Artificial Intelligence in Employment Decision-Making: Legal Challenges and Implications

Lecturer **Dana VOLOSEVICI**[1]

Abstract

This study explores the legal challenges arising from the use of artificial intelligence in employment decision-making, with particular attention to its implications for employees' rights and managerial accountability. The primary objective is to evaluate whether existing legal frameworks, most notably the General Data Protection Regulation and the Artificial Intelligence Act, offer sufficient protection against the risks associated with automated and algorithmically informed decisions in the workplace. Employing a qualitative methodology, the research is grounded in doctrinal analysis of relevant European legal instruments, supplemented by a review of academic literature in labour law, data protection, and algorithmic governance. The study adopts an interdisciplinary perspective, combining legal analysis with insights from organisational psychology and data science. The findings underscore key concerns, including the risk of indirect discrimination, the opacity of algorithmic decision-making processes, and the potential dilution of managerial responsibility. In response, the paper recommends a series of organisational measures such as targeted training, structured collective bargaining on AI deployment, and the adoption of a sustainable, rights-oriented approach to managing the workforce. The study concludes that a multidimensional governance model is essential to ensure that technological innovation remains aligned with the protection of workers' fundamental rights and the principles of democratic workplace governance.

Keywords: *artificial intelligence; employment decision-making; labour law; data protection; employee participation.*

JEL Classification: K24, K31

DOI: https://doi.org/10.62768/ADJURIS/2025/3/02

Please cite this article as:
Volosevici, Dana, „Artificial Intelligence in Employment Decision-Making: Legal Challenges and Implications", in Devetzis, Dimitrios, Dana Volosevici & Leonidas Sotiropoulos (eds.), *Digital Lawscapes: Artificial Intelligence, Cybersecurity and the New European Order*, ADJURIS – International Academic Publisher, Bucharest, Paris, Calgary, 2025, p. 25-41.

1. Introduction

Artificial Intelligence has become increasingly embedded in employment

[1] Dana Volosevici - Petroleum Gas University of Ploiesti, Romania, ORCID: 0009-0000-9109-4679, dana.volosevici@upg-ploiesti.ro.

decision-making processes due to its potential to enhance efficiency, objectivity, and predictive accuracy. In contemporary labour markets, employers face complex demands to evaluate vast amounts of candidate and employee data swiftly and consistently. AI systems offer algorithmic solutions that can process such data in ways that human decision-makers may find time-consuming or prone to cognitive bias. As noted in legal scholarship, this technological shift corresponds with a broader move toward data-driven governance in employment practices, which is reshaping traditional managerial prerogatives[2].

The integration of AI into employment settings also involves the continuous and large-scale processing of personal data, using a huge capacity of processing. This means that the volume, granularity, and frequency of data collection significantly exceed traditional methods of employee supervision, enabling employers to draw complex inferences about individual behaviour, productivity, and even psychological traits. The scale and sophistication of such data processing raise important concerns regarding the proportionality and necessity of surveillance practices, particularly when viewed through the lens of data protection principles enshrined in Article 5 of the General Data Protection Regulation[3].

Moreover, AI monitoring may be extended beyond standard working hours. For example, remote work monitoring software may continue to track device usage or location data after official working time has ended, and algorithmic scheduling tools may access personal calendars or communications to optimise future workflows. This blurs the boundary between professional and private life, potentially infringing upon the worker's right to disconnect and to a private life under Article 8 of the European Convention on Human Rights and Article 7 of the Charter of Fundamental Rights of the European Union. In this regard, the continuous reach of AI tools into employees' time and personal space not only challenges traditional notions of managerial control but also necessitates a recalibration of labour law protections in light of emerging technologies.

Against this background, the legal framework has had to evolve to address the risks of hyper-surveillance and to ensure that digital monitoring does not erode the dignity and autonomy of workers. Initially, the GDPR established a comprehensive framework aimed at safeguarding personal data and reinforcing fundamental rights in the digital age. Building upon this foundation, AI Act introduced specific obligations for high-risk AI systems, including those used in employment contexts, thereby ensuring that the development and deployment of

---

[2] Valerio De Stefano, *Negotiating the Algorithm: Automation, Artificial Intelligence and Labour Protection*, „Comparative Labor Law & Policy Journal" 41(1), 2019, p. 15–46.

[3] Paul De Hert & Vagelis Papakonstantinou, *The new General Data Protection Regulation: Still a sound system for the protection of individuals?* "Computer Law & Security Review", 32(2), 2016, p. 179–194; Antonie Aloisi & Elena Gramano, *Artificial Intelligence is Watching You at Work: Digital Surveillance, Employee Monitoring, and Regulatory Issues in the EU Context*, "Comparative Labor Law & Policy Journal", 41(1), 2019, p. 95–126.

AI technologies respect the fundamental rights and freedoms of the human person.

More specifically, GDPR introduced a series of foundational principles designed to enhance individuals' control over their personal data and to ensure that data processing activities are carried out lawfully, fairly, and transparently. Among these, the principle of transparency — enshrined in Article 5(1)(a) GDPR — requires that data subjects be clearly informed about how their personal data is collected, used, and for what purposes. Transparency is further reinforced by Articles 12 to 14 GDPR, which impose obligations on data controllers to provide information in a concise, intelligible, and easily accessible form, using clear and plain language. This principle is especially crucial in the context of automated decision-making, where individuals may otherwise lack insight into how algorithmic processes influence outcomes that affect them directly[4]. In addition to transparency, the GDPR codifies the principles of data minimisation and purpose limitation (Article 5(1)(c) and (b), respectively), which require that only personal data that is adequate, relevant, and limited to what is necessary in relation to the purposes for which it is processed be collected and used.

