

**Artificial Intelligence in Recruitment and Performance Evaluations:
International Labour Law Perspectives on Equality
and Non-Discrimination**

PhD. student **Alexandra KOVACS**¹

Abstract

This study explores the use of artificial intelligence (AI) in recruitment and performance evaluation from the perspective of international labour law, focusing on equality and non-discrimination. The main objective is to examine how AI-driven human resource tools may affect workers' rights and compliance with international standards, including ILO principles and relevant EU directives. The research employs comparative legal analysis, reviewing legislation, policy documents, and scholarly literature, supplemented by illustrative hypothetical scenarios. The legal analysis suggests that while AI can improve efficiency and objectivity, it also carries risks of bias and unequal treatment if not properly regulated. The study identifies best practices to mitigate these risks, such as algorithmic transparency, regular audits, and inclusive design. The implications highlight the importance of clear legal frameworks to ensure AI supports fair and non-discriminatory workplace practices. This research aims to provide guidance for policymakers, employers, and legal practitioners seeking to align technological innovation with international labour standards.

Keywords: artificial intelligence, recruitment and performance evaluation, international labour law, equality and non-discrimination, algorithmic bias.

JEL Classification: K24, K31, K33

DOI: 10.62768/PLPA/2026/15/1/10

Please cite this article as: Kovacs, Alexandra „Artificial Intelligence in Recruitment and Performance Evaluations: International Labour Law Perspectives on Equality and Non-Discrimination”, <i>Perspectives of Law and Public Administration</i> 15, no. 1 (March 2026): 141-157.	Article History Received: 15 January 2026 Revised: 10 February 2026 Accepted: 12 March 2026
--	---

1. Introductory Considerations

The increasing integration of artificial intelligence (AI) into employment-related decision-making processes marks a significant shift in the organization of work and the exercise of managerial prerogatives. What was once the domain of human judgment — recruitment, evaluation, promotion, and dismissal — is now, at least in part, mediated by algorithmic systems that rely on large datasets and predictive modelling. While these developments are frequently framed in terms of efficiency,

¹ Alexandra Kovacs - Deák Ferenc Doctoral School of Law, University of Miskolc, Hungary, drkovacsalexandra@gmail.com. ORCID:009-0005-2146-8774.

objectivity, and cost reduction, they simultaneously raise complex legal questions that cannot be adequately addressed without revisiting foundational principles of labour law. From a doctrinal perspective, one of the most pressing concerns relates to the compatibility of algorithmic decision-making with the principle of equality and the prohibition of discrimination. These principles, deeply embedded in international labour law, were developed in a context where decision-making was assumed to be human, intentional, and, at least in principle, explainable. The emergence of opaque, data-driven systems challenges these assumptions in fundamental ways. As a result, the question is no longer merely whether discrimination occurs, but how it manifests in environments where decisions are produced through statistical correlations rather than explicit intent.

The rise of AI in employment decision-making has been particularly pronounced in two areas: recruitment and performance evaluation. In recruitment, algorithmic tools are increasingly used to filter applications, rank candidates, and even conduct preliminary assessments through automated evaluation tasks. These systems often rely on historical data to identify patterns associated with “successful” employees, thereby introducing the risk that past inequalities are reproduced in present decision-making processes². In performance evaluation, similar technologies are employed to monitor productivity, assess behavioral patterns, and generate performance scores, which may in turn influence remuneration, promotion, or termination decisions³.

What makes these developments legally significant is not merely the use of technology as such, but the shift in the nature of decision-making. Algorithmic systems operate by identifying correlations within datasets, often without providing a transparent account of how specific variables influence outcomes. This “black box” characteristic complicates traditional legal analysis, particularly in relation to indirect discrimination. Whereas conventional cases of indirect discrimination typically involve identifiable criteria that can be scrutinized for their disproportionate impact, algorithmic systems may rely on complex combinations of variables, some of which function as proxies for protected characteristics such as gender, ethnicity, or age⁴.

Against this background, the principle of equality and non-discrimination emerges as a central analytical lens through which the use of AI in employment must be assessed. International labour law has long recognized that formally neutral practices may produce substantively unequal outcomes, thereby giving rise to indirect discrimination. This doctrinal framework is, at least in principle, well suited to addressing algorithmic bias. However, the application of these norms to AI-driven

² Solon Barocas, Andrew D. Selbst, „Big Data’s Disparate Impact”, *California Law Review*, vol. 104, issue 3, 2016: 671–732, <https://doi.org/10.15779/Z38BG31>.

³ Valerio De Stefano, „Negotiating the Algorithm: Automation, Artificial Intelligence and Labour Protection”, *International Labour Organizations*, Employment Working Paper No. 246, 2019: 15–34, <https://www.ilo.org/publications/negotiating-algorithm%E2%80%9Dautomation-artificial-intelligence-and-labour>, accessed on 9 February 2026.

⁴ Sandra Wachter, Brent Mittelstadt, Luciano Floridi, „Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, *International Data Privacy Law*, Volume 7, Issue 2, May 2017: 76–99, <https://doi.org/10.1093/idpl/ixp005>.

decision-making is far from straightforward. The opacity of algorithmic systems, the scale at which they operate, and the technical expertise required to interrogate them all pose challenges to the effective enforcement of existing legal standards⁵.

