Available online at https:
https://journal.uitm.edu.my/ojs/index.php/JIKM

e-ISSN: 2289-5337

Journal of Information and Knowledge Management (2025) Vol. 15, No. 2

Journal of Information and Knowledge Management

A Systematic Literature Review on the Impact of Data Mining on Workforce Privacy and Employment Practices: Ethical Concerns and Risks

Azratul Ain Nadiah Ismail, Ahmad Nadzri Mohamad*

Faculty of Information Science, Universiti Teknologi MARA Selangor Branch, Puncak Perdana Campus, 40150 Shah Alam, Selangor, Malaysia

Corresponding Authors' Email Address: nadzri.mohamad@gmail.com

ARTICLE INFO

Article history:
Received: 13 June 2025
Revised: 25 July 2025
Accepted: 21 August 2025
Online first
Published: 1 October 2025

Keywords:
Algorithmic decision-making
Data mining
Employee autonomy
Ethics
Privacy

https://doi.10.24191/jikm.v15i2.6962

ABSTRACT

Data mining tools enabled by artificial intelligence can change the norms in employment practices. Employers can capitalize on data mining technologies to assist in the hiring process, employee performance assessment and behavior surveillance at the workplace. While these practices are becoming more prevalent, there is a lack of consolidated research that covers the ethical concerns of data mining on workforce privacy and employment practices. Given this context, the study examines ethical concerns and risks associated with using data mining tools in the respective sphere. The research applied the PRISMA 2020 statement for a systematic literature review. A Python script was used to assist in selecting relevant articles based on selected keywords. VOSviewer was utilized as a bibliometric mapping tool to provide a preliminary understanding of the retrieved scholarly articles. The study reviewed 154 scholarly articles from four online databases. This led to the inclusion of 21 articles for the systematic literature review. The findings suggest that employees are concerned about system biases and unconsented data usage for algorithmic decisions in employment practices. This includes the lack of transparency in using data mining tools and artificial intelligence. Another concern is using emotional data for employee profiling. Emotional data can be used to monitor work performance and behaviour through wearable devices or cameras. As such, employees have a 'trust deficit' with data mining tools and systems in work-related decision making. This is when employees view these systems as having 'less empathy' in decision-making. For that reason, better mechanisms are required to enhance trust and confidence in using these systems. This includes strengthening legal aspects and frameworks to secure employees' trust and rights. Future studies can use the findings as a theoretical basis to explore the research topic in medium and large corporations across countries.

INTRODUCTION

The rapid integration of data mining and artificial intelligence into employment processes has fundamentally transformed the workplace. Artificial intelligence applications in hiring practices to monitor employee performance have become widespread in organizations. These

practices could help companies increase productivity and profits while decreasing losses associated with employee turnover. (Kamila & Jasrotia, 2023; Wang et al., 2023). Likewise, data mining could help companies to select preferred candidates in an interview (Nasiri et al., 2025), manager selections based on performance data (Saboori-Taft et al., 2025), and optimizing human resources training (Chen et al., 2024). Nonetheless, data mining in employment has a dual-edged nature. As such, there is a need to question the ethical considerations of these technologies and practices as they become integrated into human resource routines. Data mining in employment refers to the systematic organizational analysis of employee-related data using computational instruments such as machine learning, natural language processing, and behavioural algorithms to examine workforce information for decisions regarding recruitment, promotions, risk profiling, and productivity (Prikshat et al., 2022; Tursunbayeva et al., 2022).

Data mining offers efficiencies and objectivity in personnel decisions to a degree. For example, mining emotional data helps university administrators identify the psychological pressures of college students in seeking job opportunities (Chen, 2025). In another instance, data mining could help institutions analyze employee behaviours regarding absenteeism and attrition (Nawaz et al., 2024). In this sense, using historical data on absenteeism allows algorithms to analyze recurring patterns and predict future work nonattendance. However, there is always a risk related to system biases in data mining applications. These system biases might lead to non-transparent decisions that become difficult for both employers and employees to audit or contest (McCartney & Fu, 2022; Ryan et al., 2021; Weinhardt, 2021). Hence, the nature of algorithmic decisions on employment practices raises questions about justice, accountability, and privacy of employees in the workplace (Gooding & Kariotis, 2021; Rana et al., 2023).

The literature demonstrates growing concern that algorithmic instruments are not neutral. Studies have documented how biased data inputs and unregulated deployment of data mining applications can lead to exclusion and mistrust of employees (Schwarcz et al., 2020; Wang et al., 2023). Incidents involving biased resume screening algorithms and intrusive monitoring systems further emphasize the urgency of these concerns. In fact, data mining applications on open data or even social media sources might lead to imprecise analysis due to the lack of contextual details on the nature of the data made available (Nawaz et al., 2024).

One of the significant issues concerning data mining applications is the use of unethical personal data. For that reason, several national and regional initiatives covering legal aspects of using personal data have been established. One example of such an initiative in Malaysia is embedded in the National Data Protection Act 2010. Other Southeast Asian countries such as Indonesia, Singapore and Thailand have established robust data protection acts. At a regional level, the ASEAN Framework on Personal Data Protection (2016) became a point of reference to safeguard personal data usage across regional countries. Similarly, the European Union's General Data Protection Regulation (GDPR) aims to regulate the use of personal data in Europe (Ienca & Malgieri, 2022; Weinhardt, 2021). It is worth mentioning that the enforcement of GDPR framework is still limited across companies, even when the framework was introduced in 2018 (Ienca & Malgieri, 2022; Weinhardt, 2021).