Crucially, Article 22 GDPR establishes the right not to be subject to a decision based solely on automated processing, including profiling, where such a decision produces legal effects or significantly affects the individual. This provision reflects a broader concern about protecting human dignity and autonomy in the face of algorithmic decision-making. The right under Article 22 is particularly relevant in employment contexts, where automated decisions about hiring, task allocation, or performance evaluation may profoundly impact workers' professional and personal lives.

The AI Act classifies as high-risk AI systems both those intended to be used for the recruitment or selection of natural persons — particularly for the targeted placement of job advertisements, the analysis and filtering of job applications, and the evaluation of candidates — and those intended to be used for making decisions affecting the terms of employment relationships. This includes decisions relating to promotion, termination of contractual employment relationships, the allocation of tasks based on individual behaviour or personal traits or characteristics, as well as the monitoring and evaluation of the performance and conduct of individuals engaged in such relationships.

As a consequence, such AI systems must comply with a series of requirements expressly laid down in Articles 9 to 15 of the AI Act, including the obligation to ensure human oversight. This requirement, as defined in Article 14, entails the implementation of appropriate human involvement during the operation of the AI system, with the aim of preventing or minimising risks to health, safety,

[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, p. 76–99, https://doi.org/10.1093/idpl/ipx005.

or fundamental rights. Human oversight must be effective, meaning that individuals designated to supervise the system are able to understand its functioning, detect anomalies or inappropriate outcomes, and intervene when necessary to override or disable the system's operations.

The purpose of this study is to examine whether, despite the existing legislative provisions, risks related to the use of AI systems in employment-related decision-making still persist. The analysis will focus on three main aspects. First, it will explore the nature and specificity of managerial decision-making when such decisions are informed or influenced by the outcomes generated by AI systems. Secondly, the study will examine the operational characteristics of AI systems, with particular attention to those aspects of their functioning that may impact employees' rights. Finally, it will seek to identify potential organisational measures that could be implemented to mitigate the risk of infringing employees' rights, thereby promoting the responsible use of AI in the workplace.

2. The Specificity of AI-Informed Managerial Decisions in Employment

The increasing integration of artificial intelligence into human resources management has begun to reshape the nature and legitimacy of managerial decision-making. While AI systems promise efficiency, standardisation, and data-driven objectivity, their use in employment contexts raises complex questions about human agency, legal accountability, and the boundaries of managerial discretion.

a) Human Versus AI Reasoning in Employment Decision-Making. One of the foundational issues that merits legal and interdisciplinary examination is the distinction between human reasoning and AI-based reasoning in employment-related decisions. Traditionally, employment relations have evolved through decisions made by human agents, who increasingly rely on sophisticated data analytics and predictive tools. However, the decisional logic of human actors remains qualitatively different from that of algorithmic systems. Human managerial decisions are typically shaped by a complex interplay of experience, contextual judgment, organisational values, and often tacit knowledge accumulated over time[5]. Although such decisions may be partially subjective, they also allow for adaptability to specific contexts, such as the organisational climate, the socio-economic environment, or the individual characteristics of workers. In this sense, human discretion accommodates the flexibility needed to uphold key legal and ethical values in labour relations, including fairness, proportionality, and individualisation.

[5] Michael J.R. Butler, Holly L.R. O'Broin, Nick Lee and Carl Senior, *How Organisational Cognitive Neuroscience Can Deepen Understanding of Managerial Decision-Making: A Review of the Recent Literature and Future Directions,* „International Journal of Management Reviews", Vol. 18, 2016, p. 542–559.

By contrast, AI systems reason through computational logic, deriving statistical correlations from historical data and optimising decisions based on predefined objectives, such as efficiency or productivity. These systems do not interpret context or moral values; they function by identifying patterns and applying learned associations, absent a normative framework[6]. Consequently, while AI-generated outcomes may appear more precise or objective, they lack the interpretive flexibility that human reasoning provides—particularly relevant in labour law, where many decisions require case-by-case assessment[7]. This divergence creates a structural asymmetry in the decision-making process. Human reasoning allows for ethical reflection and legal proportionality, while AI reasoning, even when technically robust, may unintentionally reproduce biases embedded in training data[8]. For instance, recruitment algorithms trained on past hiring decisions may internalise and perpetuate historical patterns of gender or racial discrimination, even if unintentionally[9].

At a legal level, this distinction becomes particularly salient. Employment decisions often entail significant consequences, such as promotion, disciplinary sanctions, or termination, and must comply with substantive and procedural labour standards. These include the right to equal treatment, due process, and protection against unfair dismissal, as recognised in both national labour laws and EU law (Article 30 of the Charter of Fundamental Rights of the European Union; Council Directive 2000/78/EC). The application of generalised algorithmic logic to such decisions can undermine the required individualised assessment, risking violations of anti-discrimination law or procedural safeguards.

