

AI and fairness in the workplace: why it matters and why now

*Artificial Intelligence is tremendously useful, when applied correctly, to improve fairness, transparency, efficiency and even to correct historical biases, including ethnic and gender biases imposed by humans. However, when applied in a 'plug and play' manner, we risk exacerbating social and economic inequality. **Christine Chow, Paris Will, and Mark Lewis** discuss the implications.*

From an investor's perspective

Investors increasingly consider human capital management (HCM) to be a defining factor in delivering business outperformance. In discussions around the 'future of work', we acknowledge that the future is already here – the workforce is now more mobile, more diverse, more agile, and increasingly impacted by more data and analytics.

COVID lockdown has accelerated the digital transformation of businesses – from CV screening and AI video interviews, to virtual on boarding and digital workplace interactions. Workers' activities in a hybrid working environment provide ample opportunities for technologies to improve efficiency, but not all companies are clear and explicit about when and how analytics are used on workers. In a [recent BBC episode](#), AI technologies created for hiring were used in redundancy targeting, with unintended psychological impact on workers and possibly brand and financial impact on employers. In [Social Dilemma](#), a 2020 docudrama, interviews with tech leaders highlighted the dangers of collected information at workplace, including how it affects the human psyche by causing anxiety and depression.

Share Action, a UK-based NGO, established the Workforce Disclosure Initiative (WDI) in late 2016 to improve corporate transparency and accountability on workforce issues. WDI is supported by 68 institutions (including HSBC Asset Management) with asset managers covering over US\$10 trillion. In 2021, WDI introduced a new indicator, asking companies to describe workforce surveillance measures used to monitor workers and how the company ensures this does not have a disproportional impact on workers' right to privacy. Proprietary findings can be found in our white paper (see Notes section below).

Beyond surveillance, investors recognise that it is in general difficult to reduce human behaviour down to an algorithm. Where AI works well as a tool is when decision-making rules are quantifiable, and computers can be used to interact almost instantaneously with vast amounts of data. Where AI does not work well as a tool is when the interpretation of human characteristics is contextual and the impact on behaviour is nuanced.

Christine Chow

From an academic's perspective

A scientific approach to assessing AI decision-making in human capital management (HCM) is to evaluate the level of bias in algorithmic decisions. As AI has been proposed to eradicate human biases in areas such as hiring, promotion, and pay decisions, it must be assessed to ensure that it lessens rather than potentiates human cognitive biases. As research advances in the field and conclusions are made about the adoptability of certain types of AI for HCM, we must be wary of the way that we judge AI, as [research has shown](#) that we judge AI in comparison to "perfect" decision-making. This may explain the backlash against adopting AI. The reality is that adopting AI will replace human decisions, which are far from perfect, and often riddled with bias. As such, a better approach in assessing the adoptability of AI for HCM is to directly compare AI and human decision outcomes. AI can be adopted in contexts where it shows less biased decision-making than humans.

The capabilities of AI are thought of in two broad categories: weak AI and strong AI. Weak AI means that it is specific to a task, whereas strong AI is meant to mimic human intelligence and can generalise to other contexts. AI currently performs well in specific contexts but is limited in its generalisability. As such, for HCM, we should take caution in using AI that has been validated in a specific context's data set. We must keep considerations in mind when thinking about the capabilities of AI for HCM, especially as many of its proposed uses reflect technologies that would require strong AI. In this way companies should set their expectations about what AI can reliably accomplish.

Due to a rapidly changing workplace environment, we recommend using dynamic rather than static AI – algorithms should be continually updated with new data at regular intervals to prevent decisions from being outdated, irrelevant, and worse, unfit for purpose.

Paris Will

From a legal and regulatory perspective

It is surely in the interests of all concerned – especially those who may not have known that they should be concerned – to understand the legal and regulatory implications of introducing and deploying AI systems and processes in and around the workplace. The consequences of failing to understand and address those implications could be very serious for a range of stakeholders.

In trying to understand those implications, the problem is that there is very little law and regulation around the world applying specifically to the use of AI in human capital management. There are prospective laws and regulations in various jurisdictions that are intended to address the subject specifically, most notably a very substantial regulation proposed for the European Union – and not just in the EU, because it will have effect outside the EU, too. But, in many countries there is an existing body of law and regulation that would apply to employers' use of AI in the workplace.

Unfortunately, when, how, and why that body of law and regulation may apply, and the implications of it applying, aren't instantly obvious. What *is* obvious, however, is that there is a growing movement of disquiet about, and resistance to, the use of AI, automated decision-making, and related processes in and around the workplace, and that regulatory and legal challenges are happening – now.

Our white paper seeks to address this problem: first, by addressing board accountability and governance considerations; next by alerting readers to when, how, and why law and regulation will or may apply here; and then by outlining:

- a selection of laws and regulations that now apply specifically to the adoption and use of AI in the workplace;
- the potential impact of the proposed EU Artificial Intelligence Act; and
- certain existing laws and regulations in various countries that are most likely to apply in such circumstances, but with judicial interpretation and application.

In doing so, we hope that all those concerned – and those who should be concerned – will be better able to understand, plan for, and manage the governance, legal and regulatory, and reputational implications, risks, and exposure that could arise in adopting and relying on AI systems and processes in and around the workplace.

Mark Lewis

The paper acknowledges that AI analytics is unlikely to go away in the present and future of work. As such, to make the best and appropriate use of AI in human capital management we need to embrace it with caution and respect for workers:

- Board level oversight should focus on clarifying the intent of AI applications; establishing clear processes for AI training and testing, including back testing; ensuring transparency in the use of AI applications and accountability for their outcomes; and requiring competent and thorough due diligence of third-party products in the supply chain that may use AI. Such due diligence is currently often missing, or conducted by people

- lacking sufficient expertise;
- Companies should understand and properly manage the legal and regulatory implications of adopting and relying on AI systems and related processes in the workplace;
 - Companies should put in place a machine learning platform to create automated and repeatable data preparation processes and feature engineering to ensure consistency over time. The platform should have a data versioning function to keep track of changes in data used, different model experiments and test results. There should be a feedback loop and associated documentation that integrate outcomes and lessons learnt to improve performance over time;
 - Companies should include a means of redress and clear remediation mechanisms to an accountable human; and
 - When engaging with companies, investors can ask for the disclosure and explanation of key performance indicators (KPI) highlighted in our paper, which covers hiring, culture, and performance (Figure 1).

Figure 1. Categorisation of the human capital management process

Hiring	Culture	Performance
Workforce planning	Competency	Remuneration
Recruitment	Team collaboration	Attrition
On-boarding	Mobility	Learning



Notes:

- *This blog post marks the launch of the white paper ‘Future of Work: Investors’ Expectations of Ethical Artificial Intelligence in Human Capital Management’ on 7 April 2022. The zoom registration can be made [here](#).*
- *The post represents the views of its author(s), not necessarily those of LSE Business Review or the London School of Economics and Political Science.*
- *Featured [image](#) by [geralt](#), under a [Pixabay](#) licence*
- *When you leave a comment, you’re agreeing to our [Comment Policy](#).*