CDEI proposes a roadmap to tackle algorithmic bias

A major review from the Centre for Data Ethics and Innovation (CDEI), which draws on a detailed analysis of the use of algorithms in four sectors (financial services, local government, policing and recruitment), proposes significant measures for government, regulators and industry to act on to tackle the risks of algorithmic bias.

Key recommendations include:

- Government should place a mandatory transparency obligation on all public sector organisations using algorithms that have an impact on significant decisions affecting individuals.

*Organisations should be actively using data to identify and mitigate bias. They should make sure that they understand the capabilities and limitations of algorithmic tools, and carefully consider how they will ensure fair treatment of individuals.

- Government should issue guidance that clarifies the application of the Equality Act to algorithmic decision-making. This should include guidance on the collection of data to measure bias, as well as the lawfulness of bias mitigation techniques (some of which risk introducing positive discrimination, which is illegal under the Equality Act).

The CDEI, the UK government's advisory body on the responsible use of AI and data-driven technology, has proposed a roadmap that government, regulators and industry can take to increase fairness and reduce bias, while also ensuring that the UK regulatory ecosystem is set up to support responsible innovation.

The measures are designed to produce a step change in the behaviour of all organisations making life-changing decisions on the basis of data, with a focus on improving accountability and transparency. The report emphasises that organisations are responsible for their decisions — whether they have been made by an algorithm or a team of humans — and offers guidance for organisational leaders and boards to enhance accountability.

The CDEI recommends that the government should place a transparency obligation on all public sector organisations using algorithms which support significant decisions. This would include information about how algorithms are used in the overall decision-making process, and steps taken to ensure fair treatment of individuals.

The CDEI argues that there is an opportunity here: effective use of data can enable organisations to shine a light on practices that may otherwise go unseen, and identify the drivers of bias. The report recommends that government and regulators provide clear guidance about how organisations can

actively use data to tackle current and historic bias. This guidance should address the misconception that data protection law prevents the collection or usage of data for monitoring or addressing discrimination.

The review was informed by public engagement to gain a deeper understanding of attitudes towards algorithmic decision-making, drawing on methodologies including survey and behavioural science research. A large scale survey, conducted with Deltapoll, found that the majority of respondents were aware of the use of algorithms to support decision-making (around 6 out of 10). Of those respondents who were aware of the use of algorithms, respondents were most aware of their use in financial services (more than 5 in 10), in contrast to local government (around 3 in 10). The results suggest that the public are more concerned that the outcome of decision-making is fair, rather than whether algorithms are used to inform these judgements. There is public support for data — including age (net agreement of +59%), ethnicity (+59%) and sex (+39%) — to be used for tackling algorithmic bias in recruitment.

The review points to the need for an ecosystem of industry standards and professional services to help organisations address algorithmic bias in the UK and beyond. This presents an opportunity for the UK: leadership in this area can not only ensure fairness for British citizens, but can also unlock growth by incubating new industries in responsible technology.

To catalyse this, the CDEI has initiated a programme of work on AI assurance, in which it will identify what is needed to develop a strong AI accountability ecosystem in the UK. Other related CDEI work includes: working with the Government Digital Service (GDS) to pilot an approach to algorithmic transparency; supporting a police force and a local authority to apply lessons learnt and develop practical governance structures; and active public engagement to build understanding of the values that citizens want reflected in new models of data governance.

In its report, the CDEI calls for national leadership and coordination to drive the pace of change, and urges the government to be clear on where responsibilities sit for tracking progress.

Adrian Weller, Board Member for the Centre for Data Ethics and Innovation, said:

"It is vital that we work hard now to get this right as adoption of algorithmic decision-making increases. Government, regulators and industry need to work together with interdisciplinary experts, stakeholders and the public to ensure that algorithms are used to promote fairness, not undermine it. The Centre for Data Ethics and Innovation has today set out a range of measures to help the UK to achieve this, with a focus on enhancing transparency and accountability in decision-making processes that have a significant impact on individuals. Not only does the report propose a roadmap to tackle the risks, but it highlights the opportunity that good use of data presents to address historical unfairness and avoid new biases in key areas of life."

Simon McDougall, Deputy Commissioner — Regulatory Innovation and Technology

for the Information Commissioner's Office, said:

"We welcome and support the findings of the report and want to echo the Centre for Data Ethics and Innovation's concerns around algorithmic bias. When developed and used responsibly, algorithms can transform society for the better. But there is also significant risk that algorithms can exacerbate issues of fairness and inequality. This often impacts the most vulnerable or marginalised people. Data protection law requires fair and transparent uses of data in algorithms, gives people rights in relation to automated decision-making, and demands that the outcome from the use of algorithms does not result in unfair or discriminatory impacts. The ICO has prioritised the data protection implications of AI for a long time and has produced guidance for organisations on the use of AI."