Therefore, the use of algorithmic management tools in the workplace carries the risk of creating a false perception of neutrality, whereby decisions appear objective but are, in fact, removed from the legal and ethical principles that have traditionally underpinned human-centred employment relationships. This disconnect is further intensified by the inherent limitations of AI systems, which are unable to account for non-quantifiable factors such as empathy, moral reasoning,

---

[6] Adams-Prassl, Jeremias, *What if Your Boss Was an Algorithm? The Rise of Artificial Intelligence at Work*, „Comparative Labor Law & Policy Journal", Vol, 41(1), 2019, Available at SSRN: https://ssrn.com/abstract=3661151.

[7] Valerio De Stefano and Mathias Wouters, *AI and Digital Tools in Workplace Management Evaluation: An Assessment of the EU's Legal Framework*, European Parliamentary Research Services, Scientific Foresight Unit (PE 729.516), 2022.

[8] Sara Baiocco, Enrique Fernández-Macías, Uma Rani, and Annarosa Pesole, *The Algorithmic Management of Work and its Implications in Different Contexts,* European Commission, 2022; Katherine C Kellogg, Melissa A Valentine and Angèle Christin, *Algorithms at Work: The New Contested Terrain of Control,* „Academy of Management Annals", Vol. 14, 2020, p. 366; Mohammad Hossein Jarrahi, *Artificial Intelligence and the Future of Work: Human-AI Symbiosis in Organisational Decision Making*, „Business Horizons", Vol. 61(4), 2018, p. 577–586 https://doi.org/10.1016/j.bushor.2018.03.007.

[9] Miriam Kullmann, *Platform Work, Algorithmic Decision-Making, and EU Gender Equality Law,* „International Journal of Comparative Labour Law and Industrial Relations", Vol 34(1), 2018, p. 1–21.

or evolving organisational norms, elements that remain essential for fair and lawful decision-making in human resource management.

b) Automation Bias and Managerial Deference. Another layer of complexity in AI-informed managerial decision-making stems from documented psychological tendencies that shape how humans interact with automated systems. A central concept in this regard is automation bias - the cognitive inclination to trust and over-rely on decisions or suggestions made by automated systems, even in the face of contradictory or incorrect outputs. Research has shown that individuals tend to defer to automated recommendations because these systems are perceived as faster, more accurate, or more neutral than human judgment[10]. This bias is particularly pronounced in high-stakes or high-pressure environments, such as employment decision-making, where managerial accountability, time constraints, and organisational expectations converge. In such contexts, managers may perceive AI systems as more reliable, consistent, or legally defensible than their own subjective judgment, especially when the outputs are presented with high confidence or a veneer of objectivity.

From a legal and organisational perspective, this dynamic has significant implications. Article 22 of the GDPR grants individuals the right not to be subject to decisions based solely on automated processing, including profiling, that produces legal effects or significantly affects them. However, this right assumes the existence of genuine and informed human oversight. If managerial review is reduced to a mere formal endorsement, driven by automation bias and limited understanding, then the safeguard becomes functionally ineffective. Similarly, Article 14 of the AI Act requires that high-risk AI systems be subject to effective human oversight. However, the regulation does not explicitly define the cognitive or organisational preconditions needed for oversight to be meaningful. Research suggests that effective oversight is not merely procedural but requires a critical mental model of how the system works, as well as the confidence and autonomy to override its outputs when necessary[11].

c) Responsibility and Accountability in AI-Informed Decisions. Despite

[10] Goddard K, Roudsari A, Wyatt JC, *Automation bias: a systematic review of frequency, effect mediators, and mitigators*, „Journal of the American Medical Informatics Association", 19(1), 2012, p.121-127. doi: 10.1136/amiajnl-2011-000089; Hannah Ruschemeier and Lukas J. Hondrich, *Automation Bias in Public Administration – An Interdisciplinary Perspective from Law and Psychology*, „Government Information Quarterly", 41(3), 2024, 101953, https://doi.org/10.1016/ j. giq.2024.101953.

[11] Linda J. Skitka, Kathleen Mosier and Mark D. Burdick, *Accountability and Automation Bias*, "International Journal of Human-Computer Studies", 52(4), 2000, p. 701–717, https://doi.org/10.1006/ijhc.1999.0349; Eugenio Alberdi, Lorenzo Strigini, Andrey A. Povyakalo, Peter Ayton, *Why Are People's Decisions Sometimes Worse with Computer Support?*. In: Buth, B., Rabe, G., Seyfarth, T. (eds) „Computer Safety, Reliability, and Security". SAFECOMP 2009. Lecture Notes in Computer Science, vol 5775. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04468-7_3.

the growing reliance on AI systems in employment-related decision-making, legal responsibility remains with human actors, typically the employer and, in some instances, individual managers. Under established principles of employment law, decision-makers must be able to justify their decisions —whether those decisions are adverse, such as disciplinary action or dismissal, or favourable, such as promotion or the award of a performance bonus. The obligation to justify decisions stems from the requirement to demonstrate their legality, fairness, and proportionality, particularly in contexts where such decisions may significantly affect the employee's rights and professional trajectory.

However, when AI systems, especially those based on machine learning and characterised by limited interpretability, are involved in the decision-making process, managers may find themselves unable to explain the rationale behind a particular recommendation. The technical complexity and opacity of such systems, often referred to as "black box" models, mean that key elements of the decision may originate from a source external to the manager's control, yet the manager remains compelled, or feels pressured, to rely on those outputs.