Despite awareness of these issues, a substantial gap remains in research that comprehensively evaluates the ethical and privacy implications of data mining in employment settings. Limited studies rigorously examine how such technologies influence employee autonomy and trust, the

severity of privacy risks, and the adequacy of legal safeguards. Existing research is often fragmented across AI ethics, cognitive privacy, organizational behavior, and employment law, necessitating a more integrated synthesis. This systematic literature review addresses that gap through the following research question:

1) What are the ethical concerns and risks associated with the use of data mining techniques in employment practices that might affect employee autonomy and trust?

This study employed the PRISMA 2020 statement for a systematic literature review. Four online databases were used to extract relevant articles in the initial phase. After that, the study utilized a Python script to screen the retrieved articles based on predetermined keywords structured from the research question. The researchers used bibliometric mapping software (VOSViewer) to facilitate a preliminary understanding of the screened articles (Mohamad et al., 2024). Subsequently, the study reviewed 21 scholarly articles published between 2015 and 2025. The research extracted articles in the study with topics that span from data science, human resource management, AI ethics, legal studies, and behavioral analytics. The research contributes to a consolidated understanding of how data mining technologies affect the ethical and legal dimensions of workforce management in literature.

METHODOLOGY

Literature Screening and PRISMA Documentation

This study employed a systematic literature review (SLR) to examine ethical concerns and privacy risks associated with data mining in employment practices. The review addressed a research question and followed the PRISMA 2020 statement. The researchers retrieved 154 records, and four records were removed due to duplication. After Python-based and manual abstract screening, 22 articles were excluded due to their lack of relevance in relation to the research question. A full-text assessment of 128 articles led to 107 exclusions, and the final 21 peer-reviewed articles were identified. The authors assessed the quality of the scholarly articles and discussed concerns such as article relevancy and possible biases, such as funding origin. The PRISMA flowchart is shown in the Results section.

In this study, the researchers employed a Python script for semi-automated filtering, thereby minimizing bias in the screening process. After that, VOSviewer, a bibliometric mapping software, was utilized to identify the keywords associated with the retrieved articles. This allowed a preliminary understanding of the contents of the retrieved documents. Further reading of the retrieved documents enabled the researchers to answer the research question. The study used only publicly available secondary data, so no ethical clearance was required.

Search Strategy and Inclusion Criteria

A Boolean search string was developed to identify relevant studies. The string included key terms such as "data mining," "job market," "employee monitoring," "ethical decision-making," "performance evaluation," and "regulatory frameworks". The search string used was:

("data mining" AND ("job market" OR employment) AND ("ethical decision-making" OR "performance evaluation" OR "employee monitoring" OR "hiring" OR "recruitment") AND ("regulatory frameworks" OR "data protection laws")

The search spanned four databases: Emerald Insight, Gale Research Complete, ProQuest, and Springer Nature Link. The authors selected these databases due to their familiarity with these sources. Also, these databases are subscribed by the university library. Hence, the authors are able to retrieve the scholarly sources relevant to the research topic. The authors acknowledge that other online databases can be included for future studies. The limited nature of database coverage for the study is stated in the limitations and recommendations section. Only English peer-reviewed journal articles from 2015–2025 were included in the study.

The use of a Python script and VOSViewer

The researchers conducted a semi-automated review of titles and abstracts using a Python script (version 3.13.3) in Appendix A based on predetermined keywords aligned with the research question. A Python script was used to ensure consistency across the selection process, filtering out articles that did not address the research question before a full-text assessment. The researchers used VOSviewer (version 1.6.20) to generate network maps based on keyword co-occurrence patterns. These conceptual maps allowed the researchers to analyze the essence of the retrieved scholarly articles. The researchers used these visualization maps to understand the keywords associated with the screened articles.

FINDINGS

A total of 154 records were retrieved from four academic databases: Emerald Insight, Gale Research Complete, ProQuest, and Springer Nature Link. Research titles and abstracts were screened using a combination of Python-assisted filtering and manual review to assess relevance. After screening and full-text eligibility checks based on inclusion and exclusion criteria, 21 peer-reviewed journal articles were selected for final synthesis. The detailed selection process is illustrated in Figure 1. This figure illustrates the identification, screening, and inclusion stages of the systematic literature review. A total of 154 records were retrieved from four academic databases. After duplicate removal (n = 4), 150 records were screened through title and abstract review using Python automation and manual filtering. Of the 128 full-text articles assessed for eligibility, 20 peer-reviewed studies were included in the final synthesis. Exclusion criteria included the lack of relevance to employment contexts and the absence of discussions on ethics or privacy in relation to the research question.

