An Accountability for Algorithms Act

Executive Briefing

Executive Briefing

Algorithms are increasingly used by governments and businesses to make 'automated', data-driven decisions which can have far-reaching consequences with little transparency, scrutiny or accountability. Although algorithms' superficial appearance of objectivity appears to remove biases in human decision-making, algorithms always reflect the assumptions of those who designed them and the patterns in the data on which they are trained.

Our in-depth analysis of 'automated' decisions at work has revealed clear gaps in legal protection and existing mechanisms for accountability to workers.

Without intervention and oversight, the natural state of data-driven technologies is to replicate past patterns of structural inequality that are encoded in data, and project them into the future of work. It is vital that this is understood by policymakers, and by all who use algorithms to make decisions which can affect people's lives in multiple ways, from educational attainment and hiring, to pay and promotion.

In this report we investigate the current state of play in the use of the algorithms and high-tech software that shape work and our lives in ways we rarely notice. These algorithms are usually created, in the first instance, by developers who design computer models, deciding what data should be used to make predictions about people, and how it should be analysed, without considering the consequences for large groups of people of the decisions made or supported by these tools.

Refined by artificial intelligence (AI) and machine learning, algorithms can become so complex that they are opaque to their own designers and users, making discriminatory and unfair outcomes difficult to identify and correct. This inaccessibility through complexity may go alongside an intentional lack of transparency from firms attempting to protect commercially sensitive material by refusing to disclose details about their software and code. And these problems are compounded by the fact that the public and private sector organisations which deploy them too often do so without thinking critically about what these algorithms are, what they can and cannot do, and what the unintended consequences might be.

This failure is, in part, due to a legal framework and regulatory regime which has not kept pace with recent developments in what technology is capable of, or with how it is routinely deployed. In particular, existing equality and data protection legislation is insufficient to provide protection and redress for those disadvantaged at work by assumptions baked into algorithms. We make a series of recommendations, based around proposed new legislation, to ensure that algorithms are used in a fair and transparent way, and that people are properly accountable for their decisions about their design and use.

Our focus is work but our analysis and recommendations may inform a wider debate about AI regulation.

Part 1

The myth of neutrality

Part 1 sets out what algorithms are and how they are used in contemporary automated, data-driven decision-making systems. It discusses some of their advantages, but also some of their weaknesses and limitations, and the ways in which these can be hidden by an assumption that algorithms are inherently neutral and objective.

In fact, human agency is still at the root of the way algorithms function: human beings select the input data, design the rules by which the data is processed and rely on the outputs, even if the algorithm itself is refined by AI and machine learning without direct additional human input. And these outputs eventually shape the experiences and outcomes for those subject to algorithmic decisions: for example, when access to work and the full range of employer functions are digitised, the results end up being felt by workers, and indeed unsuccessful job applicants.

We identify seven critical choices in the design and deployment of data-driven decision-support technologies. Forms of bias, and existing inequality in the data being relied upon, can skew decision-making at any of these critical points. They are all therefore choices for which technology regulation must hold organisations which use algorithms accountable.

Part 2

Structuring accountability

Part 2 examines how the operation, language and culture of data-driven technologies make decision-making both more diffuse and less transparent. Relatively few people understand the terminology or internal workings of AI, machine learning and mass data processing. This means that human agency is often obscured. It also means that those who make and control the technology, and the data that feeds it, exercise the real power. They take decisions that will profoundly affect the lives of others on a daily basis – sometimes without realising it. But the 'invisible' nature of what they do makes it more difficult to hold them accountable, especially at work where stakes are high and information asymmetries are pronounced.

We argue that they should be accountable. Those at the heart of the data economy should consciously examine the adverse impacts of their work, especially equality impacts. That is both the burden and the privilege that comes with exercising agency on behalf of millions of other individuals. Meaningful accountability will help data-driven technologies serve the public interest, rather than becoming vehicles that reinforce unconscious biases and entrench inequality. Human agency must be affirmed, not removed.

