BIG BROTHER WATCH

DEFENDING CIVIL LIBERTIES, PROTECTING PRIVACY

Big Brother Watch's written evidence on algorithms in the justice system for the Law Society's Technology and the Law Policy Commission

February 2019

Big Brother Watch is a cross–party, non-party, independent non-profit organisation leading the protection of privacy and civil liberties in the UK. We expose and challenge threats to people's privacy, freedoms and civil liberties at a time of enormous technological change in the UK.

We welcome the opportunity to give evidence to this important policy forum on technology and the law on the issue of algorithms in the justice system.

We find it deeply concerning that algorithms are already in use in the UK's justice system, in absence of a wider consideration by parliament, or indeed Government, as to the impact.

Harm Assessment Risk Tool (HART)

In this short submission, we focus on Durham Police's use of the 'Harm Assessment Risk Tool' (HART). HART is an artificially intelligent algorithmic tool used to make recidivism risk assessments about suspects and inform prosecution decisions. It was developed by Durham Police in conjunction with academics and has been in use since 2017.

The AI risk predictions guide decisions as to whether a suspect should be charged or released onto the 'Checkpoint' rehabilitation programme. Moderate risk' suspects are informed that if they successfully complete the Checkpoint programme they will not receive a criminal conviction

We focus on this tool because it is one of the most significant examples of algorithms in the justice system in the UK, and was the topic of an investigation by Big Brother Watch in 2018. We believe that Durham Police's creation and use of HART exemplifies many of the risks associated with rapid application of algorithms in the justice system: risks of discrimination, privacy intrusion, de facto automated decision making and profiling, as well as erosion of trust in law enforcement.

Professor Luciano Floridi, Director of the Digital Ethics Lab at the Oxford Internet Institute, co-authored a paper titled 'The ethics of algorithms: Mapping the debate'. In that paper, the authors identified a number of ethical concerns raised when decisions are delegated to algorithms including **inconclusive evidence leading to unjustified actions; inscrutable evidence leading to opacity; misguided evidence leading to bias; unfair outcomes leading to discrimination; and transformative effects leading to challenges for autonomy and informational privacy**.[1] The academics involved in the HART project acknowledged that, '*The implementation of the HART model raises every single one of these concerns to a greater or lesser extent'.*[2] These serious issues, examined through the lens of the law, touch on the prohibition of discrimination (Article 14, Human Rights Act); the right to a private life (Article 8,

1 Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society. https://doi.org/10.1177/2053951716679679

2 Oswald, M., Grace, J., Urwin, S., & Barnes, G. (2017). Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and 'Experimental' Proportionality. Information & Communications Technology Law: https://ssrn.com/abstract=3029345

Human Rights Act), and the right not to be subjected to purely automated decision making and profiling (Data Protection Act 2018, s.49).

The HART algorithm is based on a random forest model, constructed from 509 separate classification and regression decision trees (CART), which are combined into the forecasting model. HART was built on a dataset using approximately 104,000 custody events over a five year period. It uses 34 different predictor variables to arrive at a forecast, 29 of which focus on the individual's history of criminal behaviour. A further variable is the number of police intelligence reports relating to the individual. The other variables include age, gender and two types of residential postcode.

Big Brother Watch was disturbed to learn of the significance of one of the postcode variables used – the Mosaic code. Mosaic is a geodemographic segmentation tool sold by marketing company Experian.

Mosaic is built on 850 million pieces of data including family composition, children, family and personal names and ethnicity inferences, online data, occupation, welfare data, health data, GCSE results, gas and electricity consumption, census data and ratio of gardens to buildings. Mosaic profiles all 50 million adults in the UK to classify postcodes, households and even individuals into one of 66 stereotypes. Examples of the 66 categories are 'Disconnected Youth', 'Asian Heritage' and 'Dependent Greys'. Experian's Mosaic code includes the 'demographic characteristics' of each stereotype, characterising 'Asian Heritage' as 'extended families' living in 'inexpensive, close-packed Victorian terraces', adding that 'when people do have jobs, they are generally in low paid routine occupations in transport or food service'.

This tool raises novel questions about big data and privacy, the right to be free from profiling and automated decisions, algorithmic discrimination, and fairness in the criminal justice system – none of which have been addressed in the development of this tool. It is unacceptable that this tool, driven by profiling data, is being used by UK law enforcement systems to inform potentially life-changing criminal justice decisions. Allowing this kind of profiling data to be used risks producing unfair and inaccurate decisions and a 'postcode lottery' of justice, reinforcing existing biases and inequality.

