# Errors in AI-assisted recruitment and the role of cybersecurity in mitigating risks

Cristina Leovaridis
*FCRP-SNSPA, Bucharest, Romania*
cristina.leovaridis@comunicare.ro

Florin-Domnel Grafu
*ROMATSA, Bucharest, Romania*
florin.grafu@romatsa.ro

## Abstract
The primary objective of this research is to demonstrate the necessity of integrating cybersecurity mechanisms as a fundamental technical foundation for mitigating processing errors in AI-assisted recruitment. The paper aims to bridge the identified gap between theoretical ethical discourse and the practical implementation of technical control solutions within organizational environments. Specialized literature extensively documents algorithmic bias and discrimination, particularly regarding gender and ethnicity, as well as the risks of ”black-box” opacity in automated hiring. Previous research emphasizes the transition from human interaction to complex digital filters, while highlighting applicants' preferences for "augmented" hiring approaches rather than purely automated ones. Our quantitative bibliometric analysis of 1338 scientific articles (2020–2026) reveals a significant ”security deficit” in Human Resources (HR) workflows. The data shows a fracture between high technological performance and limited risk control, where cybersecurity is often perceived as an external administrative challenge rather than an internal control instrument. Consequently, the paper proposes a five-layer operational framework, covering data ingestion, preprocessing, training, decision and monitoring, to ensure data integrity and decisional equity. This study redefines cybersecurity from a defensive function into an active mechanism for guaranteeing social equity and mitigating algorithmic bias. By ”pulling” the cybersecurity node into the decisional center, the research provides a pragmatic, security-by-design solution for organizations. This approach transforms abstract ethical principles into quantifiable technical safety procedures, offering a structural pillar for transparent and resilient digital recruitment.

**Keywords:** selection, algorithmic bias, human resource management, security by design, technical resilience.

## 1. Introduction
The recruitment process has evolved from traditional methods, based on direct human interaction, toward a complex digital landscape where artificial intelligence acts as a primary decisional filter. While specialized literature extensively documents the existence of algorithmic bias and discrimination in AI-assisted recruitment, there remains a significant fragmentation between ethical discourse and the implementation of technical control solutions within organizational environments.

The purpose of this article is to demonstrate, through a quantitative bibliometric analysis, the necessity of integrating cybersecurity mechanisms as a foundation for mitigating processing errors, seeking to move beyond purely philosophical approaches in favor of active technical resilience. The paper proposes an operational framework that reconciles social equity with data integrity, aiming to provide a pragmatic solution for the 'security deficit' observed in current HR workflows.

## 2. AI-Driven Recruitment: Operational Mechanisms, Benefits and Ethical Risks

Given that contemporary leaders must manage digital platforms, data-driven decision-making processes, cyber risks and virtual collaboration [1] within a smart economy where the workforce is becoming increasingly specialized in modern data techniques [2], the adoption of Artificial Intelligence (AI) technologies in recruitment and selection processes has become an essential practice for organizations aiming to streamline workflows. These applications span from drafting job advertisements and screening candidate resumes to analyzing video interviews through facial recognition software. However, as these tools have a profound impact on candidates' professional trajectories and raise significant ethical concerns, the literature [3] emphasizes the necessity of a rigorous mapping of risks and mechanisms to mitigate algorithmic bias. AI-driven recruitment represents any procedure that utilizes AI to assist organizations during the recruitment and selection of candidates, while AI itself is conceptualized as a system's ability to rigorously process and interpret external data sets, extracting learning patterns that are subsequently used through flexible adaptation to achieve specific objectives and tasks [4].

Although numerous areas within the recruitment and selection process have been identified where AI applications can be implemented, specialists currently seem to rely predominantly on three: chatbots, screening software and task automation tools; furthermore, the majority of companies adopting these AI instruments tend to be large, technology-oriented, and/or innovative firms [5]. The role of AI in automating recruitment tasks consists of automated resume scanning, candidate interaction via chatbots, online assessments and predictive hiring through machine learning (ML). The benefits for organizations include cost savings, reduced time-to-hire, increased productivity and the optimization of administrative processes [6].

The use of Natural Language Processing (NLP) for automated resume screening involves algorithms extracting semantic entities to correlate candidate competencies with job requirements. Furthermore, the implementation of machine learning models enables the forecasting of future performance based on historical success profiles, while facial and voice recognition software in video interviews evaluates personality traits or non-verbal behavioral indicators [7], [8]. Utilizing Machine Learning (ML) algorithms to predict future candidate performance relies on processing massive volumes of historical data, a process that promises objectivity but may mask opaque decision-making mechanisms [7]. Although these systems promise increased efficiency, the literature warns against algorithmic opacity, which often triggers adverse reactions from candidates; recent research indicates that applicants tend to perceive selection based exclusively on AI/ML negatively, preferring instead 'augmented' approaches, where the decision results from collaboration between the algorithm and the human element [9].