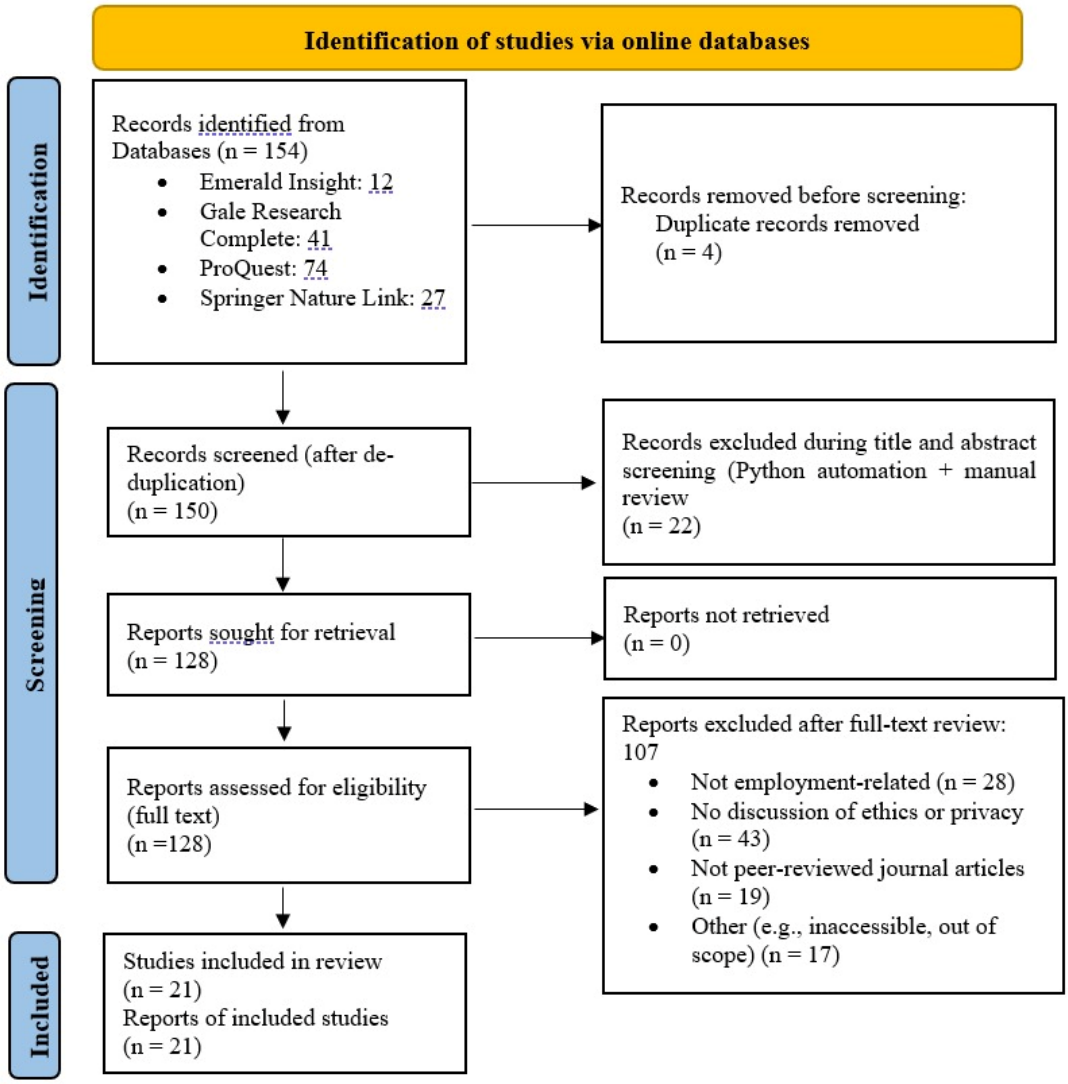

Figure 1: PRISMA 2020 diagram

Preliminary understanding through VOSviewer

VOSviewer was employed to visualize conceptual linkages across the selected literature. This allows familiarization of literature as suggested by scholars (A.N. Mohamad et al., 2022). Network visualization (Figure 2) reveals prominent co-occurrences among terms such as data mining, ethics, privacy, and machine learning.

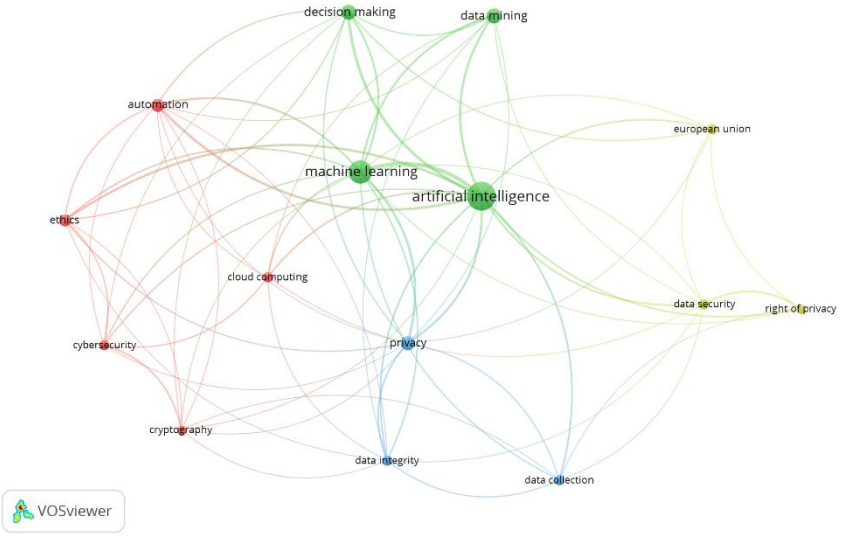

Figure 2: Network visualization

This network visualization demonstrates the conceptual relationships among 15 key terms. This network visualization illustrates the conceptual relationships among frequently co-occurring terms across the 21 peer-reviewed journal articles included in the systematic literature review. Artificial intelligence and machine learning appear at the center of the network, reflecting their centrality to discussions of data mining in employment contexts. Closely related terms, such as privacy, ethics, decision-making, and data security, are also found within the colored thematic clusters. The visualization demonstrates the multidisciplinary and interconnected nature of academic discourse on data mining applications in employment practices.