The present article proceeds from the assumption that AI does not introduce an entirely new category of legal problems but rather intensifies and reconfigures existing ones. In particular, it raises the question of whether established concepts such as indirect discrimination, burden of proof, and employer liability can be meaningfully applied in contexts where decision-making is partially or wholly automated. This inquiry is not merely theoretical. It has direct implications for the protection of workers' rights and the allocation of responsibility within increasingly complex organizational structures.

The central research questions guiding this analysis are therefore the following: First, to what extent can existing international labour law norms on equality and non-discrimination be applied to AI-assisted employment decisions? Second, does algorithmic decision-making give rise to new forms of discrimination that cannot be adequately captured by current legal categories? Third, how should responsibility be allocated when discriminatory outcomes result from the use of algorithmic systems developed or operated by third parties? These questions reflect a broader concern with the adaptability of labour law in the face of technological change.

Methodologically, the article adopts a normative and analytical approach, grounded in the interpretation of international labour law instruments and relevant doctrinal literature. This is complemented by published examples and hypothetical scenarios concerning recruitment and performance evaluation. The hypothetical scenarios illustrate possible legal problems rather than establish how particular organizations operate. Their purpose is to clarify the application of equality and non-discrimination principles to algorithmic decisions. The article therefore develops a doctrinal assessment of legal risks and regulatory responses, with the illustrative material serving to explain the reasoning.

An illustrative example may serve to highlight the nature of the problem. Consider a recruitment system that relies on historical hiring data to identify candidates who are likely to succeed in a given role. If the dataset reflects a workforce that has been predominantly male, the algorithm may implicitly learn to favour characteristics associated with male candidates, even in the absence of any explicit gender variable. As a result, female applicants may be systematically ranked lower, despite possessing equivalent qualifications. From a legal perspective, this scenario raises the question of whether such outcomes can be classified as indirect discrimination, given that the criteria used by the algorithm are not immediately transparent and may not be intentionally discriminatory⁶.

A similar issue arises in the context of performance evaluation. In platform-based work environments, workers are often assessed through automated rating systems that aggregate customer feedback and behavioural data. These ratings can have

⁵ Sandra Fredman, *Discrimination Law*, Oxford University Press, 2011, pp.: 175–198.

⁶ Jeffrey Dastin, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women*, Reuters, 2018, pp.: 1–3.

significant consequences, including reduced access to work opportunities or even deactivation from the platform. However, research suggests that such systems may be influenced by biases on the part of customers, which are then amplified by the algorithmic logic of the platform⁷. In this way, ostensibly neutral evaluation mechanisms may produce systematically unequal outcomes for different groups of workers.

These examples underscore the need for a careful reassessment of the relationship between technology and legal norms. They suggest that the challenges posed by AI are not limited to issues of technical design but extend to the core principles that underpin labour law. They call into question the adequacy of existing frameworks for identifying and remedying discrimination in contexts where decision-making is diffuse, data-driven, and difficult to scrutinize.

At the same time, it is important to avoid overstating the novelty of these challenges. Labour law has long grappled with practices that produce unequal outcomes without explicit discriminatory intent. The doctrine of indirect discrimination, in particular, provides a conceptual tool for addressing such situations. The key issue, therefore, is not whether the law is entirely unprepared for AI, but whether it can be effectively applied in contexts characterized by technological complexity and informational asymmetry.

In this respect, the article adopts a cautiously critical perspective. It acknowledges the potential of AI to enhance efficiency and consistency in employment decision-making, while also emphasizing the risks it poses to fundamental rights. Rather than advocating for entirely new legal categories, the analysis focuses on the interpretation and adaptation of existing norms. This approach reflects the view that the core values of labour law — fairness, equality, and the protection of vulnerable workers — remain relevant, even as the mechanisms through which decisions are made undergo significant transformation.

Ultimately, the objective of this study is to contribute to an ongoing scholarly and policy debate on the regulation of AI in the workplace. By examining the intersection of algorithmic decision-making and the principle of non-discrimination, it seeks to clarify the legal issues at stake and to identify potential pathways for future development. In doing so, it aims to demonstrate that the challenges posed by AI are not merely technical, but fundamentally normative, requiring careful consideration of the values that labour law is intended to protect.

2. The Use of Artificial Intelligence in Employment: Functioning and Risks

The integration of artificial intelligence into employment-related decision-making is often described in abstract or technical terms. However, from a labour law perspective, what matters is not the sophistication of the technology itself, but the way in which it restructures decision-making processes that have traditionally been carried

⁷ Alex Rosenblat, Luke Stark, „Algorithmic Labor and Information Asymmetries: A Case Study of Uber’s Drivers”, *International Journal of Communication*, vol. 10, 2016: 3758–3784.

out by human actors. In order to assess its legal implications, it is therefore necessary to reconstruct, in simplified terms, the operational logic of algorithmic systems and to examine how this logic manifests in concrete employment contexts.

At its most basic level, algorithmic decision-making in employment can be understood as a three-step process: data collection, model construction, and decision output. First, large amounts of data are gathered, often from internal company records or external sources. These data may include information about previous employees, their performance, career progression, and, in some cases, behavioural patterns. Second, this data is used to train a model, typically through machine learning techniques, which identifies patterns and correlations associated with desired outcomes. Finally, the model is applied to new cases, generating predictions or recommendations that inform or determine employment-related decisions⁸.

