



**EMPLOYMENT PRACTICE GROUP**

**NEW CLIENT PUBLICATION**

<b>Date</b>	11 May 2023
<b>Title</b>	Retrenched by HAL 9000: what employers should know before using AI in dismissals based on operational requirements
<b>Author(s)</b>	Mehnaaz Bux, Keah Challenor

*Employers using AI systems to identify employees for retrenchment need to be cautious of potential discrimination and should ensure that human bias is not systematised. Employers should consider the human-centred AI Principles adopted by the Organisation for Economic Co-operation and Development for best practice guidance and ensure compliance with local labour laws.*

When employers contemplate dismissing one or more employees for operational requirements, section 189(2)(b) of the Labour Relations Act (LRA) requires employers to engage in a meaningful joint consensus-seeking process with the affected employees. In this process, they should attempt to reach consensus on, among other things, the method for selecting which employees to dismiss. The consulting parties must agree to the selection criteria, or, if no criteria were agreed, criteria that are fair and objective.

According to a recent survey conducted by the Society for Human Resource Management, nearly one in four employers uses automation and artificial intelligence (AI) to support human resource-related tasks. AI systems enable the automated processing of numerous types of data, producing outcomes and recommendations rapidly and at scale. At first glance, using AI to decide which employees are to be selected for retrenchment may appear to be the perfect way to ensure fairness and objectivity. However, unless employers can prove that the algorithm(s) used to make such decisions are unbiased, they may unintentionally find themselves falling foul of the LRA.

Once the criteria are established, employers may consider using an AI system to identify which employees should be retained and which should be retrenched. This would certainly remove any scope for favouritism or human error. However, employers need to guard against the use of AI systems that may recommend a result which could be construed as discriminatory.

Some employers have found that developing a 'neutral' programme is easier said than coded. For example, Amazon abandoned the development of a CV analysis algorithm which unintentionally showed a bias against female candidates. The algorithm was designed to scan CVs and pick out those that were similar to CVs submitted by candidates that were ultimately hired. However, given that the majority of the CVs provided to the AI system as examples of 'good' CVs were those of men, the algorithm inadvertently preferred CVs submitted by men (over women). The algorithm penalised CVs that included the word "women", for example, "captain of women's soccer team". While AI systems have the potential to improve fairness in the workplace, there is also a risk that human bias may be multiplied and systematised.

Existing legislation on anti-discrimination, data protection, and rights to due process in the workplace must of course be enforced when AI systems are used in the workplace, for retrenchments or other tasks.

While employers may not have databases that include information such as an employee's religion or political opinions, the possibility of discrimination creeping into algorithms remains. Consider the following example: following consultations, employers and employees have agreed that retention of essential skills is a valid criterion for determining which employees will be dismissed. If, in that workplace, the majority of the holders of those essential skills have never taken maternity leave, the employer will need to ensure that the algorithm does not interpret pregnancy as an indicator that an employee does not possess essential skills.

A dismissal is automatically unfair when it is *directly or indirectly* based on an arbitrary ground, including race, gender, sex, ethnic or social origin, colour, sexual orientation, age, disability, religion, conscience, belief, political opinion, culture, language, marital status or family responsibility. In May 2019, the Organisation for Economic Co-operation and Development member states adopted human-centred AI Principles. These principles are a useful guide for employers navigating the implementation of AI systems in the workplace. They include inclusivity, human-centred values and fairness, transparency, robust security and safety, and accountability when it comes to decision-making. Various cases in the US and EU have required employers to disclose data / algorithmic programming IP used in their AI systems, or reinstate individuals dismissed solely based on those algorithms.

With the risk of discrimination in mind, any employer using AI systems to identify employees for retrenchment would be advised not to give an algorithm full discretion. If an employee alleges that they were selected for retrenchment based on the use of a biased AI tool, the employer may be faced with: (1) an allegation that it did not follow a fair procedure when dismissing for operational requirements; or (2) unfair dismissal claims (potentially automatically unfair dismissal claims, depending on the circumstances).

Even if AI systems do not involve full automation and humans are involved in various ways, human decision-making is likely to be profoundly affected by AI systems that encourage new ways of approaching, understanding, and acting upon information. Learning to work AI is an unavoidable reality that employers and their legal teams must navigate with caution. The rate at which AI technology is developing is likely to pose significant implications for employers, particularly because AI can be perceived as leading to job losses. Successfully adapting to new ways of working is essential for employers. This could include implementing measures and strategies to upskill and reskill workers.