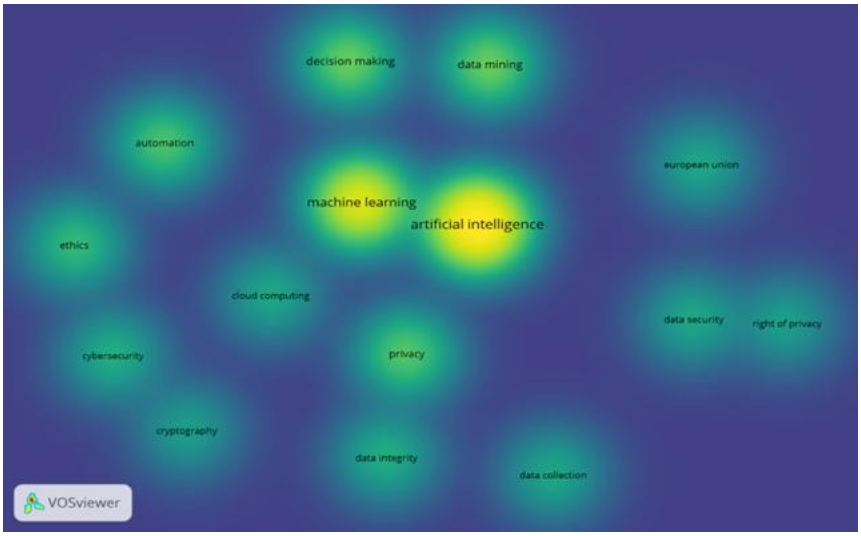

Figure 3: Density Visualization of Keyword Co-occurrences in Selected Studies

Figure 3 shows visualization of keyword co-occurrences in selected studies. This density visualization gives an overview of central themes such as artificial intelligence, machine learning, and privacy as the main concepts of the retrieved literature. These terms appear in the densest areas, indicating their prominence in the current academic literature about data mining and employment practices.

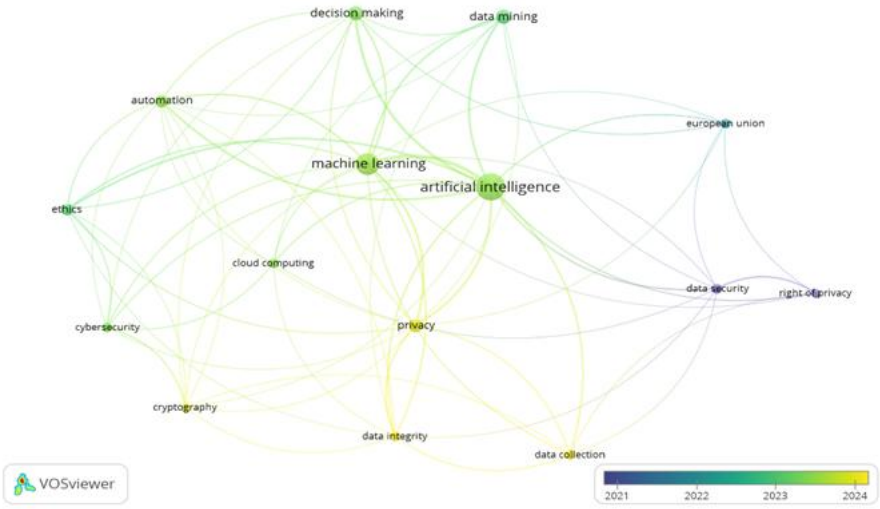

Figure 4: Overlay Visualization of Keyword Co-occurrence by Year in the Selected Studies

Figure 4 shows overlay visualization on the temporal trends in keyword use exponentially per year from 2021 to 2024. We can be informed that there has been an increase in attention to terms such as "right of privacy", "European Union", and "data security" in recent years. This escalated attention infers the constructs of increased academic interest in legal protections contextualized around the regulatory conversation.

Table 1 summarizes the presence of fifteen recurring ethical concerns: surveillance, consent violation, bias or discrimination, opacity or transparency, fairness and accountability, autonomy loss, governance gaps, lack of explainability, power imbalance or control, chilling effect or silencing, ethical invisibility, inequity in outcomes, mental health impact, consent ambiguity, and trust and reliability. Each concern was ticked based on its explicit discussion or thematic centrality in the reviewed studies.

