BIG BROTHER WATCH
DEFENDING CIVIL LIBERTIES, PROTECTING PRIVACY

Science and Technology Inquiry into Algorithms in decision-making

Written evidence submitted by Big Brother Watch

October 2017

1. We have long warned against the adoption of Big Data analytics as the panacea to society's woes. We have raised concern about bias, assumptions and the potential for negative or restrictive conclusions being drawn, which could cause individuals harm. Algorithms, the tools used to analyse big data, are clearly part of this concern.

2. Firstly, not all data is good data. Rather than encouraging "big data" analytics, we should be encouraging "relevant data" and "proportionate data" analytics, especially in the context of those made by or carried out in relation to Government and public bodies.

3. Whilst some algorithms are designed to provide yes/no answers, increasingly that is not the sole purpose. Indeed as we have seen from the recent "Growing the Artificial Intelligence Industry in the UK"[1] paper published by the Professor Dame Wendy Hall and Jerome Pesenti, the push appears to be for algorithms to infiltrate and impact every corner of our data driven lives. We feel profoundly nervous of such proposals and stress that caution is applied to the optimism of a perfect world which the paper tends to err towards.

4. Firstly human beings are nuanced. There is no such thing as one size fits all in terms of people. All legislative decisions will have positive and negative outcomes depending on how the law impacts the individual. Whilst broad laws are necessary to help society to function, the role of using algorithms to offer precision to those laws is not necessarily going to provide better outcomes.

5. Making an appropriate decision about an individual requires an appreciation of nuance, opinion and personality. As most of us know an internet search rarely provides us with the perfect or indeed the right answer first time; we have to scroll and read a number of different links to a variety of webpages, analysing, reading and assessing the range of differing views and opinions, before we are able to draw conclusion or be equipped to ask further questions to determine a more complete picture.

6. Using algorithms therefore to answer questions which impact people's lives therefore will require nuance to address the failings of one size fits all. However pinning down nuance in data runs the risk of requiring, not just more data, but the potentially scrutinising the minutia of people's lives in order to improve outcomes, not just of the algorithms but of the individuals themselves. In a completely connected society that scrutiny comes in the form of real time monitoring and tracking of citizens. If such scrutiny were replicated in the physical world citizens would be in uproar.

7. In order to address the concern of intrusion, algorithms will predict behaviour. But with prediction comes the issue of bias or prediction based on inaccurate data which is lacking detail.

[1] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf

8. It is inevitable that if humans are building systems, human bias will, even unwittingly, find their way into those systems. Arguably an algorithm engineered by a majority white, middle class, highly educated, male workforce in Silicon Valley will be steered not just by the data but by the background and experience of those people. These are people whose backgrounds are far removed from many, if not most, of the people whose lives will be impacted by the algorithm or system.

9. Organisational bias can also be problematic with algorithms. Algorithms built to address policing issues must be completely clear of bias and be 100% accurate. Without such guarantees, the risk of arresting the wrong person, holding someone in custody or determining areas of crime hotspots are just two obvious negative consequences.

10. Obviously such guarantees are impossible. The consequences can already be seen. In May of this year, Durham Police began to use an algorithm called HART to determine whether a suspect should be kept in custody or released on bail. HART bases its decision on the predicted likelihood of a suspect re-offending if released. The algorithm uses data from five years of offending histories, as well as a suspect's postcode and gender, to classify suspects as either low, medium, or high-risk of offending. The force stated that the algorithm was 88% correct in relation to its high-risk predictions, and 98% accurate in relation to low-risk predictions.[2]

11. Whilst such statistics sound appealing, if only 88% are correct, 12% of cases are therefore incorrect. 12% is clearly not an insignificant figure, especially in the context of someone being deprived of their civil liberty.

12. The algorithm was criticised by RUSI in its report on Big Data and Policing, on the basis that because the system was only using Durham Police's data, offences committed in other areas would not be considered, and dangerous criminals might not be identified.[3]

13. A further example of algorithmic technology which to date provides far from accurate responses is the recent trial of facial biometric recognition by the Metropolitan Police at Notting Hill Carnival. The trial was the second one undertaken at the Carnival. It was deemed to be a success on the basis that one person had been correctly matched over the two days of use. However, when we were invited to watch the trail on its second day, we witnessed 35 people be wrongly identified by the technology. One specific example of a false positive we witnessed was of a female with long hair being picked out of the crowd as a match to an algorithm of a bald male. Without human intervention the issue of false positives and inaccuracies could have been potentially very damaging.