While this sequence may appear neutral, each stage involves choices that have legal significance. The selection of data, the definition of “success,” and the parameters of the model all reflect underlying assumptions that may not be immediately visible. An employer may perceive these systems as objective precisely because they rely on quantitative data, without examining how that data was selected. Such a perception may obscure the extent to which algorithmic outputs depend on historically contingent and potentially biased inputs.

The use of AI in recruitment provides a particularly clear illustration of this dynamic. Many organizations now rely on automated tools to filter large volumes of applications, especially for entry-level or high-turnover positions. These systems typically rank candidates based on their similarity to previously successful employees. In practice, this means that historical hiring decisions—however imperfect—become embedded in the logic of the algorithm.

A frequently cited example in the literature concerns a large technology company that developed an internal recruitment tool trained on historical CV data. The system was designed to identify candidates most likely to succeed, but it was eventually discovered that it systematically downgraded applications containing indicators associated with female candidates⁹. Although the system did not explicitly use gender as a variable, it relied on proxies derived from past hiring patterns. What is particularly striking about this case is not only the presence of bias, but the fact that it remained undetected for a significant period of time, despite the company’s technical expertise.

Beyond such high-profile cases, similar dynamics can be observed in more routine recruitment practices. For a hypothetical illustration, consider smaller firms using commercially available CV-screening tools that assign scores based on educational background, employment history, and linguistic patterns. Even if these tools are presented as neutral and efficiency-enhancing, their functioning raises important questions. For instance, if a model is trained on data reflecting a workforce with limited diversity, it may implicitly learn to favour candidates who resemble the existing

⁸ Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, and Harlan Yu, „Accountable Algorithms”, *University of Pennsylvania Law Review*, vol. 165, issue 3, 2016: 633–705.

⁹ Jeffrey Dastin, *op. cit.*, pp.: 1–3.

employee base. This does not necessarily result from intentional discrimination, but it can nonetheless produce systematically unequal outcomes.

The risks associated with algorithmic decision-making are not confined to recruitment. They are equally, if not more, pronounced in the context of performance evaluation. Here, AI systems are used to monitor worker behaviour, assess productivity, and generate performance metrics that may influence managerial decisions. This is particularly evident in platform-based work environments, where the entire labour process is mediated through digital interfaces.

In the case of ride-hailing or delivery platforms, for example, workers are typically evaluated through a combination of automated metrics and customer ratings. These ratings are aggregated into scores that determine access to future work opportunities, bonuses, or, in some cases, continued access to the platform itself. From a legal standpoint, the crucial issue is that these systems operate continuously and at scale, often without meaningful human oversight.

Research has shown that customer ratings may be influenced by factors unrelated to the quality of the service provided, including biases linked to race, gender, or accent¹⁰. When such ratings are incorporated into algorithmic evaluation systems, these biases can be amplified rather than mitigated. In effect, the algorithm does not merely reflect individual prejudices but transforms them into structured and consequential patterns of decision-making.

A similar logic can be observed in more traditional employment settings through the use of productivity tracking systems. These tools monitor various aspects of employee activity, such as time spent on tasks, keystroke frequency, or responsiveness to communications. The resulting data is then used to generate performance scores or rankings. While these systems are often justified on the grounds of efficiency and accountability, they raise concerns about the reduction of complex human activities to quantifiable indicators.

Consider a hypothetical company that introduces a monitoring system to track the online activity of its remote workforce. Employees whose recorded activity falls below a threshold would be flagged as underperforming and referred for managerial review. If the system fails to account for differences in job roles or working styles, employees engaged in tasks requiring sustained concentration or offline work could receive lower scores despite meeting their responsibilities. Although the same criteria would apply to everyone, their effects could differ substantially, illustrating why performance indicators require assessment in their specific employment context.

These examples illustrate a broader point: algorithmic systems tend to privilege what can be measured, often at the expense of what cannot. In doing so, they reshape the criteria by which workers are evaluated, sometimes in ways that are neither transparent nor contestable. From a labour law perspective, this raises important questions about fairness, accountability, and the ability of workers to challenge adverse decisions.

Another characteristic feature of AI systems in employment is their opacity.

¹⁰ Alex Rosenblat, Luke Stark, *op. cit.*, pp. 3758–3784.

Even where employers are willing to disclose the use of algorithmic tools, they may not be able to provide a clear explanation of how specific decisions are reached. This is partly due to the complexity of machine learning models, but also to the fact that many systems are developed by external vendors and operate as proprietary “black boxes.” As a result, both employers and employees may have limited insight into the decision-making process.

This lack of transparency has direct implications for the enforcement of non-discrimination norms. In traditional cases, it is often possible to identify a specific rule or criterion that produces a disparate impact. By contrast, algorithmic systems may rely on a multitude of interacting variables, making it difficult to isolate the source of any given outcome. This complicates not only the identification of discrimination but also the allocation of responsibility.

In light of these considerations, it becomes clear that the legal challenges posed by AI in employment are closely linked to its operational logic. The reliance on historical data, the use of predictive modelling, and the opacity of decision-making processes all contribute to a situation in which unequal outcomes may arise without explicit discriminatory intent. At the same time, the scale and speed at which these systems operate mean that such outcomes can affect large numbers of workers in a relatively short period of time.

The purpose of this section has been to demonstrate that AI in employment is not an abstract or hypothetical phenomenon, but a set of concrete practices that are already reshaping the labour market. By examining how these systems function in recruitment and performance evaluation, it becomes possible to identify the points at which legal principles—particularly those relating to equality and non-discrimination—are likely to be engaged. The following section builds on this analysis by outlining the relevant legal framework and assessing its applicability to the scenarios described above.

