justice disposals, how are individuals notified of its use (as appropriate in the context of the tool’s operation and purpose)?</td>
</tr>
<tr>
<td>A</td>
<td>Accuracy</td>
<td>Does the specification match the policing aim and decision policy? Can the stated accuracy of the algorithm be validated reasonably periodically? Can the percentage of false positives/negatives be justified? How was this method chosen as opposed to other available methods? What are the consequences of inaccurate forecasts? Does this represent an acceptable risk (in terms of both likelihood and impact)? Is the algorithmic tool deployed by those with appropriate expertise?</td>
</tr>
<tr>
<td>R</td>
<td>Responsible</td>
<td>Would the operation of the algorithm be considered fair? Is the use of the algorithm used in an advisory capacity? Does a human officer retain decision-making discretion? What other decision-making by human officers will add objectivity to the decisions (partly) based on the algorithm?</td>
</tr>
</tbody>
</table>
algorithm transparent (taking account of the context of its use), accountable and placed under review alongside other IT developments in policing? Would it be considered to be for the public interest and ethical?

| E | Explainable | Is appropriate information available about the decision-making rule(s) and the impact that each factor has on the final score or outcome (in a similar way to a gravity matrix)? Is the force able to access and deploy a data science expert to explain and justify the algorithmic tool (in a similar way to an expert forensic pathologist)? |

---

Fig. 2

**Brief explanatory notes and additional considerations**

The Algorithms in Policing – Take ALGO-CARE ™ framework is intended to provide guidance for the use of risk-assessment, predictive, forecasting, classification, decision-making and assistive policing tools which incorporate algorithmic machine learning methods and which may impact individuals on a micro or macro level.

| A | Advisory | Care should be taken to ensure that an algorithm is not inappropriately fettering an officer's discretion, as natural justice and procedural fairness claims may well arise. Consider if supposedly advisory algorithmic assessments are in practice having undue influence. If it is proposed that an algorithmic decision be automated and determinative, is this justified by the factors below? Data protection rights in regard to automated decisions may then apply. |

| L | Lawful | The algorithm’s proposed functions, application, individual effect and use of datasets (police-held data and third party data) should be considered against necessity, proportionality and data minimisation principles, in order to inform a ‘go/no-go’ decision. In relation to tools that may inform criminal justice disposals, regard should be given to the duty to give reasons. |

| G | Granularity | Consideration should be given to common problems in data analysis, such as those relating to the meaning of data, compatibility of data from disparate sources, missing data and inferencing. Do forces know how much averaging or blurring has already been applied to inputs (e.g. postcode area averages)? |

| O | Ownership | Consider intellectual property ownership, maintenance of the tool and whether open source algorithms should be the default. When drafting procurement contracts with third party software suppliers (commercial or academic), require disclosure of the algorithmic workings in a way that would facilitate investigation by a third party in an adversarial context if necessary. Ensure the force has appropriate rights to use, amend and disclose the tool and any third party data. Require the supplier to provide an ‘expert’ witness/evidence of the tool’s operation if required by the force. |

| C | Challengeable | The results of the analysis should be applied in the context of appropriate professional codes and regulations. Consider whether the application of the algorithm requires information to be given to the individual and/or legal |
Written evidence submitted by Marion Oswald, Senior Fellow in Law and Director of the Centre for Information Rights, University of Winchester, and Sheena Urwin, Head of Criminal Justice, Durham Constabulary (ALG0030)

<table>
<thead>
<tr>
<th>A</th>
<th>Accuracy</th>
<th>How are results checked for accuracy, and how is historic accuracy fed back into the algorithm for the future? Can forces understand how inaccurate or out-of-date input data affects the result?</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>Responsible</td>
<td>It is recommended that ethical considerations, such as consideration of the public good and moral principles (so spanning wider concerns than legal compliance) are factored into the deployment decision-making process. Administrative arrangements such as an ethical review committee incorporating independent members could be established for such a purpose (such as Cleveland &amp; Durham Joint External Ethics Committee or NSDEC).</td>
</tr>
<tr>
<td>E</td>
<td>Explainable</td>
<td>The latest methods of interpretable and accountable machine learning systems should be considered and incorporated into the specification as appropriate. This is particularly important if considering deployment of ‘black box’ algorithms, where inputs and outputs are viewable but internal workings are opaque (the rule emerges from the data analysis undertaken). Has the relevant Policing &amp; Crime Commissioner been briefed appropriately?</td>
</tr>
</tbody>
</table>

