27, Dingley Place
London
EC1V 8BR
www.futureadvocacy.org
@FutureAdvocacy

Submission to Science and Technology Committee's inquiry on 'Algorithms in decision-making' - April 2017

Matthew Fenech, Olly Buston

A: Summary:
Algorithmic transparency is not only ethically desirable, but also commercially advantageous. There are various global examples of positive social outcomes that followed from algorithmically-driven decision-making. The development of a global directory of 'best practice' in the application of algorithmic decision-making to public service delivery would aid the implementation of this best practice in the UK.

B: Recommendations:

- **The UK Government should seek out examples of good practice in the use of algorithms in decision-making, and apply the lessons learned from these use cases to the UK situation.**
- **The UK Government should support the development of a global directory of (i) 'best practice' in policies around algorithmic decision-making to public service delivery and (ii) examples of successful applications.**
- **The UK Government should promote transparency and accountability in AI decision-making by supporting research that facilitates an opening of the 'black box' of intelligent algorithms and supporting open data initiatives.**
- **The UK Government should adhere to the commitment laid out in the UK Digital Strategy to implement the General Data Protection Regulation by May 2018, regardless of the outcome of Brexit negotiations.**

C: Introduction:
Future Advocacy is a think tank focused on making sure the United Kingdom is best positioned to capitalize on the opportunities and mitigate the risks presented by artificial intelligence, big data, and similar technologies driving the Fourth Industrial Revolution. We are optimistic about these technologies and the tremendous opportunities they present, both in terms of economic growth and in improving social and personal well being. In tandem, we are aware of the inevitable socio-economic challenges these 'disruptive' technologies provide, and we advocate for policies that will minimise any negative impacts.

Over the last year, we've contributed to the Science and Technology Committee's inquiry on Robotics and AI; participated in discussions at United Nations level; briefed Downing Street staff; written a scoping paper for Sir Tim Berners-Lee's Web Foundation on 'Maximising the

FUTURE ADVOCACY

27, Dingley Place
London
EC1V 8BR
www.futureadvocacy.org
@FutureAdvocacy

Opportunities and Minimising the Risks of Artificial Intelligence in Low and Middle Income Countries', and released a report called 'An Intelligent Future?' (available at futureadvocacy.com/s/An-intelligent-future-3.pdf), which makes 12 policy recommendations to the Government. Since then, we've established a global network of partners in industry and academia to begin to action these proposals.

We welcome the Commons Science and Technology Select Committee's inquiry into the use of algorithms in public and business decision-making. In this document, we outline our position on the points identified by the Select Committee that are relevant to our remit. **We are happy to be contacted** if you have any questions about our response (on the contact details below), and **we are happy for our response to be published in full**.

D: Responses to Consultation Questions:

The extent of current and future use of algorithms in decision-making in Government and public bodies, businesses and others, and the corresponding risks and opportunities

As part of our work for the Web Foundation on 'Artificial Intelligence in Low and Middle Income Countries' we identified many examples of good use of algorithms to guide public and private decision-making worldwide. For example:

- In 2015, Colombia conducted its first agricultural census for 45 years. 43,000 interviews were carried out to obtain data on various metrics such as yield size, crop type and land use - **the vast resulting dataset was then analysed algorithmically to identify patterns and to distil essential points for policy makers to act on**. The major positive outcome of this project was that 3 million hectares of land that were underutilised or had been seized by the state as a result of criminal activity were redistributed to small landowners and landless farmers. The fairer ownership of land that resulted from this intervention contributed to improved relations between the Government and the FARC rebel group.[1]
- In Togo, the Red Cross/Red Crescent Climate Centre runs the FUNES project aimed at reducing the impact of flooding caused by overspilling of the Nangbéto Dam. This is a machine-learning algorithm embedded in a digital tool that supports Forecast-Based Financing.[2] In the past, models were poor at predicting the likelihood of overspill, due to a lack of good quality data (meteorological information, for example, was only available on a monthly basis). Using a combination of crowdsourced information (including by mobile phone) and AI techniques, **an improved model of overspill prediction has been**

[1] Albertus, M. and Kaplan, O. (2015), 'The Key To Peace In Colombia Is In Its Countryside', Forbes, available at https://www.forbes.com/sites/realspin/2015/10/29/the-key-to-peace-in-colombia-is-in-its-countryside/#de89ae966deb