3. Equality and Non-Discrimination in International Labour Law

The principle of equality occupies a central place in international labour law, reflecting the enduring commitment of the international community to ensure that all workers are treated fairly, irrespective of their personal characteristics. The doctrinal distinction between formal and substantive equality indicates that this principle is more than a requirement of identical treatment: it also addresses barriers to genuine parity of opportunity across diverse groups. Understanding this distinction is crucial for evaluating the impact of modern employment practices, particularly in contexts where algorithmic or data-driven systems are employed.

At a conceptual level, equality can be understood in two complementary dimensions: formal and substantive. Formal equality entails that individuals in similar circumstances should be treated alike, without explicit preference or disadvantage based on protected characteristics such as gender, age, ethnicity, or religion¹¹. In practice,

¹¹ Sandra Fredman, *op. cit.*, pp. 23–54.

formal equality is often reflected in statutory prohibitions against direct discrimination—where an overt criterion explicitly disadvantages a particular group. For example, a policy requiring that only male candidates may apply for a specific position constitutes a clear violation of formal equality principles. From a legal perspective, prohibiting such explicit criteria is an essential safeguard, but their absence cannot establish compliance where other recruitment conditions may still disadvantage a protected group without objective justification.

Substantive equality, by contrast, recognizes that equal treatment in formal terms does not necessarily result in equitable outcomes. Structural and historical disadvantages may persist even when policies appear neutral. As an illustration, consider the widespread requirement for prior managerial experience in recruitment for certain executive positions. While the criterion does not mention gender, it may disproportionately disadvantage women if historical patterns have limited their access to such roles. In a hypothetical recruitment process, a CV-screening tool could reproduce these patterns by privileging candidates whose profiles resemble the historical workforce, despite applying the same selection criteria to every applicant¹². Substantive equality thus requires a focus not only on the rules themselves but on their effects.

The distinction between formal and substantive equality is closely linked to the legal concepts of direct and indirect discrimination. Direct discrimination occurs when an explicit criterion is applied in a manner that disadvantages a protected group. In contrast, indirect discrimination arises when a neutral policy or practice disproportionately impacts a particular group and cannot be objectively justified. A hypothetical recruitment policy requiring particular qualifications or lengthy experience may disadvantage minority candidates if access to those credentials has been unequal. Whether this amounts to indirect discrimination requires examining the disadvantage and any objective justification. Statistical comparisons can help assess possible disparities, which supports the use of appropriately designed monitoring and review mechanisms¹³.

International labour law provides a comprehensive framework for addressing these issues. The International Labour Organization (ILO) has been a central actor in establishing principles of equality and non-discrimination in employment. Convention No. 111 concerning Discrimination in Respect of Employment and Occupation (1958) defines discrimination broadly, covering any distinction, exclusion, or preference that affects employment opportunities on prohibited grounds¹⁴. Importantly, the Convention recognizes both direct and indirect forms of discrimination, thereby embedding the

¹² Bob Hepple, *Equality: The Legal Framework*, Second edition. Oxford: Hart Publishing, 2014, pp. 67–98

¹³ Myrtle P. Bell and Daphne P. Berry, „Viewing Diversity Through Different Lenses: Avoiding a Few Blind Spots”, *Academy of Management Perspectives*, Vol. 21, No. 4, 2007: 21-25, <https://doi.org/10.5465/amp.2007.27895336>.

¹⁴ International Labour Organization (1958), *Convention No. 111 concerning Discrimination in Respect of Employment and Occupation*, ILO Publications, pp.: 1–12, <https://www.ilo.org/media/27816/download>, accessed on 9 February 2026.

concept of substantive equality into the international legal corpus. As a matter of normative interpretation, these standards provide a reference point for domestic legislation and organizational compliance policies, including the assessment of recruitment and evaluation practices in multinational enterprises.

In addition to ILO instruments, United Nations human rights conventions reinforce these principles. The International Covenant on Civil and Political Rights (1966) and the Convention on the Elimination of All Forms of Discrimination Against Women (1979) articulate rights to non-discrimination that extend into the sphere of employment. While these treaties are framed in general human rights terms, their interpretation by treaty bodies has increasingly emphasized the practical obligations of states and employers to monitor outcomes and to take positive measures where structural inequalities exist¹⁵. Applied to a hypothetical workplace inclusion policy, these principles offer criteria for assessing recruitment, promotion, and retention practices, without presupposing that any particular organization has achieved compliance.

Regional legal systems provide additional mechanisms to operationalize equality. The European Union, for instance, has developed a sophisticated body of legislation prohibiting discrimination on multiple grounds, including gender, racial or ethnic origin, religion, disability, age, and sexual orientation. Directives such as 2000/78/EC establish general frameworks for equal treatment in employment, explicitly addressing indirect discrimination and placing obligations on employers to justify neutral practices that produce disproportionate adverse effects¹⁶. In applying these rules to an illustrative dispute, statistical comparisons may help assess whether a neutral criterion causes a particular disadvantage, while leaving the factual evaluation and any objective justification to the competent decision-maker.

Analogies from traditional discrimination cases are instructive for understanding the challenges posed by modern employment practices. Consider the historical case of a company that required a minimum height for certain operational roles. While the policy applied equally to all applicants, statistical analysis revealed that it disproportionately excluded female candidates. Courts treating this as indirect discrimination highlighted the need to examine not only the formal wording of employment criteria but also their real-world impact¹⁷. The legal implication for human resources policies is therefore that apparently neutral requirements must be assessed in light of their effects on protected groups and any relevant objective justification.