This situation raises important questions regarding the allocation of legal responsibility. In other regulatory domains, such as product liability, one might consider the role of the AI system developer or manufacturer. However, in the field of labour law, responsibility for employment decisions continues to rest with the employer, including responsibility for the choice and implementation of AI tools used within the organisation. Given the inherently asymmetrical nature of the employment relationship, where the employer holds structural and legal power over the employee, it is the employer who must bear responsibility for organising work, including decisions informed by AI[12]. Accordingly, even when decisions are mediated or informed by algorithmic systems, the notion of a "responsibility gap"[13], where no actor is held accountable, is not admissible within the framework of labour law. The employer must ensure that the systems deployed in human resource management are compatible with legal standards and that adequate mechanisms for oversight and contestability are in place.

Nevertheless, when decisions are based on highly complex AI systems, even where these systems meet the formal requirements of the AI Act, employers may face significant challenges in demonstrating that such decisions are proportionate, lawful, and non-discriminatory. This difficulty may undermine the employer's ability to defend its actions before labour courts or data protection authorities, especially in jurisdictions where the burden of proof in discrimination cases lies with the employer.

---

[12] Valerio De Stefano and Simon Taes, *Algorithmic Management and Collective Bargaining,* „Transfer: European Review of Labour and Research", 29(1), 2022, https://doi.org/10.1177/10242589221141055.

[13] Andreas Matthias, *The responsibility gap: Ascribing responsibility for the actions of learning automata*, „Ethics and Information Technology",6, 2004, p. 175–183. https://doi.org/10.1007/s10676-004-3422-1.

Furthermore, the diffusion of decision-making agency between algorithmic systems and human actors fundamentally challenges the principle of personal accountability that underlies managerial functions. Employment law has traditionally operated on the assumption that decision-makers act with deliberation, autonomy, and legal awareness. When managers are unable to interrogate, override, or meaningfully interpret algorithmic outputs, this assumption no longer holds. As a result, the integrity of decision-making processes, and the legal accountability attached to them, risks being substantially weakened, unless clear governance frameworks and robust safeguards are instituted to preserve human oversight and legal responsibility.

3. Algorithmic Design and the Protection of Workers' Rights

A central concern in the governance of artificial intelligence is the distinction between opaque and transparent AI systems, a distinction with significant implications for the protection of fundamental rights, particularly in the employment context. These categories refer not only to the technical architecture of AI models but also to their legal relevance in determining accountability, explainability, and compliance with data protection and non-discrimination obligations.

Opaque AI systems, often referred to as "black box" models, are characterised by their limited interpretability. Such systems, typically based on advanced machine learning techniques such as neural networks, generate outputs without providing a comprehensible rationale that could be understood by non-expert users, including employers or affected individuals. This opacity inhibits the ability to interrogate, justify, or challenge decisions, thereby undermining core principles of transparency, accountability, and procedural fairness[14]. Conversely, transparent AI systems are either inherently interpretable, such as those employing decision trees or rule-based logic or are supplemented by explainability tools that enable users to understand the logic behind outputs. Transparent systems are more easily aligned with legal standards that require decisions to be explainable, contestable, and reviewable, particularly when they have significant effects on individuals, as is often the case in employment. The AI Act directly addresses these concerns by imposing a series of obligations on high-risk AI systems. Under Article 13 AI Act, providers of high-risk AI systems must ensure that the system is designed and developed in a manner that allows for an appropriate level of transparency. This includes the obligation to inform users about the system's intended purpose, its decision-making logic (where possible), and any limitations that may affect its reliability or fairness.

However, even in the case of transparent AI systems, the decision-mak-

[14] Amedeo Santosuosso and Giovanni Sartor (2024), *Decidere con l'IA: Intelligenze artificiali e naturali nel diritto*, Il Mulino, p. 75.

ing process remains inherently complex. It typically involves multiple stages, including data collection, data preprocessing, algorithmic analysis, and the generation of recommendations. These phases, which are not under the direct control of the system user, must ensure accuracy, relevance, and legal compliance from the system's initial design phase and throughout its operational lifecycle[15].

For example, the data collection phase involves the systematic aggregation of relevant data from multiple sources, both within the organisation and external to it. These sources include structured data, such as employee attendance records, performance metrics, and financial transactions, as well as unstructured data, such as emails, feedback reports, or written evaluations. AI-driven recruitment platforms, for example, utilize resumes, online job applications, interview transcripts, and previous hiring decisions to assess candidate suitability[16]. Similarly, compliance monitoring systems integrate data from internal policy documents, regulatory frameworks, and past legal rulings to evaluate employees' adherence to workplace regulations. A significant aspect of data collection may involve real-time data streaming, where AI continuously receives updates from workplace monitoring tools, biometric attendance systems, customer feedback surveys, and other dynamic sources[17].

For algorithmic analysis to be statistically valid and unbiased, it must process a large and representative dataset that includes heterogeneous sources and temporally distributed data points. The inclusion of varied sources ensures that the algorithm captures diverse patterns and avoids biases introduced by homogeneous data. Processing data from different time periods accounts for temporal shifts, preventing models from becoming outdated or overfitting to short-term trends. These factors are critical for ensuring algorithmic fairness, predictive accuracy, and robustness in dynamic workplace environments[18]. For statisticians, data must be valid and reliable, meaning that the input data should have a verified level of accuracy to reduce the likelihood of Type I errors (false positives). To obtain such data, it is mandatory to ensure rigorous data collection processes.