[2] Coughlan de Perez, E., van den Hurk, B., van Aalst, M., Jongman, B., Klose, T. and Suarez, P. (2014) 'Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts', Nat. Hazards Earth Syst. Sci. Discuss., 2, 3193–3218

27, Dingley Place
London
EC1V 8BR
www.futureadvocacy.org
@FutureAdvocacy

developed which can function in spite of imperfect data, allowing for better, algorithmically-driven decisions on resource distribution in advance of flooding.[3,4]

- The United States Defense Advanced Research Projects Agency (DARPA) launched the Memex programme in 2015, with the objective of building tools that can collect and analyse content ignored by, or unavailable to, commercial search engines (such as sites on the 'Deep' or 'Dark Web'), and build models to predict behaviour. By using machine learning to analyse pricing data of online adverts for paid sexual services, for example, they have improved the identification of underage sex workers under the control of organised trafficking rings.[5] Similarly, a collaboration between researchers at Leuven Catholic University, the Netherlands, the National Research University's Higher School of Economics, Russia, and the Amsterdam police force has developed software that automates the processing and analysis of police reports, learning to look for indicators that the crime in question is in some way linked to human trafficking.[6] **In both these cases, algorithmic tools were used to deploy limited police resources more effectively**.

However, there are also examples of algorithmic decision-making that had a negative impact. Perhaps the most well-known is that of recidivism software widely used by American courts to assess the likelihood of an individual re-offending. This was found to be two times less likely to falsely flag white people and two times more likely to falsely flag black people as likely to reoffend.[7]

Therefore, there is much for the the UK Government to learn from global examples of the implementation of algorithmic decision-making in public services. At Future Advocacy, **we are scoping support for the development of a global directory of 'best practice' in policies around algorithmic decision-making in public service delivery, and examples of successful applications.** Such a resource would provide data on i) the use cases/scenarios for which algorithms were used to guide decision-making, ii) the algorithm used and its design, iii) the institution(s) responsible for the algorithm and its implementation, and iv) the outcomes of its use. We are confident that the development of such a resource will be immensely helpful to public and private bodies seeking to learn from the experiences of their counterparts in other countries and implement best practice in the UK.

---

[3] Salvarez C., (2017), 'How to smartly utilize the window between forecast and hazard', RCRC Climate Centre, available at https://media.ifrc.org/innovation/2017/02/01/forecast-based-financing-an-amazing-initiative-of-the-rcrc-climate-centre/

[4] IHC Communications Team (2017) 'How can the hydropower sector and the Red Cross work together to protect flood-prone communities?', IHC, available at https://www.hydropower.org/blog/how-can-the-hydropower-sector-and-the-red-cross-work-together-to-protect-flood-prone

[5] Couch, C. (2016) "How Artificial Intelligence Can Stop Sex Trafficking" NovaNext. Available at: http://www.pbs.org/wgbh/nova/next/tech/sex-trafficking/ (accessed on 29 March, 2017)

[6] "Artificial Intelligence Helps Combat Human Trafficking" (2014), press release by National Research University Higher School of Economics, Russia. Available from https://iq.hse.ru/en/news/177666531.html (accessed 29 March, 2017)

[7] Angwin, J., Larson, J., Mattu, S., and Kirchner, L. (2016) 'Machine Bias', Pro Publica, available at https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

27, Dingley Place
London
EC1V 8BR
www.futureadvocacy.org
@FutureAdvocacy

Whether 'good practice' in algorithmic decision-making can be identified and spread, including in terms of:

— The scope for algorithmic decision-making to eliminate, introduce or amplify biases or discrimination, and how any such bias can be detected and overcome;

— Whether and how algorithmic decision-making can be conducted in a 'transparent' or 'accountable' way, and the scope for decisions made by an algorithm to be fully understood and challenged;

— The implications of increased transparency in terms of copyright and commercial sensitivity, and protection of an individual's data