Table 1: Thematic centrality of the reviewed studies

Author(s) / Year	Surveillance	Consent Violation	Bias / Discrimination	Transparency	Fairness and Accountability	Autonomy Loss	Governance Gaps	Lack of Explainability	Power Imbalance / Control	Chilling Effect / Silence	Ethical Invisibility	Inequity in Outcomes	Mental Health Impact	Consent Ambiguity	Trust and Reliability
Chhabra Roy & Prabhakaran (2023)			✓	✓	✓		✓		✓		✓	✓			
Tursunbayeva et al. (2022)	✓	✓	✓		✓	✓	✓								
Wang et al. (2023)			✓	✓	✓			✓			✓				
Milovanović et al. (2022)			✓	✓	✓		✓								✓
Ryan et al. (2021)		✓				✓		✓					✓		✓
Torbert (2021)		✓							✓		✓				
Lu (2020)			✓	✓			✓								
Islam et al. (2022)			✓	✓	✓			✓							
Sessa et al. (2024)			✓	✓	✓	✓	✓		✓						
Prince & Schwarcz (2020)	✓	✓	✓	✓	✓				✓	✓	✓	✓		✓	
McCartney & Fu (2022)		✓				✓					✓		✓		✓
Black et al. (2024)				✓	✓		✓								

Table 1: Thematic centrality of the reviewed studies

Author(s) / Year	Surveillance	Consent Violation	Bias / Discrimination	Transparency	Fairness and Accountability	Autonomy Loss	Governance Gaps	Lack of Explainability	Power Imbalance / Control	Chilling Effect / Silence	Ethical Invisibility	Inequity in Outcomes	Mental Health Impact	Consent Ambiguity	Trust and Reliability
Kabashkin et al. (2023)			✓	✓	✓										
Lu (2021)	✓	✓		✓											
Abdul Mateen et al. (2023)			✓	✓			✓	✓							✓
Ienca & Malgieri (2022)		✓				✓	✓				✓		✓		
Verma et al. (2022)			✓	✓	✓			✓							✓
Kamila & Jasrotia (2025)				✓	✓										✓
Rana, Azizul & Awan (2023)		✓	✓	✓											
Gutierrez (2022)	✓	✓							✓		✓				
Weinhardt (2021)	✓	✓							✓		✓				

Note. *A tick (✓) indicates that the ethical concern is explicitly discussed in the reviewed studies.*

DISCUSSION

The study poses a research question that examines the ethical concerns and risks associated with using data mining techniques in employment practices. This section discusses research findings in light of the existing literature.

Some ethical concerns are discussed far more frequently than others, based on Table 1.0. For instance, "bias and discrimination" is one of the main topics discussed in the reviewed literature. Also, "transparency" issues have become a central topic in the reviewed scholarly articles. "Fairness and accountability" also receive substantial attention. Other concerns do not share this spotlight. For example, "consent violation" is undeniably foundational to ethical data governance. Yet in many studies, it appears as a procedural reference — mentioned, yes, but rarely interrogated. Another theme, which is "autonomy loss" follows a similar pattern. Although this concern deals directly with how individuals experience algorithmic control in the workplace, it remains underexplored. This is surprising. Given how algorithms now influence hiring, monitoring, and even termination, one would expect a richer discussion.

The reviewed literature displays inconsistencies across authors due to their research interests. For example, Prince & Schwarcz (2020) offer detailed coverage of "bias/discrimination", "transparency", and "fairness and accountability". Their work also addresses "consent violation" and "chilling effect/silence", and "inequity in outcomes" as themes. By contrast, Lu (2020) emphasizes "governance gaps" while again engaging with "bias/discrimination" and "transparency" as concerning themes. Another study, McCartney and Fu (2022) explored "trust and reliability", "consent ambiguity", and "autonomy loss". This fragmented focus may reflect disciplinary differences or research interests between authors in the reviewed scholarly journals.

One of the primary concerns found in the research involves the utilization of algorithms for unethical work-related decisions. This concern is reflected in some of the articles that fall under the main themes, such as "transparency" and "fairness and accountability". For example, one of the scholarly articles we reviewed noted that many people's analytics systems fail to highlight transparency in human resource decisions (Tursunbayeva et al., 2022). Our study also found evidence highlighting the lack of understanding of algorithmic decisions on work-related matters (Prikshat et al., 2022). This situation might affect the opportunities for employees to contest decisions that might affect their careers (Prikshat et al., 2022). In fact, the findings suggest that the absence of informed consent in data mining applications might lead to power imbalances between employers and employees (Sessa et al., 2024). Such a situation might create relationship disharmony and dissatisfaction at the workplace. The systematic literature review acknowledged the importance of data mining applications in the emergence of artificial intelligence in the workplace. The study also realized from the reviewed literature that employees are aware of the risks of using these applications in human resource decisions.

Algorithmic models generate concerns about bias or discrimination, which aligns with one of the themes that we found in the study. The findings suggested that employees are concerned about biases associated with data mining in employment practices, such as hiring and performance evaluation. For instance, when an algorithm receives training on historical or unbalanced datasets, the model merely reinforces discriminatory hiring and performance evaluation against employees (Gutierrez, 2021; Rana et al., 2023; Schwarcz et al., 2020). Relatedly, one of the scholarly articles

we reviewed emphasizes that AI systems often magnify existing social inequalities due to a lack of intersectional fairness models, leading to discriminatory impacts on marginalized groups (Islam et al., 2023). To make matters worse, the use of artificial intelligence tools to monitor productivity and emotional responses raises mental health concerns. Such a concern coincides with another theme we discovered in the reviewed literature, which is "mental health impact". In this context, using artificial intelligence tools for work performance might contribute to mental stress, fear, and decreased job satisfaction (Gooding & Kariotis, 2021; Lu, 2021). In some instances, facial recognition, wearable technology and employability sensors for monitoring are considered invasive of individual privacy (McCartney & Fu, 2022). To this end, the research acknowledged that data mining applications can create biases in human resource practices. Additionally, employees considered that data mining tools might invade their privacy at the workplace. Also, the reviewed literature insinuates that there are mental health concerns for employees regarding the research topic.