14. These technologies are being rolled out by law enforcement without any debate in Parliament. Any such algorithmic or automated decision-making must be subject to regulation, legislation and oversight.

15. Additionally, we must also be aware of placing too much trust in algorithms to carry out tasks on our behalf. To paraphrase Franklin Foer: when we outsource thinking to algorithms, we are really outsourcing our thinking to the organisations that run the algorithms.[4]

16. Evidence of how unquestioned trust can impinge negatively on society is evident in the use of algorithms to make stock market decisions. In 2016 high-frequency trading algorithms were blamed

---

[2] http://www.bbc.co.uk/news/technology-39857645

[3] RUSI (2016), Big Data and Policing An Assessment of Law Enforcement Requirements, Expectations and Priorities, September 2017 https://rusi.org/sites/default/files/201709_rusi_big_data_and_policing_babuta_web.pdf

[4] Foer, F (2017), "World Without Mind", Penguin: UK, pg 72

for a *"flash crash"* of Sterling.[5][6] This resulted from algorithms assessing the reaction to Brexit in the mass media and placing value on negative headlines based on opinion, as opposed to news based on quantifiable fact. Whilst the use of "black box" algorithms in trading are long-standing, the move towards assessing data based on social media activity, rather than objective considerations, is a worrying trend, particularly with the influence of "fake news" or sponsored tweets by rogue states seeking to disrupt through influence. This is a further example of poor data resulting in negative algorithmic decisions.

17. We acknowledge that the process of learning what data is relevant in a particular circumstance often comes from trial and error, but when it comes to building algorithms for use by Government or public bodies, the luxury of trial and error does not exist: the decisions being made impact directly on people's lives.

18. As life is now lived on and off-line concurrently, discussion must be had to determine if the same rules of law, morality and ethics which we live by in the physical world can be embedded into the online world with the same success. This is a complex point because, as we can see across a wide range of technologies, the rules are rarely as clear-cut. The same rules that apply in the physical world may not work in the digital world.

19. Transparency and accountability are two different things. Transparency of an algorithm will relate to how the algorithm works, whilst accountability will relate to the ethics and rule of law.

20. With regards to transparency, the complexity of algorithms and the data they use mean that the inputs, processes and outputs can all be inaccessible to the citizens whose data is being used.

21. In some cases, even the engineers who create an algorithm do not know its inner-workings, as the evolution of the algorithm can move quickly. Certain algorithms cannot be transparent by their very nature. This is due to 'black-boxing', a process whereby you can only see the input or output of a certain type of algorithm, and not its inner workings. The inner workings of such algorithms (known as neural networks) are constantly changing as it learns based on the data it is given. This leaves the algorithm almost impossible to audit, challenge, and hold accountable for the decision made.

22. With regards to accountability, currently, we see the Government ask business to address these types of complexities and find solutions – we have seen this recently in relation to the issues of extremist content online, fake news and child sex exploitation in the Internet Safety Strategy – Green Paper[7].

23. It may seem logical on the surface that the UK Government seeks the guidance and technical abilities of the big tech companies to assist with finding solutions to these problems. However, is it right that such processes, which could impact the laws of the country and the rights and freedoms of citizens, be determined by big business? Such powers should surely be determined by Parliament and be subject to independent scrutiny. Such proposals raise very serious questions regarding accountability.

24. Consideration must also be given to establish exactly what accountability of an algorithm will mean. After all who owns an algorithm? Is it the individual engineer or the organisation? It may sound far-

[5] https://www.technologyreview.com/s/602586/algorithms-probably-caused-a-flash-crash-of-the-british-pound/

[6] https://www.theguardian.com/business/2016/oct/07/what-caused-pound-flash-crash-brexit-fallen-sterling

[7] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/650949/Internet_Safety_Strategy_green_paper.pdf

fetched but should an algorithm be taken to court? Can an algorithm be sued if the consequence of it has led to a harmful impact on an individual or individuals?