Other illustrative cases concern educational requirements, physical fitness tests, or seniority-based criteria, which often disproportionately affect ethnic minorities or women due to structural inequities in access to prior opportunities. These analogies are directly relevant for assessing AI-mediated decision-making, because algorithmic

¹⁵ United Nations (1979), *Convention on the Elimination of All Forms of Discrimination Against Women*, UN Publications, pp.: 5–23, <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-elimination-all-forms-discrimination-against-women>, accessed on 9 February 2026.

¹⁶ European Union (2000), Council Directive 2000/78/EC establishing a general framework for equal treatment in employment and occupation, OJ L 303, 2.12.2000, pp. 16–22.

¹⁷ Case law analogy: *Robinson v. Secretary of State for Employment* (UK, 1997), pp.: 101–108.

models often rely on historical data reflecting precisely these past inequalities. For example, a hypothetical automated CV-screening system that rewards prior employment at prestigious firms or attendance at elite universities could reproduce unequal access to those institutions, creating a risk that requires assessment under indirect discrimination rules.

Importantly, these examples demonstrate that legal principles developed to address traditional forms of bias remain applicable in contemporary employment contexts. They also underscore the need for careful design, monitoring, and auditing of organizational practices, whether human- or algorithm-mediated, to ensure that equality and non-discrimination obligations are met. Regular statistical audits could help an employer identify possible disparities in recruitment and performance evaluation, document the reasons for challenged criteria, and assess whether alternative arrangements would reduce unjustified disadvantages.

The overarching lesson from this analysis is that equality and non-discrimination are not abstract ideals, but practical obligations that require continuous attention to outcomes and effects. Both formal and substantive equality must be considered, and employers need to be aware that neutral policies can produce unequal impacts. The doctrinal analysis suggests that integrating these considerations into human resources policies, training, and monitoring systems is essential for compliance with international labour law and for fostering genuinely inclusive workplaces.

Furthermore, understanding the interplay between traditional legal frameworks and contemporary employment practices provides a foundation for examining emerging issues in algorithmic decision-making. While AI and other digital tools introduce new mechanisms for decision-making, they do not invalidate the principles established in international labour law. Rather, they magnify the need for rigorous attention to equality and non-discrimination, particularly in contexts where historical patterns of exclusion are encoded in data and may be reproduced at scale.

In conclusion, the principle of equality in international labour law encompasses both formal and substantive dimensions, with direct and indirect discrimination representing the primary mechanisms through which violations occur. International instruments from the ILO, UN, and regional systems provide a robust framework for addressing these issues, and analogies from traditional discrimination cases offer practical guidance for contemporary practice. The normative analysis supports applying these principles to organizational decisions concerning recruitment, evaluation, and broader employment practices, while recognizing that any finding of discrimination would depend on evidence and the applicable legal test. This analysis provides the foundation for the subsequent exploration of algorithmic decision-making, highlighting the points at which legal norms intersect with evolving employment technologies.

4. Algorithmic Decision-Making and Discrimination: A Legal Assessment

The increasing reliance on algorithmic systems for employment decision-making raises questions that are both practical and legal. From a doctrinal perspective, the challenges presented by such systems often manifest less in overtly discriminatory

intent and more in subtle, structural biases embedded in data and model design. Algorithmic decision-making, by its nature, translates complex human and organizational histories into quantifiable patterns, which can inadvertently reproduce pre-existing inequalities¹⁸.

At the core of these challenges is the phenomenon of algorithmic bias. Bias arises when the data used to train a model reflects historical disparities or selective sampling. In a hypothetical recruitment process, a CV-screening tool could prioritize candidates from elite educational institutions or particular previous employers, using those attributes as indicators of likely suitability. Although these criteria may appear neutral, they can reproduce patterns of access that historically favoured certain socio-economic or demographic groups. Consequently, candidates from less representative groups may be systematically disadvantaged, illustrating a form of indirect discrimination¹⁹.

Indirect discrimination in this context aligns closely with the legal definitions established in international and regional labour law. A neutral criterion, if it produces a disproportionate negative effect on a protected group and cannot be objectively justified, constitutes unlawful discrimination. The doctrinal difficulty lies in identifying whether an apparently neutral output reproduces disadvantage associated with historical hiring practices. Consider a hypothetical recruitment algorithm that ranks male candidates higher for technical roles despite excluding gender as an explicit input. If the training data reflects an overrepresentation of men in those positions, the resulting rankings may perpetuate that imbalance. This scenario illustrates why the selection criteria, their effects, and any objective justification require separate examination²⁰.

Beyond recruitment, algorithmic systems are increasingly used to evaluate employee performance. In platform-based environments, such as delivery or ride-sharing services, workers are assessed continuously through performance scores derived from customer ratings and task completion metrics. As a hypothetical illustration, customer ratings might incorporate prejudices unrelated to service quality, including reactions to a worker's accent or communication style. If an algorithm uses such ratings to allocate work or bonuses, those prejudices could produce material disadvantages²¹. The legal assessment would then require examining whether the system causes an unjustified adverse effect on a protected group.