Automation and prediction in recruitment lead to greater efficiency in the hiring process. However, the findings suggest that data mining tools used to evaluate interview candidates may introduce algorithmic biases. These tools might use personal data such as past work experiences and educational backgrounds as a scoring mechanism. Hence, it is possible to exclude candidates in an interview through a system without split seconds, leading to rejections without knowing why. Organizations often lack a meaningful way to analyze algorithmic logic within such hiring exercises (Rana et al., 2023). In our study, one research highlights that hiring technologies often prevent stakeholders from understanding how employment decisions are made. This situation could leave interview candidates being rejected without recourse or explanation (Lu, 2021). In addition, the systematic literature review also revealed that data mining tools are used to monitor or predict work misconduct. This situation can be considered an ethical dilemma, as it could stigmatize and misjudge individuals based on certain inferred traits and behaviours (Gooding & Kariotis, 2021).

Similarly, the findings suggest that digital phenotyping or sentiment analysis systems can collect behavioural and emotional data without consent, potentially leading to invasive employee profiling (Ienca & Malgieri, 2022). These systems can collect real-time data, such as employee keystrokes or facial expressions, which raises ethical dilemmas about whether mental states can be recorded and retained as analyzable corporate records (Ryan et al., 2021). One of the journal articles that we reviewed cautions that behavioural authentication systems may collect extensive personal data and be misused for continuous surveillance (Wang et al., 2024). As such, employees might not be well aware of the data collected for performance evaluation (Rana et al., 2023). Consequently, this undermines accountability and deprives employees of the ability to appeal unjust or inappropriate outcomes. The abovementioned findings fall under several themes of the study, mainly "consent ambiguity", concern violation", "trust and reliability" and "fairness and accountability". In brief, the findings advocate that data mining tools can be convenient for employers. Nonetheless, some risks could lead to unfavourable decisions against employees. As such, some practical guidelines must be in place to guide data mining for human resource management.

Additionally, employees can be constantly monitored through wearable devices or software-based performance management applications. Although these applications may offer greater objectivity, they may lack consideration of contextual factors surrounding a team or the emotional

well-being of individuals. One of the journal articles that we reviewed acknowledges that pervasive monitoring practices risk creating an atmosphere of digital micromanagement, where productivity is prioritized over well-being. Furthermore, systems of this nature might dismiss opportunities for empathy and discretion in work-related decision-making (Gooding & Kariotis, 2021; Torbert, 2020). Moreover, employee monitoring systems that track email, tone of voice, and visual presence increase the ethical dilemmas associated with affective computing tools, which make inferences about employee mental states without permission. Also, predicting burnout or misconduct through continuous monitoring is another ethical dilemma. It may stigmatize individuals based on inferred traits and exclude individuals without any basis (Gooding & Kariotis, 2021).

Scholars have emphasized the development of ethical systems that prioritize fairness, accountability, and transparency. As such, scholars recommend that employers design data mining applications in employment practices with an accountable internal structure to balance employee-employer rights (Ryan et al., 2021; Tursunbayeva et al., 2022). This is when employee involvement can help rebuild trust and protect their rights (Ryan et al., 2021). This recommendation is echoed in other scholarly articles that we reviewed. For example, Lu (2020) proposed that companies adopt algorithmic disclosure efforts similar to financial reporting to enhance corporate responsibility. In some instances, scholars associated the need to develop legal frameworks for data utilization, such as the General Data Protection Regulation (Ienca & Malgieri, 2022; Islam et al., 2023). Moreover, scholars recommend mechanisms such as Mental Data Protection Impact Assessments and internal ethics committees to ascertain employee rights (Gooding & Kariotis, 2021; Tursunbayeva et al., 2022). Regulations such as the General Data Protection Regulation (GDPR) typically focus on safeguarding individuals from detrimental effects through personal data usage. However, these regulations and frameworks are sporadically adopted in practice. In short, companies must ensure employee consultation, including the utilization of algorithmic audits or framework adoptions to maintain fairness across employment practices.

Some researchers believe it is better to deal with AI fairness before these systems can be fully built. In this sense, Rana et al. (2023) mentioned that problems with system biases are easier to manage when caught early in the process. If AI-based systems are running, it may already be too late or too complicated to fix them. But even with technical solutions in place, that is just not enough. People affected, such as employees, should also be part of the conversation to improve AI-assisted data mining applications. Tursunbayeva et al. (2022) pointed out something similar. They informed that rules around ethics work best when developed together with the people involved in the process. It is not just about assisting employers in human resource matters with technologies. It is about making sure employees are treated fairly at the workplace.

LIMITATIONS AND RECOMMENDATIONS

While this review synthesizes ethical concerns and privacy risks in employment-related data mining, several limitations are noted. The study focused solely on English-language peer-reviewed articles. It is also important to note that most studies originated from Western contexts. Future research should incorporate primary data collection and expand its geographic and linguistic coverage. Furthermore, a field expert did not validate the keywords selected for Python scripting. Also, the researchers did not use inter-rater reliability to evaluate the selected articles in the screening process. Future research might consider a field expert to provide some advice on

the selection of keywords (Mohamad, 2025). Also, scholars can consider including other online databases to capture the essence of the topic. This includes using abstract and citation databases such as Scopus and Web of Science. While doing so, it should be noted that not all scholarly articles can be accessed from these databases due to subscription limitations. Scholars might also develop practical frameworks that emphasize algorithmic transparency, auditing, and employee participation to address unethical data mining applications in the workplace.