As data complexity increases, so does the risk of violating legal provisions designed to protect the rights of data subjects, particularly employees in this context. This situation creates an inherent tension between two legitimate interests: statisticians, who require large-scale, detailed data to develop statistically

---

[15] Bruno Lepri, Nuria Oliver, Emmanuel Letouzé, Alex Pentland, Patrick Vink, *Fair, Transparent, and Accountable Algorithmic Decision-making Processes*. „Philosophy and Technology", 31, 2018, p. 611–627. https://doi.org/10.1007/s13347-017-0279-x; Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi, *The ethics of algorithms: Mapping the debate*, "Big Data & Society", 3(2), 2016, https://doi.org/10.1177/2053951716679 679.

[16] Aaron Rieke and Miranda Bogen, *Help wanted: An examination of hiring algorithms, equity, and bias,* Labor and Employment. Upturn, 2018, December 10. https://www.upturn.org/reports/2018/hiring-algorithms/, accessed 16 March 2025.

[17] Ifeoma Ajunwa, *The quantified worker: Law and technology in the modern workplace* Cambridge University Press, 2023. https://doi.org/10.1017/9781316888681.

[18] Sandra Wachter, Brent Mittelstadt, Luciano Floridi, *op.cit.*

robust models, and legal professionals, who work to ensure that data processing remains transparent, proportional, and compliant with privacy and data protection regulations. While statisticians prioritize data completeness and granularity to enhance the accuracy and predictive power of their models, legal experts emphasize the need to minimize data collection, limit processing to strictly necessary information, and ensure fairness and accountability in its use.

The data preprocessing phase, which transforms raw data into a usable format, could also raise legal concerns. Data cleaning, which involves correcting errors or filling in missing values, may (unintentionally) alter employee records, raising issues of data integrity and legal accountability, especially in sensitive areas as disciplinary or performance-related disputes. Normalization and standardization, used to harmonize data formats, can facilitate accurate analysis but may inadvertently perpetuate systemic biases, especially if contextual differences (e.g., regional pay scales) are not properly accounted for. In feature selection and engineering, the identification of variables for algorithmic decision-making must be carefully scrutinized. Including or proxying protected characteristics (e.g., gender or ethnicity) risks indirect discrimination, unless justified under specific legal frameworks such as affirmative action policies. Employers must ensure that these design choices comply with both anti-discrimination law and the AI Act's data governance requirements (Article 10 AI Act). However, fulfilling these obligations presupposes that employers possess sufficient technical and legal knowledge to identify, request, and verify compliance with such requirements. This represents a significant shift in the traditional scope of managerial responsibility, requiring not only legal awareness but also a degree of algorithmic literacy. Moreover, trade unions and employee representatives should also be equipped to engage with these systems meaningfully, particularly in their roles concerning consultation, co-determination, and the protection of workers' rights.

Once data has been preprocessed, AI systems apply algorithmic models to identify patterns and classify behaviours. While such techniques are intended to increase the efficiency and consistency of managerial decision-making, they also introduce significant legal risks, particularly in relation to discrimination, data protection, and procedural fairness. A brief review of several common AI techniques helps to illustrate the potential risks at stake. Predictive analytics, used to anticipate outcomes such as employee attrition or performance decline, typically relies on historical datasets. However, these datasets may reflect past inequalities or biased decision-making, thereby embedding and perpetuating indirect discrimination against protected groups[19]. Classification models, which assign employees or behaviours into predefined categories (e.g., "low risk," "high risk"), may produce disparate impacts when seemingly neutral input variables correlate

[19] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Publishing, 2016.

with protected characteristics such as age, gender, or ethnicity. This raises concerns regarding algorithmic discrimination, which may infringe the principles of equal treatment and proportionality in employment law.

Anomaly detection systems, often employed in compliance or fraud monitoring, identify deviations from normative behaviour. However, where such models are poorly calibrated or based on incomplete contextual information, they may generate false positives. The resulting actions — such as unwarranted disciplinary investigations — may infringe on workers' reputational rights, procedural guarantees, and the broader principle of fair treatment under both labour law and fundamental rights frameworks.

Following the algorithmic analysis phase, AI systems generate outputs in various forms — such as recommendations, rankings, or alerts — tailored to the specific decision-making context in which they are deployed. While these outputs are intended to enhance managerial efficiency and consistency, they raise significant concerns regarding interpretability, contextual relevance, and legal accountability. These concerns arise primarily from the fact that AI-generated outputs, though often perceived as objective or neutral, may in fact be opaque, offering limited insight into the reasoning behind specific recommendations —thereby increasing the risk of discriminatory outcomes[20], particularly when the underlying data or model design reflects existing societal or institutional biases.

As previously noted, the transition from algorithmic output to human intervention is not without legal risks. Human decision-makers may adjust AI-generated outputs, override recommendations, or request additional information before reaching a final decision. However, this presupposes that the human agent treats the AI system as a standard decision-support tool, engaging with it through critical reasoning rather than passive acceptance. For such oversight to be meaningful, the human operator must possess the requisite technical and legal knowledge to interpret the system's output and must be provided with sufficient information by the AI system itself — such as the underlying rationale, confidence levels, and limitations — to enable an informed and lawful decision.