The interplay between algorithmic bias and legal accountability is further complicated by issues of transparency and explainability. Traditional labour law relies on the ability to identify specific criteria or decisions that can be evaluated for compliance with non-discrimination norms. Algorithmic systems, particularly those based on complex machine learning models, often operate as “black boxes,” providing

¹⁸ Solon Barocas, Andrew D. Selbst, *op. cit.*, pp. 671–732.

¹⁹ Manish Raghavan, Solon Barocas, Jon Kleinberg, Karen Levy, „Mitigating Bias in Algorithmic Hiring: Evaluations and Lessons”. In : Mireille Hildebrandt, Carlos Castillo, Elisa Celis, Salvatore Ruggieri, Linnet Taylor, Gabriela Zanfir-Fortuna (eds.), *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ACM, 2020, pp. 469–481, <https://doi.org/10.1145/3351095.3372828>.

²⁰ Jeffrey Dastin, *op. cit.*, pp. 1–3.

²¹ Alex Rosenblat, Luke Stark, *op. cit.*, pp. 3758–3784.

output without a clear explanation of the underlying reasoning. For the purposes of legal analysis, this lack of transparency presents a significant challenge: when a worker contests a decision, the employer may struggle to demonstrate whether or how discrimination occurred. Such opacity heightens the burden of proof for both parties and raises the question of how legal norms can adapt to technological complexity²².

The burden of proof is a critical issue in this context. Traditional employment law often allows for a shift of the burden of proof once a *prima facie* case of discrimination is established. However, when algorithmic decisions are involved, establishing this initial case is difficult, given the complexity of the data and model interactions. In a hypothetical grievance procedure, limited access to relevant data or an explanation of the model's operation could restrict an employee's ability to demonstrate disparate impact. This asymmetry underscores the need for procedural mechanisms that facilitate transparency, such as mandated audits or access to algorithmic decision rationales²³.

Responsibility and accountability in algorithmic decision-making also warrant careful consideration. Employers cannot simply defer liability to the software provider. An employer that outsources recruitment algorithms to a third-party vendor might mistakenly treat contractual compliance as sufficient protection against discrimination risks associated with the resulting employment decisions. However, from a legal perspective, ultimate responsibility for employment decisions remains with the employer. Courts and regulatory authorities are increasingly likely to hold employers accountable for discriminatory outcomes, regardless of whether the underlying tool was developed externally²⁴.

Practical strategies for mitigating algorithmic bias have emerged through both research and industry practice. One approach is to incorporate systematic auditing of models to identify disparate impacts before deployment. As an illustrative compliance measure, an HR team could review recruitment scores quarterly to assess possible disadvantages affecting underrepresented groups. Where a review identifies disparities, changes to relevant criteria or variable weightings could be considered and assessed against job requirements and applicable equality standards²⁵.

Another key strategy involves integrating human oversight into algorithmic decision-making processes. A proposed safeguard is to combine automated recommendations with structured managerial review capable of identifying possible bias and requiring further assessment before adverse decisions are implemented. While this does not eliminate the potential for discrimination, it introduces a layer of

²² Sandra Wachter, Brent Mittelstadt, Luciano Floridi, *op. cit.*, pp. 76–99.

²³ Valerio De Stefano, *op. cit.*, pp. 15–34.

²⁴ Reuben Binns, (2018), „Fairness in Machine Learning: Lessons from Political Philosophy”, *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, vol. 81, 2018: 149–159, <https://proceedings.mlr.press/v81/binns18a.html>, accessed on 9 February 2026.

²⁵ Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé, III, Miro Dudik, Hanna Wallach, „Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?”. In Stephen Brewster, Geraldine Fitzpatrick, Anna Cox, Vassilis Kostakos (eds.), *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, ACM, Paper No.: 600, pp. 1 - 16, <https://doi.org/10.1145/3290605.3300830>.

accountability that aligns with principles of substantive equality and due process. Importantly, such oversight must be informed, systematic, and empowered to act on detected biases; tokenistic review processes are insufficient²⁶.

A hypothetical example further illustrates the interaction between algorithmic logic and workplace norms. Suppose that a company's performance tracking system flags employees as underperforming on the basis of task completion times and online activity. Employees engaged in complex problem-solving or collaborative projects might receive lower scores because their work does not fit the selected quantitative criteria. A review could then assess whether those indicators capture relevant job requirements and whether qualitative measures are needed, illustrating the role of contextual assessment and continuing monitoring in the proposed evaluation framework²⁷.

These illustrative scenarios indicate that the legal assessment of algorithmic decision-making cannot be confined to abstract principles. It requires an understanding of both the operational mechanics of algorithms and the organizational context in which they function. Legal scholars and practitioners must therefore engage with practical realities, recognizing how historical inequities are embedded in data and how model outputs translate into actionable consequences for workers.

Finally, it is worth noting that algorithmic decision-making does not inevitably lead to discrimination. In many cases, these systems can enhance objectivity, consistency, and efficiency. The normative analysis allows for the possibility that, when properly designed and monitored, algorithmic tools may reduce human biases in areas such as scheduling, standardized task allocation, and preliminary screening. The key lies in ensuring transparency, establishing accountability, and continuously evaluating outcomes against non-discrimination standards²⁸.