25. If all algorithms were open source it would enable greater scrutiny and transparency. Whilst we appreciate business will oppose unique selling points to be exposed, shared or used by competitors, the right within the General Data Protection Regulation (GDPR) for a data subject to ask for human intervention when it comes to an automated decision or profiling poses a question mark over how this can be done if access to how decisions are made is not transparent. Without the ability for the algorithm behind the decision to be scrutinised, the data subject, data controller and data processor may all be left in the dark, leading to an unacceptable accountability deficit.

26. The intention to use algorithms as a way of policing the internet must not lead to a way of restricting citizens' human rights, such as the right to a private life or the right to freedom of thought, conscience and religion.

27. Already Instagram are using algorithms to read comments and determine via patterns and association if the words used are abusive.[8] If the algorithm decides the comment is offensive the comment isn't posted, but the individual who typed it doesn't know that the comment was blocked. Whilst it is important to protect people online, caution is needed to protect against machine-learning algorithms behaving as online thought police.

28. Furthermore we are concerned by the negative language being used when discussing protection and privacy of data. In the 'Growing the Artificial Intelligence Industry in the UK' [9] paper. Privacy and protection of data are seen to be *"barriers to overcome"*. Privacy and protection of data should not be viewed as a barrier, nor as a restriction of opportunity, inhibitor of innovation or any other negative consequence which they are so often billed as. We must begin to accept that in a connected society, protection of data through privacy and security is absolutely fundamental to the security of the citizen and the security of society as a whole. With cybercrime on the rise from criminal individuals, nation states and terror organisations, the ability to limit vast databases of sensitive data is not a barrier, but a fundamentally sensible approach to the role of security any government should seek to establish for its citizens.

29. We are pleased to see the new Data Protection Bill apply the offence of re-identifying de-identified personal data.

30. In relation to the protection of individuals' personal data, the issue of re-identification or de-anonymisation is significant. The analytical capabilities of algorithms in the 'big data' environment have the potential to completely undermine formed notions of privacy, especially in the context of 'anonymised' data. We have consistently raised concern about the promises of anonymisation as a panacea.

31. Whilst we appreciate that big data analysis of health data, statistical data and historical research for example can provide fascinating, evolutionary and revolutionary solutions, we believe there should be the option of people to opt out of having their personal data analysed in such a way – especially in the context of such data being proven to be re-identified.

---

[8] https://www.theverge.com/2017/6/29/15892802/instagram-ai-offensive-comment-filter-launches

[9] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf

32. This would not disrupt a data set but would offer the right to a vulnerable member of society whose participation in such research could – if they were re-identified, even accidentally – cause considerable harm or distress.

33. We welcome the right of a citizen to challenge automated decision making and profiling in the GDPR. We note the same rights are applied in the Data Protection Bill in relation to law enforcement processing and intelligence services processing but the number of exemptions then applied are likely to lead to very little opportunity for a citizen to be notified that they have been a subject of profiling or automated decision making and therefore rarely, if ever, get the opportunity to raise a challenge.

34. Whilst we are pleased that the GDPR is to be enacted into UK law pre-Brexit we remain concerned about future data protection for UK citizens post Brexit. Particularly as the GDPR will be subject to the proposed EU (Withdrawal) Bill, leaving it open to being removed from UK law, amended or diminished.

35. We caution very firmly against the removal or diminishment of the GDPR in a post-Brexit UK. As the Information Commissioner has said, the GDPR is the *"gold standard"* of data protection regulation and enforcement and one we should maintain.[10]

About Big Brother Watch

Big Brother Watch is a civil liberties and privacy campaign group that was founded in 2009. We have produced unique research exposing the erosion of civil liberties in the UK, looking at the dramatic expansion of surveillance powers, the growth of the database state and the misuse of personal information.

[10] Elizabeth Denham, House of Lords EU Home Affairs Sub-Committee, Wednesday 8 March 2017 http://www.parliamentlive.tv/Event/Index/125e1463-62ed-41bb-ab64-811d0f94bfee