4. Internal Governance of Workplace AI Systems

As the deployment of AI systems in employment contexts becomes increasingly prevalent, the need for organisational safeguards to prevent the infringement of employees' rights has become a pressing legal and ethical concern. While regulatory frameworks such as the GDPR and AI Act provide general obligations concerning data governance, transparency, and human oversight, these must be complemented by internal organisational measures to ensure compliance and accountability at the operational level.

[20] Jeremias Adams-Prassl, Reuben Binns and Aislinn Kelly-Lyth, *Directly Discriminatory Algorithms*, Modern Law Review, 86(1), 2023, p. 144-175. https://doi.org/10.1111/1468-2230.12 759.

a) AI Literacy in the Workplace. A first and essential organisational measure to ensure the responsible use of AI in the workplace is the education and training of both managers and employees on the structure, functioning, and legal implications of the AI systems in use. This is particularly important given the increasing complexity and opacity of algorithmic decision-making tools, which often function as black boxes and produce outputs that may appear objective but are not easily interpretable.

Managers, in particular, must develop a basic level of AI literacy — the ability to understand how algorithmic systems operate, the types of data they use, the logic of their decision-making processes, and the contexts in which their outputs are reliable or problematic. This knowledge is critical not only for effective oversight but also for complying with legal obligations, such as those laid down in Article 14 of the AI Act, which requires human oversight mechanisms for high-risk systems. Without sufficient training, managers may over-rely on AI outputs, treating them as authoritative even in cases where critical contextual judgment is required.

Workers, likewise, must be informed about the presence and role of AI in decisions affecting their working lives, including areas such as recruitment, task allocation, performance evaluation, and disciplinary action. The GDPR reinforces this obligation through its principles of transparency (Article 5(1)(a)) and the right to be informed (Articles 13 and 14). Workers should be made aware of what data is being collected, how it is processed, and for what purposes it is used. This transparency is not merely procedural but essential to enable workers to contest adverse decisions, exercise their right to explanation (under Article 15 GDPR), and assert their rights under Article 22 GDPR, which prohibits decisions based solely on automated processing that produce significant effects unless specific safeguards are in place.

Moreover, training initiatives should not be limited to isolated sessions or technical documentation. Instead, they should be integrated into broader organisational learning strategies and include interactive elements — such as case studies, simulations, and critical discussions — that foster a culture of accountability and ethical awareness. As argued by De Stefano and Taes[21], embedding these capacities within both managerial practice and workforce participation is essential to prevent the erosion of labour rights in algorithmically managed environments.

b) Worker Participation in the Regulation of Algorithmic Management. A second, and equally essential, organisational safeguard is the integration of collective bargaining and social dialogue into the governance of AI systems used in employment contexts[22]. The deployment of algorithmic tools in the workplace should not be left to unilateral managerial discretion, particularly given

[21] Valerio De Stefano and Simon Taes, *op.cit*.
[22] Idem.

their capacity to affect fundamental rights such as privacy, non-discrimination, and the right to fair working conditions. As emphasised by the European Trade Union Confederation[23] and recognised in broader EU labour law frameworks, worker participation through collective bargaining is a key mechanism to ensure democratic oversight over technological change.

Collective agreements can and should serve as regulatory instruments capable of defining the scope, purpose, and limitations of AI use in the workplace. This includes stipulating the types of decisions for which AI tools may be used, setting clear boundaries around invasive monitoring practices, and determining when human oversight must intervene. By formalising these parameters, collective bargaining helps to operationalise the principles of proportionality and necessity, which underpin both the GDPR and the AI Act.

Moreover, collective agreements can establish procedures for transparency and auditing, particularly for high-risk systems used in areas such as recruitment, task allocation, disciplinary measures, and performance evaluation. These procedural safeguards are essential for ensuring compliance with Article 22 GDPR, which prohibits decisions based solely on automated processing unless specific conditions and safeguards are met, including the right to obtain human intervention and to contest the decision. In this context, algorithmic auditing — an internal or external review of how the AI system functions, how data is processed, and whether outputs are fair and legally compliant — can be embedded into collective agreements as a recurring obligation, not merely a one-time procedural formality.

c) A Risk-Based Approach to AI Governance. Organisations should adopt a risk-based approach to the deployment of AI in the workplace, recognising that not all processes require algorithmic optimisation. Employers ought to limit the use of AI systems to contexts where their benefits are demonstrably proportionate to the potential legal, ethical, and social risks. This approach aligns with the principle of data minimisation under Article 5(1)(c) GDPR, as well as the risk management framework established by Article 9 of the AI Act, which requires providers of high-risk systems to implement continuous, iterative processes to identify, evaluate, and mitigate risks.

This cautious deployment strategy is also consistent with the objectives of Sustainable Development Goal 8 (SDG 8), which promotes decent work and inclusive economic growth. In particular, SDG 8 calls for the protection of labour rights and the promotion of safe and secure working environments for all workers. The indiscriminate or opaque use of AI systems — particularly in hiring, monitoring, or disciplinary decisions — risks undermining these commitments by eroding transparency, fairness, and accountability in the employment relationship. If employees are to be recognised as legitimate stakeholders whose rights

[23] European Trade-Union Confederation (ETUC), *ETUC Resolution Calling for an EU Directive on Algorithmic Systems at Work,* 3 December 2021, https://www.etuc.org/en/document/etuc-resolution-calling-eu-directive-algorithmic-systems-work accessed 17 March 2025.

and dignity are respected within organisational structures, then ethical employment practices must include a responsible approach to the use of AI. Given that the AI Act explicitly classifies employment-related AI systems as high-risk (Annex III, Section 4), it is incumbent upon employers to demonstrate that their use of such systems is not only lawful but also aligned with broader principles of corporate social responsibility and sustainable workplace governance.