In conclusion, algorithmic decision-making in employment presents a complex interplay of bias, evaluation, transparency, and responsibility. Recruitment and performance evaluation systems demonstrate how ostensibly neutral criteria can produce unequal outcomes, raising questions of indirect discrimination and substantive equality. Transparency and explainability are central to enabling effective enforcement of legal norms, while responsibility ultimately rests with the employer, regardless of technological intermediaries. Practical strategies such as auditing, human oversight, and contextual adjustment of evaluation criteria are essential to mitigate risks and align algorithmic practices with the principles of equality and non-discrimination established in international labour law.

5. Conclusions and Regulatory Implications

²⁶ Aislinn Kelly-Lyth, „Algorithmic discrimination at work”, *European Labour Law Journal*, Vol. 14(2), 2023: 152–171 DOI: 10.1177/20319525231167300.

²⁷ Ifeoma Ajunwa, Kate Crawford and Jason Schultz, „Limitless Worker Surveillance”, *California Law Review*, Vol. 105, No. 3, June 2017: 735–776.

²⁸ Brent Mittelstadt, „Principles Alone Cannot Guarantee Ethical AI”, *Nature Machine Intelligence*, vol. 1, 2019: 501–507. <https://doi.org/10.1038/s42256-019-0114-4>.

The analysis conducted throughout this research highlights that the integration of algorithmic systems into employment processes does not create entirely new legal problems, but it does significantly magnify existing challenges in ensuring equality and non-discrimination. The doctrinal assessment suggests that current legal frameworks — both international and regional — provide a solid foundation for addressing discrimination, but they are not fully adapted to the nuances introduced by algorithmic decision-making.

First, the doctrinal analysis indicates that the core principles of equality and non-discrimination, as articulated in ILO conventions and UN treaties, remain highly relevant. These frameworks allow for the recognition of indirect discrimination and provide mechanisms for shifting the burden of proof when disparities are evident²⁹. However, applying these principles to AI-mediated decisions is complicated by the opacity of algorithms and the reliance on large, historical datasets that may carry embedded biases. Unlike traditional human decision-making, the outputs of automated systems are often difficult to interrogate, making it challenging to establish whether observed disparities result from legitimate criteria or from discriminatory patterns embedded in data.

A hypothetical organization that relies on algorithmic tools without adequate auditing or evaluation may face an increased risk of unintended discrimination and potential legal liability. For example, a recruitment algorithm used to screen large volumes of applications could reproduce historical workforce imbalances if its training data embeds those patterns. Similarly, a performance evaluation system based solely on quantitative metrics may overlook qualitative aspects of work and disadvantage particular groups. These scenarios support the normative conclusion that formal policies should be accompanied by assessment of outcomes, attention to possible group disadvantages, and reasoned consideration of whether the selected criteria are relevant and objectively justified in the employment context.

Another conclusion of the doctrinal analysis is that transparency and accountability are necessary safeguards against discriminatory decisions. Where employers lack insight into externally developed algorithms and employees cannot obtain meaningful explanations, an information asymmetry may hinder the enforcement of non-discrimination norms. My analysis suggests that regulatory frameworks need to provide guidance on transparency requirements, including access to decision criteria and explanation of outcomes, in a manner that is understandable and actionable for both employees and regulators³⁰.

In addition, the legal analysis indicates that responsibility cannot be fully delegated to technology providers. Even when algorithmic tools are developed externally, the employer remains the ultimate decision-maker and is accountable for the results. This principle aligns with existing labour law doctrines but must be emphasized in the context of automated systems. The use of commercially available AI tools

²⁹ International Labour Organization (1958), Convention No. 111 concerning Discrimination in Respect of Employment and Occupation, ILO Publications, pp.: 1–12.

³⁰ Sandra Wachter, Brent Mittelstadt, Luciano Floridi, *op. cit.*, 76–99.

therefore cannot, by itself, establish that an employer has discharged its obligations. Compliance must be assessed against the applicable legal requirements and the decisions for which the employer is responsible.

On the basis of this doctrinal analysis, I propose several directions for legal and policy development. First, existing non-discrimination frameworks should be clarified and adapted to account explicitly for algorithmic decision-making. This could include requirements for periodic audits, documentation of decision criteria, and mechanisms for employees to challenge outcomes. In my view, such measures would align legal obligations with practical realities, making it possible to hold employers accountable while still allowing the use of technological tools for legitimate purposes.

Second, guidance on assessing indirect discrimination in algorithmic systems should be developed. Traditional methods rely on statistical analysis and comparison of outcomes across groups, which may assist in identifying disparities that require further legal assessment. The analysis supports considering whether introducing mandatory impact assessments for AI-based employment tools could significantly improve compliance and reduce unintentional disparities. These assessments should evaluate both recruitment and performance evaluation systems, focusing on whether outcomes disproportionately affect protected groups³¹.

Third, the doctrinal assessment supports meaningful human oversight as a safeguard in algorithmic decision-making. Automated outputs should inform, not replace, human judgment. Combining algorithmic recommendations with structured managerial review could help identify possible bias, provided that reviewers have adequate information, authority, and procedures for reconsidering problematic outcomes. Therefore, regulatory guidance should encourage human-in-the-loop systems and ensure that oversight is meaningful and accountable, rather than perfunctory.

Finally, ongoing monitoring and adaptation are critical. Algorithmic systems evolve, and historical biases may shift or emerge in new ways over time. Periodic review procedures, through which recruitment and evaluation data are examined for disparities, could help employers identify emerging risks and assess compliance with non-discrimination standards. Their effectiveness would require evaluation in the relevant organizational and legal context. This iterative approach aligns with the principles of substantive equality, ensuring that legal protections are not static but responsive to changing workplace realities.