**Transparency, accountability and commercial confidentiality**

15. The authors would suggest that concerns around transparency and accountability cannot be addressed in a one-size-fits-all way. It would not be appropriate, for instance, for the functionality of a tool used in the investigative process to be ‘transparent’ in the sense of the detailed functionality being publicly available. Where a tool assists in the decision-making about an out-of-court disposal, however, information about its use should be made available to the affected individual and/or his legal adviser (to address Article 6/right to a fair hearing concerns).

16. *Algo-care* identifies that, at the very least, the public body should be able to explain the decision-making rule(s) and the impact that each factor has on the final score or outcome, and ensure that it has access to and can deploy a data science expert to explain and justify the algorithmic tool (in a similar way to an expert forensic pathologist). The framework also notes that development specifications should incorporate as appropriate the latest methods of interpretable, interactive and accountable machine learning systems (see for instance Kroll et al., 2017).

17. These factors necessitate the careful drafting of procurement contracts with third party software suppliers (commercial or academic). Contracts should require disclosure of the algorithmic workings in a way that would facilitate investigation by a third party in an adversarial context if necessary (and the provision of an expert witness/evidence of the tool’s operation). In addition to appropriate rights to use, amend and disclose the software tool, public sector bodies should pay attention to rights over any third party data that have been used as inputs, such as mosaic postcodes. Although *Algo-care* identifies that open source software as default should be considered, it is appreciated that
access to the source code does not necessarily, of itself, result in an appropriately understandable and challengeable tool. Such access could however aid validation exercises for accuracy and bias.

18. The Wisconsin court in Loomis dismissed the argument that unless he could review how the factors are weighed and how risk scores are determined by the tool, the accuracy of the assessment cannot be verified. The court stated:

‘Although Loomis cannot review and challenge how the COMPAS algorithm calculates risk, he can at least review and challenge the resulting risk scores set forth in the report.’

19. The corresponding author has previously argued that it seems unlikely that the UK Supreme Court would allow the secret workings of a proprietary risk-assessment algorithm to form part of a court’s sentencing deliberation as the Wisconsin Supreme Court has done (Oswald, Grace, 2016). As the Doody decision held that a Home Secretary must show ‘how his mind is working’ in setting the tariff for life sentence prisoners, the same is likely to be true of an algorithm working in a similar context. Where risk-based decisions have an immediate and direct impact upon individuals, again it seems unlikely that commercial confidentiality would be permitted to be a barrier to appropriate scrutiny.

Audit and validation

20. The Algo-care framework identifies the need for regular audit and validation of the chosen model in terms of accuracy and in order to identify any bias or discrimination. The Durham model has highlighted the difficulties of validating situations where there is a significant discrepancy between the assessments i.e. the custody officer assesses 'high', the algorithm 'low' or vice versa. If the decision in these cases is always to act upon the high risk assessment, it will then be difficult to assess accuracy if the individual receives a custodial sentence (so during that period they cannot reoffend). (A recent US paper which compares an algorithm to bail decisions by judges suggests inputting outcomes for jailed defendants using outcomes of offenders with similar observables who the more lenient judge released: Kleinberg et al., 2017).

Advisory

21. The very nature of these tools means that an individual is being compared to the past behaviour of others with a similar profile. Using home address, for instance, has been criticised by Cathy O'Neil in her book ‘Weapons of Math Destruction’ for assuming that people in dangerous neighbourhoods are ‘birds of a feather’. As Alpaydin comments, ‘there are always other factors that affect the output; we cannot possibly record and take all of them as input, and all these other factors that we neglect introduce uncertainty.’ This is why Algo-care emphasises that, from a public law perspective, an algorithmic tool should not fetter a police officer’s discretion. The custody officer should, both in theory and in practice, be able to override the algorithm’s assessment where other factors suggest a different outcome. The overall decision-making process should ensure that those factors that led the human decision-making to override the algorithm are documented fully.