To institutionalise ethical and legal oversight, organisations should adopt an internal code of conduct governing the use of AI in employment. Such a code should clearly articulate key principles — transparency, fairness, non-discrimination, and human oversight — and set out internal procedures and responsibilities for ensuring compliance. Beyond policy statements, organisations may establish AI ethics committees or appoint AI compliance officers tasked with reviewing proposed AI systems, evaluating their impact on workers' rights, and overseeing their ongoing use. These internal governance mechanisms can serve not only to mitigate legal liability but also to reinforce organisational accountability and cultivate trust among employees.

5. Discussions

The findings of this paper contribute to the evolving discourse on how artificial intelligence reshapes the nature of decision-making in employment, particularly by influencing or informing managerial judgments. While AI has introduced unprecedented capabilities in terms of data processing, pattern recognition, and predictive analytics, its deployment in employment contexts, especially in high-stakes areas such as recruitment, performance evaluation, and disciplinary decisions, raises complex legal challenges.

At the heart of the debate is the epistemological divergence between human and algorithmic reasoning. Human managers operate within a normative framework that values fairness, proportionality, and contextual judgment, as embedded in both national labour law and EU legal standards. By contrast, AI systems optimise outputs based on statistical correlations in training data, which may be shaped by historical patterns of discrimination or systemic inequalities. This divergence is particularly problematic when algorithmic outputs are perceived as neutral or objective, despite being generated through opaque processes that lack moral or legal reasoning. As a result, there is a risk that AI systems could normalise and obscure discriminatory practices, while simultaneously disempowering human decision-makers who defer to them.

This leads to a second area of concern: the risk of automation bias and the erosion of managerial accountability. Users frequently place undue trust in algorithmic systems, especially when the systems are marketed as tools for enhancing efficiency and fairness. In practice, however, this reliance may result in rubber-stamping AI outputs without proper scrutiny. The AI Act's requirement for effective human oversight (Article 14) and the GDPR's Article 22 safeguards

both assume a level of AI literacy and critical engagement that is not guaranteed in most workplace settings. If decision-makers do not understand how a system functions, or cannot interrogate the logic behind its outputs, they cannot fulfil their legal duties, nor can they meaningfully justify their decisions in the event of litigation.

Another key insight emerging from this study is the essential role of collective bargaining and worker participation in the governance of AI systems in the workplace. As emphasised by the European Trade Union Confederation (ETUC, 2021), the adoption of algorithmic management tools should not occur through unilateral employer decisions. Rather, social dialogue must serve as the framework for democratically negotiating the scope, conditions, and safeguards of AI deployment. One of the most critical areas for negotiation concerns the limitation of AI use, particularly in light of the significant volume of data processing that even transparent AI systems require. The functioning of such systems often entails the collection and analysis of extensive employee-related data, which, in certain cases, results in excessive or disproportionate monitoring practices.

Although AI systems rely on access to large and organisation-specific datasets to deliver relevant outputs, this technical requirement cannot justify the continuous surveillance of workers. The protection of privacy and dignity in the workplace must remain a core principle, and any AI deployment must be scrutinised to ensure that it does not place employees under pervasive or intrusive monitoring regimes.

Furthermore, the discussion has placed the responsible use of AI within the broader context of sustainable and ethical workplace governance. As suggested by the risk-based approach in the AI Act (Article 9) and the principle of data minimisation under the GDPR (Article 5(1)(c)), organisations should not pursue algorithmic optimisation indiscriminately. Instead, they must assess whether AI use is proportionate to the task at hand and whether the legal, ethical, and social risks outweigh the operational benefits. This approach also resonates with Sustainable Development Goal 8 (SDG 8), which advocates for decent work, labour protections, and inclusive economic growth. AI deployment that undermines workers' autonomy, privacy, or procedural rights is incompatible with these broader commitments.

Finally, the discussion affirms that regulation alone is not sufficient. While the GDPR and AI Act establish key principles and obligations, their effectiveness depends on the extent to which organisations internalise and operationalise these norms. This requires not only legal compliance but also a shift in organisational culture, one that values human agency, transparency, and responsibility in the face of increasingly automated infrastructures.

In sum, the findings suggest that a multidimensional governance strategy is needed, one that integrates legal safeguards, technical design, organisational

practices, and participatory mechanisms. Only through such an integrated approach can the transformative potential of AI be harnessed in a way that upholds the fundamental rights and dignity of workers.