Predictive policing and risk assessment systems

In fact, a number of UK police forces are investing in commercial software, or building their own systems, to predict crime.

In addition to undermining privacy and engaging a myriad of rights issues, the use of commercial machine-learning and 'black box' AI in the criminal justice system raises very serious accountability issues, as the decision-making processes cannot be understood or analysed. If an individual is subject to a decision, prediction or risk assessment, but cannot be told the reasons for the decision nor challenge it, there is an unacceptable accountability deficit.[3] In such a context, it is difficult to ensure the protection of individuals' rights and even their right to a fair trial.

PredPol

PredPol, is a geographic crime prediction tool that feeds crime and location information to a machine-learning algorithm to calculate predictions.[4] However, multiple studies have found that these systems can lead to areas being disproportionately over-policed, resulting in self-perpetuating feedback loops where predictions become self-affirming.[5] Similar systems have been or are currently being considered by Greater Manchester Police, West Midlands Police, Yorkshire Police and the Metropolitan Police. However, it is clear that these experimental predictive systems could have discriminatory impact.

National Data Analytics Solution (NDAS)

In addition, the new National Data Analytics Solution (NDAS), piloted by West Midlands Police but intended for all police forces to use from March 2019,[6] uses data about individuals taken from a number of public bodies to predict the risk of someone committing a crime in future, in order to pre-emptively intervene.

An independent review of the system said that there were "serious ethical issues" in particular in relation to inaccurate prediction and "the potential reversal of the presumption of innocence".[7] We share those concerns. It also raised questions around privacy rights and data protection, specifically the repurposing of data collected by public services for policing, the accuracy of the data, and people's ability to "meaningfully consent" to their data being used.

---

3 https://publications.parliament.uk/pa/cm201719/cmpublic/dataprotection/memo/dpb06.pdf

4 https://www.predpol.com/

5 Lyria Bennett Moses & Janet Chan (2018) Algorithmic prediction in policing: assumptions, evaluation, and accountability, Policing and Society, 28:7, 806-822, DOI: 10.1080/10439463.2016.1253695 ; Ensign et al, (2017) 'Runaway Feedback Loops in Predictive Policing', Cornell University Library, 29 June 2019 https://arxiv.org/abs/1706.0984)

6 https://www.newscientist.com/article/2186512-exclusive-uk-police-wants-ai-to-stop-violent-crime-before-it-happens/

7 https://www.turing.ac.uk/sites/default/files/2018-11/turing_idepp_ethics_advisory_report_to_wmp.pdf

Ineffective legal frameworks

The Data Protection Act 2018 contains broad exemptions for law enforcement purposes, and as such fails to sufficiently protect citizens' rights – including the right to be free from purely automated decisions

The GDPR safeguards individuals against significant decisions based solely on automated processing.[8] However, the UK's Data Protection Act 2018 makes exemptions from this important GDPR right. Section 14 of the Data Protection Act 2018 permits purely automated decisions with legal or similar significant effects to be made about a subject, in absence of the subject's consent – so long as the subject is notified that the decision was purely automated after the fact. The subject is then to be afforded just one month to request a new decision if they wish.

However, we are not aware of individuals being notified of purely automated decisions by police, despite the amount of automated processing in use.

This is likely because under section 14 of the Data Protection Act 2018, automated decisions that have significant legal or similar effects on a subject are not necessarily classified as "purely automated" if a human has administrative input. For example, if a human merely ticks to accept and thus enact a serious automated decision, the decision would not need to be classified as "purely automated" under law and as such, the minimal safeguards of notification and re-evaluation would not even apply.

Therefore, justice decisions could be being made that are for all intents and purposes automated decisions, without individuals being notified of this fact or of their right to appeal. We raised concerns about this during the passage of the (then) Data Protection Bill 2018, which were echoed by the Deputy Counsel to the Joint Committee on Human Rights who said, "There may be decisions taken with minimal human input that remain de facto determined by an automated process".[9]

The Data Protection Act 2018 in fact throws open the door for authorities to make justice decisions based on big data and automated processing – and weak legal definitions mean that the few safeguards there are may not even apply.

Big Brother Watch believes that two important amendments are required to the Data Protection Act 2018. First, decisions that engage individuals' human rights must never be purely automated decisions; second, automated decisions should be more clearly defined as those lacking *meaningful* human input.

Silkie Carlo

8GDPR, Article 22

9Note from Deputy Counsel, 'The Human Rights Implications of the Data Protection Bill', 6 December 2017 (https://www.parliament.uk/documents/joint-committees/human-rights/correspondence/2017-19/Note_Deputy_Counsel_DPBill.pdf)