The study offers practical recommendations to human resource practitioners, policymakers and organizational leaders. Human resource practitioners and organizational leaders need to engage employees to develop a mutual understanding of the use of data mining technologies for performance management. Equally important are continuous assessments of system effectiveness in human resource matters such as hiring and personalized training development. Such a measure is necessary because these systems might use outdated data for decision-making (Saboori-Taft et al., 2025). Also, system developers and operators must recognize system shortcomings early in implementation. Employee consultation at this stage would help to improve the systems further. Also, an accountable internal structure can be assigned to balance employer-employee rights at an institutional level. At a national level, it is ideal for policymakers to design personal data protection frameworks as a guide for wider implementation based on identical guidelines from other regions (European Economic and Social Committee, 2021).

CONCLUSION

This systematic literature review examined the ethical concerns and privacy risks associated with data mining applications in employment practices. Based on the findings of twenty-one peer-reviewed articles, it is often determined that algorithmic decision-making is perceived as non-transparent and violates employees' autonomy and privacy. Data mining may deliver efficiencies but raises ethical issues about fairness, accountability, and potential discriminatory practices. The research believes that transparency within algorithms, audits, and employee involvement is integral to a strategy for building trust. Employee participation is key to fair and accountable use of AI-based employment practices. Future research should examine employee experience with data mining applications and how ethical models can be used to ensure responsible and equitable use of technology. Scholars may also include other stakeholders triangulate qualitative data, such as from employers and organizational leaders. This could provide a comprehensive understanding of the implementation of related systems from the employer and employee perspectives. Future research can focus on specific ethical concerns raised in this study and how best to address them through system and people-oriented improvements.

ACKNOWLEDGEMENTS

The authors would like to thank everyone who spent time, offered support, and shared their expertise towards this research. The study received no funding assistance from any public or private institution.

REFERENCES

Abdul Mateen, Adia Khalid, Lee, S., & Nam, S. Y. (2023). Challenges, issues, and recommendations for blockchain and cloud-based automotive insurance systems. *Applied Sciences*, 13(6), 3561. https://doi.org/10.3390/app13063561

Black, E., Koepke, J. L., Kim, P. T., Barocas, S., & Hsu, M. (2024). Less discriminatory algorithms. *Georgetown Law Journal*, 113(1), 53–120. https://doi.org/10.2139/ssrn.4590481

Chen, G. (2025). The application of text mining and deep learning in identifying the psychological pressure of college students in employment. *Discover Artificial Intelligence*, 5(117). https://doi.org/10.1007/s44163-025-00372-z

Chen, X., Chen, H., & Xu, J. (2024). An innovative approach to corporate HR training based on deep learning. *Applied Mathematics and Nonlinear Sciences*, 9(1), 1–17. https://doi.org/10.2478/amns-2024-0419

Chhabra Roy, Neha & Prabhakaran, Sreeleakha. (2023). Sustainable response system building against insider-led cyber frauds in banking sector: A machine learning approach. *Journal of Financial Crime*, 30(1), 48–85. https://doi.org/10.1108/JFC-12-2021-0274

Gooding, P., & Kariotis, T. (2021). Ethics and law in research on algorithmic and data-driven technology in mental health care: Scoping review. *JMIR Mental Health*, 8(6). https://doi.org/10.2196/24668

Gutierrez, M. (2021). Algorithmic gender bias and audiovisual data: A research agenda. *International Journal of Communication*, 15, 439–461.

Ienca, M., & Malgieri, G. (2022). Mental data protection and the GDPR. *Journal of Law and the Biosciences*, 9(1), 1–19. https://doi.org/10.1093/jlb/lsac006

Islam, Rashidul, Keya, K. N., Pan, S., Sarwate, A. D., & Foulds, J. R. (2023). Differential fairness: An intersectional framework for Fair AI. *Entropy*, 25(4), 1–44. https://doi.org/10.3390/e25040660

Kamila, Manoj Kumar & Jasrotia, Sahil Singh. (2025). Ethical issues in the development of artificial intelligence: Recognizing the risks. *International Journal of Ethics and Systems*, 41(1), 45–63. https://doi.org/10.1108/IJOES-05-2023-0107

Lu, S. (2021). Algorithmic opacity, private accountability and corporate social disclosure in the age of artificial intelligence. *Vanderbilt Journal of Entertainment and Technology Law*, 23(1). https://scholarship.law.vanderbilt.edu/jetlawAvailableat:https://scholarship.law.vanderbilt.edu/jetlaw/vol23/iss1/3

Madinier, F. S. (2021). *A guide to artificial intelligence at the workplace - Your rights on algorithms*. https://www.eesc.europa.eu/sites/default/files/files/qe-03-21-505-en-n.pdf

McCartney, S., & Fu, N. (2022). Promise versus reality: A systematic review of the ongoing debates in people analytics. *Journal of Organizational Effectiveness*, 9(2), 281–311. https://doi.org/10.1108/JOEPP-01-2021-0013