Bibliography

1. Adams-Prassl, Jeremias, Binns, Reuben and Kelly-Lyth, Aislinn, *Directly Discriminatory Algorithms*, "Modern Law Review", 86(1), 2023, pp. 144–175. https://doi.org/10.1111/1468-2230.12759.
2. Adams-Prassl, Jeremias, *What if Your Boss Was an Algorithm? The Rise of Artificial Intelligence at Work*, "Comparative Labor Law & Policy Journal", Vol. 41(1), 2019. Available at SSRN: https://ssrn.com/abstract=3661151.
3. Ajunwa, Ifeoma, *The Quantified Worker: Law and Technology in the Modern Workplace*, Cambridge University Press, 2023. https://doi.org/10.1017/9781316888681.
4. Alberdi, Eugenio, Strigini, Lorenzo, Povyakalo, Andrey A. and Ayton, Peter, *Why Are People's Decisions Sometimes Worse with Computer Support?*, in Buth, B., Rabe, G. and Seyfarth, T. (eds), "Computer Safety, Reliability, and Security", SAFECOMP 2009, Lecture Notes in Computer Science, vol. 5775, Springer, Berlin, Heidelberg, 2009. https://doi.org/10.1007/978-3-642-04468-7_3.
5. Aloisi, Antonio and Gramano, Elena, *Artificial Intelligence is Watching You at Work: Digital Surveillance, Employee Monitoring, and Regulatory Issues in the EU Context*, "Comparative Labor Law & Policy Journal", 41(1), 2019, pp. 95–126.
6. Baiocco, Sara, Fernández-Macías, Enrique, Rani, Uma and Pesole, Annarosa, *The Algorithmic Management of Work and Its Implications in Different Contexts*, European Commission, 2022.
7. Butler, Michael J.R., O'Broin, Holly L.R., Lee, Nick and Senior, Carl, *How Organisational Cognitive Neuroscience Can Deepen Understanding of Managerial Decision-Making: A Review of the Recent Literature and Future Directions*, "International Journal of Management Reviews", Vol. 18, 2016, pp. 542–559.
8. De Hert, Paul and Papakonstantinou, Vagelis, *The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals?*, "Computer Law & Security Review", 32(2), 2016, pp. 179–194.
9. De Stefano, Valerio and Taes, Simon, *Algorithmic Management and Collective Bargaining*, "Transfer: European Review of Labour and Research", 29(1), 2022. https://doi.org/10.1177/10242589221141055.
10. De Stefano, Valerio and Wouters, Mathias, *AI and Digital Tools in Workplace Management Evaluation: An Assessment of the EU's Legal Framework*, European Parliamentary Research Services, Scientific Foresight Unit (PE 729.516), 2022.
11. De Stefano, Valerio, *Negotiating the Algorithm: Automation, Artificial Intelligence and Labour Protection*, "Comparative Labor Law & Policy Journal", 41(1), 2019, pp. 15–46.
12. Goddard, Kate, Roudsari, Ali and Wyatt, Jeremy C., *Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators*, "Journal of the

American Medical Informatics Association", 19(1), 2012, pp. 121–127. https://doi.org/10.1136/amiajnl-2011-000089.
13. Jarrahi, Mohammad Hossein, *Artificial Intelligence and the Future of Work: Human-AI Symbiosis in Organisational Decision Making*, "Business Horizons", 61(4), 2018, pp. 577–586. https://doi.org/10.1016/j.bushor.2018.03.007.
14. Kellogg, Katherine C., Valentine, Melissa A. and Christin, Angèle, *Algorithms at Work: The New Contested Terrain of Control*, "Academy of Management Annals", Vol. 14, 2020, p. 366.
15. Kullmann, Miriam, *Platform Work, Algorithmic Decision-Making, and EU Gender Equality Law*, "International Journal of Comparative Labour Law and Industrial Relations", 34(1), 2018, pp. 1–21.
16. Lepri, Bruno, Oliver, Nuria, Letouzé, Emmanuel, Pentland, Alex and Vink, Patrick, *Fair, Transparent, and Accountable Algorithmic Decision-Making Processes*, "Philosophy and Technology", 31, 2018, pp. 611–627. https://doi.org/10.1007/s13347-017-0279-x.
17. Matthias, Andreas, *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata*, "Ethics and Information Technology", 6, 2004, pp. 175–183. https://doi.org/10.1007/s10676-004-3422-1.
18. Mittelstadt, Brent Daniel, Allo, Patrick, Taddeo, Mariarosaria, Wachter, Sandra and Floridi, Luciano, *The Ethics of Algorithms: Mapping the Debate*, "Big Data & Society", 3(2), 2016. https://doi.org/10.1177/2053951716679679.
19. O'Neil, Cathy, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Publishing, 2016.
20. Rieke, Aaron and Bogen, Miranda, *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias*, Labor and Employment, Upturn, 10 December 2018. https://www.upturn.org/reports/2018/hiring-algorithms/ accessed 16 March 2025.
21. Ruschemeier, Hannah and Hondrich, Lukas J., *Automation Bias in Public Administration – An Interdisciplinary Perspective from Law and Psychology*, "Government Information Quarterly", 41(3), 2024, 101953. https://doi.org/10.1016/j.giq.2024.101953.
22. Santosuosso, Amedeo and Sartor, Giovanni, *Decidere con l'IA: Intelligenze Artificiali e Naturali nel Diritto*, Il Mulino, 2024.
23. Skitka, Linda J., Mosier, Kathleen and Burdick, Mark D., *Accountability and Automation Bias*, "International Journal of Human-Computer Studies", 52(4), 2000, pp. 701–717. https://doi.org/10.1006/ijhc.1999.0349.
24. Wachter, Sandra, Mittelstadt, Brent and Floridi, Luciano, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, "International Data Privacy Law", 7(2), 2017, pp. 76–99. https://doi.org/10.1093/idpl/ipx005.