The lack of transparency surrounding modern machine learning algorithms, particularly neural networks, has often been alluded to using the term 'black box'.[8] Such decision-making systems are often deployed as a background process, unknown and unseen by those they impact. The use of this technology in this way raises significant and justifiable concerns; for example, the recidivism algorithm referred to earlier was protected under Intellectual Property Laws and was not open to scrutiny.[9] Furthermore, significant stores of data that are used to train and run these algorithms are not in the public domain, meaning it is impossible to test or challenge results, and these data may contain significant bias, reflecting the situation they were collected in. For example, if data on job applications was gathered from an industry that systematically hired men over women and this data was then used to help select likely strong candidates in the future, this could then reinforce sexism in hiring decisions. Indeed, Google ads promising help getting jobs paying more than $200,000 were shown to significantly fewer women than men.[10] This bias can reinforce prejudice in an insidious way that is often hard to detect. As Kate Crawford puts it, 'histories of discrimination can live on in digital platforms, and if they go unquestioned, they become part of the logic of everyday algorithmic systems'.[11]

It is encouraging, however, that the computer/data science research community is investing time and money into research on improved algorithmic transparency. The correct approach to the use of algorithms to take important decisions that have significant social impact may involve a range of different machine learning algorithms, from the 'good old fashioned AI' using logical/rule-based systems, to statistical techniques that have received a lot of attention recently, such as deep learning. Research should focus on developing a diverse range of solutions, and making existing techniques more transparent, which is becoming increasingly

[8] Castelvecchi, D. (2016) 'Can we open the black box of AI?' Nature News, available at http://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731

[9] Angwin, J., Larson, J., Mattu, S., and Kirchner, L. (2016) 'Machine Bias', Pro Publica, available at https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[10] Spice, B., (2015) "Questioning the Fairness of Targeting Ads Online." Carnegie Mellon University, available at http://cmu.edu (accessed 10 March, 2017)

[11] Crawford, K. (2016), "Artificial Intelligence's White Guy Problem", The New York Times, available from https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html (accessed 21 March, 2017)

27, Dingley Place
London
EC1V 8BR
www.futureadvocacy.org
@FutureAdvocacy

technically possible.[12,13,14] Moreover, in our research around the area, we have uncovered various situations where increased algorithmic transparency was not only ethically desirable, but a commercial necessity. We were struck, for example, by the testimony of Ed Bishop, Co-founder and CTO of CheckRecipient, an algorithmic tool that aims to reduce sensitive information being misdirected by email, at a recent conference held at Imperial Business School. His company has realised that the relationship with their clients has improved as a result of more information being provided to them about the reasons for the algorithmic decisions taken by the software. We have therefore **called for the UK Government to promote transparency and accountability in AI decision-making by supporting research that facilitates an opening of the 'black box' of intelligent algorithms and supporting open data initiatives**.[15]

Methods for providing regulatory oversight of algorithmic decision-making, such as the rights described in the EU General Data Protection Regulation 2016.

We agree with the principle set out in Article 22(1) of the General Data Protection Regulation (Regulation (EU) 2016/679, GDPR), which was adopted by the European Union (EU) in April 2016 and to be implemented in all EU member states by 2018. This Article states "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her".[16] Thus, European citizens now have the right to question and fight decisions that affect them that have been made on a purely algorithmic basis. We welcome the commitment by the Government, laid out in its UK Digital Strategy published on 1st March 2017, to "implement the General Data Protection Regulation by May 2018', thus ensuring "a shared and higher standard of protection for consumers and their data across Europe and beyond." **We recommend that this commitment is adhered to regardless of the outcomes of Brexit negotiations that will take place over the next two years.**

Contact:
Name: Matthew Fenech
Email: mattfenech@futureadvocacy.org
Phone: 07828 798522

[12] Castelvecchi, D. (2016) 'Can we open the black box of AI?' Nature News, available at http://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731
[13] Knight, W. (2017), 'The Dark Secret at the Heart of AI', MIT Technology Review, available at https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/
[14] Hardesty, L. (2016), 'Making computers explain themselves', MIT News, available at http://news.mit.edu/2016/making-computers-explain-themselves-machine-learning-1028
[15] Future Advocacy (2016) 'An Intelligent Future? Maximising the opportunities and minimising the risks of artificial intelligence in the UK', available at futureadvocacy.com/s/An-intelligent-future-3.pdf
[16] Regulation (EU) 2016/679 of the European Parliament and of the Council (2016) OJ L119/46