Regulatory oversight of algorithmic decision-making

22. The Algo-care framework recommends that ethical considerations, such as consideration of the public good and moral principles are factored into the deployment decision-making process.
Administrative arrangements such as an ethical review committee incorporating independent members could be established for such a purpose (such as Cleveland & Durham Joint External Ethics Committee or National Statistician’s Data Ethics Advisory Committee).

23. To keep in line with the likely expansion of the use of algorithmic tools within policing, it is also suggested that national guidance be developed by the College of Policing, overseen by HMIC.

April 2017
Written evidence submitted by Marion Oswald, Senior Fellow in Law and Director of the Centre for Information Rights, University of Winchester, and Sheena Urwin, Head of Criminal Justice, Durham Constabulary (ALG0030)

References

Ethem Alpaydin, ‘Machine Learning’ (MIT Press, 2016)


Doody [1993] UKHL 8

Jamie Grace and Marion Oswald, “‘Being on our radar does not necessarily mean being under our microscope’: The Regulation and Retention of Police Intelligence”, (2016) 22(1) EJoCLI
http://webjcli.org/article/view/441/621


Sarah Kramer, ‘An Algorithm is replacing bail hearings in New Jersey’ Motherboard 23 Feb 2017

https://www.pennlawreview.com/print/?id=553

Loomis, 2016 WI 68


Cathy O’Neil, ‘Weapons of Math Destruction’ (Allen Lane, 2016)


Marion Oswald and Jamie Grace, (2016) ‘Norman Stanley Fletcher and the case of the proprietary algorithmic risk assessment’ Policing Insight, available at http://repository.winchester.ac.uk/305/
http://www.predpol.com/how-predpol-works/
http://www.northpointeinc.com/
Algorithms in Policing – Take ALGO-CARE™
A proposed decision-making framework for the deployment of
algorithmic assessment tools in the policing context

Advisory

Is the assessment made by the algorithm used in an advisory capacity? Does a human officer retain decision-making discretion? What other decision-making by human officers will add objectivity to the decisions (partly) based on the algorithm?

On a case-by-case basis, what is the policing purpose justifying the use of algorithm, both its means and ends?

Is the potential interference with the privacy of individuals necessary and proportionate for legitimate policing purposes? In what way will the tool improve the current system and is this demonstrable? Are the data processed by the algorithm lawfully obtained, processed and retained, according to a genuine necessity with a rational connection to a policing aim? Is the operation of the tool compliant with national guidance?

Does the algorithm make suggestions at the right level of detail/ granularity, given the purpose of the algorithm and the nature of the data processed? Do the benefits outweigh any technological or data quality uncertainties or gaps? Is the provenance and quality of the data sufficiently sound? Consider how often the data should be refreshed. If the tool takes a precautionary approach towards false negatives, consider the justifications for this.

Who owns the algorithm and the data analysed? Does the force need rights to access, use and amend the source code and data analysed? How will the tool be maintained and updated? Are there any contractual or other restrictions which might limit accountability or evaluation? How is the operation of the tool kept secure?

What are the post-implementation oversight and audit mechanisms e.g. to identify any bias?

Where an algorithmic tool informs criminal justice disposals, how are individuals notified of its use (as appropriate in the context of the tool’s operation and purpose)?

Does the specification match the policing aims and decision policy? Can the stated accuracy of the algorithm be validated reasonably periodically? Can the percentage of false positives/negatives be justified? How was this method chosen as opposed to other available methods? What are the consequences of inaccurate forecasts?

Does this represent an acceptable risk (in terms of both likelihood and impact)? Is the algorithmic tool deployed by those with appropriate expertise?

Accuracy

Would the operation of the algorithm be considered fair? Is the use of the algorithm transparent (taking account of the context of its use), accountable and placed under review alongside other IT developments in policing?

Would it be considered to be for the public interest and ethical?

Responsible

Is appropriate information available about the decision-making rule(s) and the impact that each factor has on the final score or outcome (in a similar way to a gravity matrix)? Is the force able to access and deploy a data science expert to explain and justify the algorithmic tool (in a similar way to an expert forensic pathologist)?

Explainable

© Marion Oswald (James Irvine 2017)
Developed in collaboration with Durham Constabulary
Written evidence submitted by Marion Oswald, Senior Fellow in Law and Director of the Centre for Information Rights, University of Winchester, and Sheena Urwin, Head of Criminal Justice, Durham Constabulary (ALG0030)