Milovanović, S., Bogdanović, Z., Labus, A., Despotović-Zrakić, M., & Mitrović, S. (2022). Social recruiting: An application of social network analysis for preselection of candidates. *Data Technologies and Applications*, 56(4), 536–557. https://doi.org/10.1108/DTA-01-

2021-0021

Mohamad, Ahmad Nadzri, Sylvester, A., & Campbell-Meier, J. (2022). Towards a taxonomy of emerging topics in open government data: A bibliometric mapping approach. *Proceedings of the 55th Hawaii International Conference on System Sciences*, 7, 2553–2562. https://doi.org/10.24251/hicss.2022.317

Mohamad, Ahmad Nadzri, Sylvester, A., & Campbell-Meier, J. (2024). Towards a taxonomy of research areas in open government data. *Online Information Review*, 48(1), 67–83. https://doi.org/10.1108/OIR-02-2022-0117

Mohamad. Ahmad Nadzri (2025). Investigating the actualization of open data affordances for start-up entrepreneurs. *Information Discovery and Delivery*, 53(3), 366–374. https://doi.org/10.1108/IDD-03-2024-0050

Nasiri, Hamidreza, Moshiri, Behzad & Ardestani, Farshad Fatemi (2025). Empirical implementation of extraction of decision-maker's evaluation method in job candidates selection based on interpretable data mining tools. *Operational Research*, 25(2), 1–36. https://doi.org/10.1007/s12351-025-00930-4

Nawaz, M. Saqib, Nawaz, M. Zohaib, Fournier-Viger, P., & Luna, J. M. (2024). Analysis and classification of employee attrition and absenteeism in industry: A sequential pattern mining-based methodology. *Computers in Industry*, 104106, 159–160. https://doi.org/10.1016/j.compind.2024.104106

Prince, A. E. R., & Schwarcz, D. (2020). Proxy discrimination in the age of artificial intelligence and big data. *Iowa Law Review*, 105(3), 1257–1318.

Rana, Saadia Afzal, Azizul, Zati Hakim & Awan, Ali Afzal. (2023). A step toward building a unified framework for managing AI bias. *PeerJ Computer Science*, 9:e1630, 1–27. https://doi.org/10.7717/peerj-cs.1630

Ryan, M., Antoniou, J., Brooks, L., Jiya, T., Macnish, K., & Stahl, B. (2021). Research and practice of AI ethics: A case study approach juxtaposing academic discourse with organisational reality. *Science and Engineering Ethics*, 27(2), 1–29. https://doi.org/10.1007/s11948-021-00293-x

Sessa, F., Chisari, M., Esposito, M., Karaboue, M. A. A., Salerno, M., & Cocimano, G. (2024). Ethical, legal and social implications (ELSI) regarding forensic genetic investigations (FGIs). *Journal of Academic Ethics*, 23(3), 617–637. https://doi.org/10.1007/s10805-024-09582-z

Torbert, P. (2021). Because it is wrong: The immorality and illegality of the online service contracts of Google and Facebook. *Journal of Law, Technology & the Internet*, 12(1), 1–173. https://doi.org/10.2139/ssrn.3537227

Tursunbayeva, A., Pagliari, C., Di Lauro, S., & Antonelli, G. (2022). The ethics of people analytics: Risks, opportunities and recommendations. *Personnel Review*, 51(3), 900–921. https://doi.org/10.1108/PR-12-2019-0680

Verma Prikshat, Parth Patel, Arup Varma, & Ishizaka, A. (2022). A multi-stakeholder ethical framework for AI-augmented HRM. *International Journal of Manpower*, 43(1), 226–250. https://doi.org/10.1108/IJM-03-2021-0118

Wang, C., Tang, H., Zhu, H., Zheng, J., & Jiang, C. (2024). Behavioral authentication for security and safety. *Security and Safety*, 3, 2024003. https://doi.org/10.1051/sands/2024003

Weinhardt, M. (2021). Big data: Some ethical concerns for the social sciences. *Social Sciences*, 10(2), 1–14. https://doi.org/10.3390/socsci10020036

APPENDIX A

Python script (version 3.13.3)

```
import pandas as pd
import re

# Step 1: Load the CSV file containing the article titles and abstracts
df = pd.read_csv('C:/Users/70076083/Downloads/Command/mendeley_export.csv')

# Step 2: List of selected keywords based on your research objectives
keywords = [
'data analytics', 'data collection', 'data management', 'data quality', 'data-driven',
'accuracy', 'data mining', 'data processing', 'information retrieval', 'cybersecurity',
'data integrity', 'ethics', 'privacy', 'digital transformation', 'efficiency'
]

# Step 3: Function to check if any of the keywords appear in the title or abstract
def keyword_match(text, keywords):
# Convert text to lower case and check for keyword matches
text = str(text).lower()
for keyword in keywords:
if re.search(r'\b' + re.escape(keyword) + r'\b', text):
return True
return False

# Step 4: Apply the function to filter rows where keywords are found in the title or abstract
filtered_df = df[df['title'].apply(lambda x: keyword_match(x, keywords)) |
                 df['abstract'].apply(lambda x: keyword_match(x, keywords))]

# Step 5: Save the filtered data to a new CSV file
filtered_df.to_csv('filtered_articles.csv', index=False)

# Step 6: Print the filtered articles
print(filtered_df)
```