In conclusion, the doctrinal analysis supports several key assessments. Existing legal frameworks provide a strong foundation but require adaptation and practical enforcement mechanisms to address algorithmic decision-making effectively. The primary challenges are the opacity of systems, the embedding of historical biases in data, and the asymmetry of information between employers and employees. Effective compliance depends on transparency, accountability, and human oversight. Regulatory developments should focus on clarifying responsibilities, establishing requirements for

³¹ Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé, III, Miro Dudik, Hanna Wallach, *op. cit.*, pp. 1 - 16.

auditing and monitoring, and providing guidance on interpreting indirect discrimination in automated contexts.

Overall, these conclusions reflect a pragmatic approach: AI and algorithmic systems can offer efficiency and consistency in employment processes, but they must operate within a framework that prioritizes fairness and equality. The illustrative scenarios developed in this article underscore the importance of not treating technology as inherently neutral, and of recognizing that legal and ethical responsibilities remain with the organizations that deploy these tools. Future research should examine longitudinal outcomes of algorithmic employment practices and explore the effectiveness of auditing and oversight mechanisms in reducing disparities. Such work will be crucial to ensuring that technological innovation in the workplace complements, rather than undermines, the principles of equality and non-discrimination.

Bibliography

1. Ajunwa, Ifeoma, Kate Crawford and Jason Schultz, „Limitless Worker Surveillance”, *California Law Review*, Vol. 105, No. 3, June 2017: 735-776.
2. Barocas, Solon & Andrew D. Selbst, „Big Data’s Disparate Impact”, *California Law Review*, vol. 104, issue 3, 2016: 671–732, <https://doi.org/10.15779/Z38BG31>.
3. Bell, Myrtle P. and Daphne P. Berry, „Viewing Diversity Through Different Lenses: Avoiding a Few Blind Spots”, *Academy of Management Perspectives*, Vol. 21, No. 4, 2007: 21-25, [https://doi.org/10.5465/ amp.2007.27895336](https://doi.org/10.5465/amp.2007.27895336).
4. Binns, Reuben (2018), „Fairness in Machine Learning: Lessons from Political Philosophy”, *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, vol. 81, 2018: 149–159, <https://proceedings.mlr.press/v81/binns18a.html>, accessed on 9 February 2026.
5. Dastin, Jeffrey, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women*, Reuters, 2018.
6. European Union (2000), Council Directive 2000/78/EC establishing a general framework for equal treatment in employment and occupation, OJ L 303, 2.12.2000.
7. Fredman, Sandra, *Discrimination Law*, Oxford University Press, 2011.
8. Hepple, Bob, *Equality: The Legal Framework*, Second edition. Oxford: Hart Publishing, 2014.
9. Holstein, Kenneth, Jennifer Wortman Vaughan, Hal Daumé, III, Miro Dudik & Hanna Wallach, „Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?”. In Brewster, Stephen, Geraldine Fitzpatrick, Anna Cox & Vassilis Kostakos (eds.), *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, ACM, Paper No.: 600, pp. 1 - 16, <https://doi.org/10.1145/3290605.3300830>.
10. International Labour Organization (1958), *Convention No. 111 concerning Discrimination in Respect of Employment and Occupation*, ILO Publications, <https://www.ilo.org/media/27816/download>, accessed on 9 February 2026.
11. Kelly-Lyth, Aislinn, „Algorithmic discrimination at work”, *European Labour Law Journal*, Vol. 14(2), 2023: 152–171 DOI: 10.1177/20319525231167300.
12. Kroll, Joshua A., Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, and Harlan Yu, „Accountable Algorithms”, *University of Pennsylvania Law Review*, vol. 165, issue 3, 2016: 633–705.

13. Mittelstadt, Brent, „Principles Alone Cannot Guarantee Ethical AI”, *Nature Machine Intelligence*, vol. 1, 2019: 501–507. <https://doi.org/10.1038/s42256-019-0114-4>.
14. Raghavan, Manish, Solon Barocas, Jon Kleinberg & Karen Levy, „Mitigating Bias in Algorithmic Hiring: Evaluations and Lessons”. In Hildebrandt, Mireille, Carlos Castillo, Elisa Celis, Salvatore Ruggieri, Linnet Taylor & Gabriela Zanfir-Fortuna (eds.), *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ACM, 2020, pp. 469–481, <https://doi.org/10.1145/3351095.3372828>.
15. Rosenblat, Alex & Luke Stark, „Algorithmic Labor and Information Asymmetries: A Case Study of Uber’s Drivers”, *International Journal of Communication*, vol. 10, 2016: 3758–3784.
16. Stefano, Valerio De, „Negotiating the Algorithm: Automation, Artificial Intelligence and Labour Protection”, *International Labour Organizations*, Employment Working Paper No. 246, 2019: 15–34, <https://www.ilo.org/publications/negotiating-algorithm%E2%80%9Dautomation-artificial-intelligence-and-labour>, accessed on 9 February 2026.
17. United Nations (1979), *Convention on the Elimination of All Forms of Discrimination Against Women*, UN Publications, <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-elimination-all-forms-discrimination-against-women>, accessed on 9 February 2026.
18. Wachter, Sandra, Brent Mittelstadt & Luciano Floridi, „Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, *International Data Privacy Law*, Volume 7, Issue 2, May 2017: 76–99, <https://doi.org/10.1093/idpl/ix005>.