Part 3 and 4

The regulatory ecosystem and key challenges and gaps

Parts 3 and 4 analyse the existing legal and regulatory frameworks that are relevant in this area. These two parts of the report are built around a series of case studies that illuminate both the strengths and weaknesses of the existing framework of law, and demonstrate that more can and should be done. They conclude that there is a regulatory lacuna: existing statutory and regulatory frameworks do not provide sufficient accountability for the key individuals at the critical stage in the design and use of data-driven technologies. The Equality Act 2010 has much to offer, but it was not designed with AI or machine learning systems in mind. There is no single regulatory or other body charged with overseeing the use of AI, machine learning and data-driven technologies in the workplace: protection is piecemeal, and the existing regulators lack effective levers to hold big data companies to account when they take decisions that have unfair and discriminatory results.

Part 5

A new path forward

Part 5 sets out our recommendations for a new approach to the governance and regulation of algorithms, including AI and machine learning. This approach must be principle-driven and human-centred, with equality between citizens and social groups a central pillar of the new regime. And it must work across the entire innovation and deployment cycle, shift our emphasis to preventative action, and align our legal regimes and regulators.

Our central proposal is an overarching **Accountability for Algorithms Act (AAA)**. This would direct and inform policy, standards and behaviours of those involved in the design and deployment of data-driven technologies to ensure it serves the public interest. New dedicated legislation is the clearest and most pragmatic way to ensure that the specific allocation of responsibility and actions needed are clearly understood, undertaken and effectively enforced. And it will offer the direction that actors across the technology cycle – and the public – are demanding.

The AAA's key features are shown on the next page.

Part 5 *continued*

The AAA's key features

Corporate duties

New corporate duties, marking a shift in regulatory emphasis to pre-emptive governance and action in the public interest. These new duties would include duties on actors who are developing and/ or deploying algorithmic systems, as well as on other key actors across the design cycle and supply chain, **to pre-emptively evaluate those systems' equality impacts, and to make reasonable adjustments to eliminate unlawful discrimination and advance equality of opportunity and fair outcomes**.

These duties would be triggered by a risk-based contextual threshold. Algorithmic impact assessments should be rigorous and ongoing through the life cycle of the system. Strategic decisions would need to consider the desirability of reducing socio-economic and 'postcode' disadvantage too.

Increased transparency

A focus on increased transparency about key decisions across the innovation cycle and supply chain, including **a mandatory transparency obligation to record and report on facts of, purposes and outcomes of algorithmically-assisted decision-making in the workplace, and a freestanding right to know**, extending beyond technical design decisions to human decisions. Summary algorithmic impact assessments, including assessment of equality impacts, would be disclosed.

A new group right to know, for trade unions, other collectives and NGOs to enforce employers to understand and explain the impact on equality of their uses of data, and a **right for workers to be involved to a reasonable level** in the development and application of systems.

Intersectional regulatory forum

The establishment of a new intersectional regulatory forum to **coordinate, drive and align the work of existing regulators and enforce the AAA's new duties**, which would otherwise lie between the Equality and Human Rights Commission and the Information Commissioner's Office. New powers and additional resources would be needed to support specialist joint investigations and test cases, improve access to justice, and provide for cross-cutting statutory guidance. New powers would enable regulators to seek full algorithmic impact assessments with sources, methods and processes, as well as to suspend use of algorithms pending investigation.

Public consultation and dialogue

In addition, we propose wide public consultation and dialogue to contribute to the development, refinement and implementation of the AAA. This should start as soon as possible, and relevant areas for consultation should extend beyond the remit of this report.

The full report *Mind the gap: how to fill the equality and AI accountability gap in an automated world* **is available at www.ifow.org**

Institute *for the*
Future of Work

We research and develop practical ways to
improve work and working lives. We do that
by understanding how work is changing
and how we can make the future better
– in the face of technological change and
economic turmoil.

We believe that bringing people together
with different perspectives and experiences
enriches our understanding and ideas.
Through this collaborative approach, we
aim to create innovative, practical and
inclusive solutions.

We use the best interdisciplinary research
to shape policy and decision-making
in government and business to build
approaches that put people first.

Somerset House, Strand
London WC2R 1LA
T +44 (0)20 3701 7633

www.ifow.org
@FutureWorkInst