Review into Bias in Algorithmic Decision-Making

12th March 2021

Lara Macdonald, Senior Policy Advisor, CDEI
Ghazi Ahamat, Senior Policy Advisor, CDEI
Dr Adrian Weller, CDEI Board Member and Programme Director for AI at the Alan Turing Institute
Government’s role in reducing algorithmic bias

- As a **major user** of technology, government and the public sector should **set standards, provide guidance and highlight good practice**

- As a **regulator**, government needs to **adapt existing regulatory frameworks** to incentivise ethical innovation
Bias, discrimination & fairness

- We interpreted bias in algorithmic decision-making as: the use of algorithms which can cause a systematic skew in decision-making that results in unfair outcomes.

- Some forms of bias constitute discrimination under UK equality (anti-discrimination) law, namely when bias leads to unfair treatment based on certain protected characteristics.

- There are also other kinds of algorithmic bias that are non-discriminatory, but still lead to unfair outcomes.

- Fairness is about much more than the absence of bias

- There are multiple (incompatible) concepts of fairness
How should organisations address algorithmic bias?

**Guidance to organisation leaders and boards**

- Understand the **capabilities and limits** of algorithmic tools
- Consider carefully whether individuals will be **fairly treated by the decision-making process** that the tool forms part of
- Make a **conscious decision on appropriate levels of human involvement** in the decision-making process
- Put structures in place to **gather data and monitor outcomes** for fairness
- Understand your legal obligations and carry out **appropriate impact assessments**

**Achieving this in practice will include:**

- Building (diverse, multidisciplinary) internal capacity
- Understand risks of bias by measurement and stakeholder engagement
- Creating organisational accountability and transparency around fair decisions
Policing

How are algorithms used?

Algorithms are built on historical data to derive insights, prioritise resources, and assess risks associated with individuals for example:

- Predictive mapping
- Individual risk assessment
- Data scoring tools
- Other

Advice to police forces

Conduct an integrated impact assessment before adopting a new data analytics software

Classify the output of statistical algorithms as a form of police intelligence

Ensure that they have appropriate rights of access to algorithmic software

Recommendation

The UK Government’s Home Office (justice department) should define clear roles and responsibilities for national policing bodies with regards to data analytics
Recommendation: **Government** should place a mandatory transparency obligation on all public sector organisations using algorithms that have a significant influence on significant decisions affecting individuals.

**Why this recommendation?**

- (Perception of) lack of transparency is a problem for trust
- Transparency drives quality
- Democratic accountability for public sector decisions
- We already publish transparency information on non-algorithmic decision-making processes (e.g. caseworker manuals)
- An opportunity for governments to use its buying power to influence the wider market
- Hard for individual teams/projects to be transparent on their own
Government’s role in reducing algorithmic bias

- As a **major user** of technology, government and the public sector should set standards, provide guidance and highlight good practice.

- As a **regulator**, government needs to **adapt existing regulatory frameworks** to incentivise ethical innovation.
# The Regulatory Framework

## I. Cross-cutting principles of fairness set out in equality and data protection law.

- **UK equality law regulator (& other regulators under their Public Sector Equality Duty):**

- **UK data protection regulator:**
  - *Data Protection Act 2018, GDPR*

## I. Sector regulators establish and enforce standards of fairness in particular sectors.

<table>
<thead>
<tr>
<th>Private Sector</th>
<th>Public Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UK financial conduct regulator:</strong> <em>Principles for fair treatment</em></td>
<td><strong>UK policing and fire &amp; rescue inspectorate</strong></td>
</tr>
<tr>
<td><strong>UK competition regulator:</strong> <em>Consumer Rights Act</em></td>
<td><strong>UK education inspectorate</strong></td>
</tr>
<tr>
<td>Others...</td>
<td>Others...</td>
</tr>
</tbody>
</table>
Recruitment

**SOURCING**
- Job description review software
- Targeted advertising
- Recruiting chatbots
- Headhunting software

**SCREENING**
- Qualification screening tools
- CV matching
- Psychometric tests and games
- Ranking algorithms

**INTERVIEW**
- Voice and face recognition in video interviewing

**SELECTION**
- Background check software
- Offer predicting software
Data is needed to monitor outcomes and identify bias, but data on protected characteristics is not available often enough.

Why? Various reasons, including concerns that:

- Collecting protected characteristic data is not permitted or hard to justify in data protection law.
- Customers will not want to share the data.
- Data could provide an evidence base that organisational outcomes were biased.

We recommended a push to incentivise greater usage of such data, and are now doing some follow-up work on opportunities for doing this in more innovative ways than repeated collection.
Next steps: Developing an AI Assurance Ecosystem

Assurance covers a number of governance mechanisms for third parties (which in the case of AI systems could even be the developer) to develop trust in the compliance and risk of a system or organisation.

### Compliance assurance

| 1) Formal verification       |
| 2) Audit (as commonly used in business) |
| 3) Certification            |
| 4) Accreditation            |

### Risk assurance

| 1) Impact assessment     |
| 2) Audit (as often used in ML fairness) |
| 3) Impact evaluation     |
| 4) Ongoing testing (incl. red teaming) |

CDEI is actively looking at what it will take to build a mature ecosystem of assurance services that enable the trustworthy deployment of algorithmic decision-making.
Framework for Trustworthy Deployment *in context*

**IDENTIFY requirements**
Engage with users, practitioners and stakeholders: What do people need? What are their concerns?

**BUILD in requirements**
Theoretical and technical foundations to build in requirements into AI systems

**CHECK requirements**
Enforceable governance which is legally sound and technically feasible