Companies implement algorithmic decision-making to avoid or even transcend human biases. However, the integration of AI into candidate recruitment and selection introduces a range of errors, the most widely discussed being algorithmic bias: algorithms are vulnerable to prejudice regarding gender, ethnicity, sexual orientation or other characteristics, if they rely on inaccurate, biased or non-representative input and training data; algorithms replicate biases if the input data is already flawed [10]. This automated

prejudice does not stem from malicious intent, but is rather the result of training models on historical datasets that already contain latent social inequities.

Furthermore, if an algorithm 'learns' exclusively from current employees, it will tend to seek individuals who resemble the existing workforce (in terms of personality, academic background, communication style etc., to ensure team compatibility), thereby rejecting eminent but 'atypical' candidates and eliminating diversity and innovation [11]. The risk of discrimination intensifies in technical professions, where algorithms may systematically penalize female candidates if historical training data is dominated by male profiles, as observed in the Amazon case, or may limit ethnic minorities' access to opportunities through 'algorithmic objectives' that, in an attempt to minimize average prediction error, tend to favor majority groups over underrepresented ones. More recently, although attempts have been made to circumvent these biases by removing names or gender from resumes, it has been found that algorithms can infer these sensitive and protected traits from seemingly harmless and legitimate data. AI can 'guess' such characteristics from other details (e.g., hobbies, address, university attended), thus replicating historical discrimination under the guise of an objective technical selection [12]. These deficiencies are exacerbated by the 'black-box' phenomenon: the lack of computational transparency makes it impossible to explain why a candidate was rejected, rendering the employer's decisions opaque to the applicant [13] and triggering a transparency crisis in human resource management.

To prevent negative organizational repercussions, such as the erosion of the employer brand, decreased productivity or diminished employee motivation, it is imperative that firms constantly optimize the degree of perceived equity within talent management workflows, from recruitment to development programs [10]. Given that algorithmic biases lead to discriminatory hiring practices based on gender, race, color, personality traits and more, biases stemming from limited raw datasets and the subjectivity of algorithm designers, some authors [14] recommend implementing both technical measures, such as utilizing unbiased datasets and improving algorithmic transparency and managerial measures, such as internal ethical corporate governance and external oversight, to mitigate these errors.

## 3. Research Methodology

The present research adopts a quantitative bibliometric approach to present the conceptual evolution of the intersection between artificial intelligence, recruitment processes, and cybersecurity. The methodology utilizes keyword co-occurrence techniques. The data were extracted from the Web of Science (WoS) database. The choice of query terms was guided by the need to cover the multidimensionality of the researched topic, being structured on three conceptual dimensions that define the current recruitment process, using the following key concepts:

1. Dimension 1, technological (support technology): the terms ”AI” OR ”artificial intelligence” OR ”automated decision making” OR ”machine learning” were selected to include all forms of intelligent automation used in HR.

2. Dimension 2, functional (field of applicability): The use of the group ”recruitment” OR ”talent acquisition” OR ”hiring” OR ”human resources” refers to anchoring the research in the organizational environment. Since the candidate selection process is often treated either as an isolated function (recruitment/hiring) or as part of an integrated management strategy (human resources/talent acquisition), this selection of terms allowed for the capture of the entire data flow circulating through recruitment systems.

3. Dimension 3, risks and security (critical issues): this is the central dimension of the article, where the terms ”cybersecurity” OR ”ethics” OR ”bias” were used to highlight the duality of risks: technical risks, represented by cybersecurity, aiming at data integrity and system resilience in the face of attacks; and algorithmic risks, represented by ethics and bias, aiming at processing errors and unintentional discrimination. The justification for grouping these terms using the OR operator resides in the fact that mitigating risks in AI recruitment requires a hybrid approach: cybersecurity protects the data, while ethics and the elimination of bias protect the fairness of the process.

The Boolean intersection (synthesis) obtained by using the AND operator between these dimensions is essential for the validity of the sample. This logic was applied to exclude ”informational noise” (e.g., articles regarding AI in medicine or cybersecurity in the banking system). Thus, the resulting articles do not merely represent isolated studies on technology or HR, but constitute the body of knowledge located exactly at the intersection of AI-driven recruitment errors and the security mechanisms necessary for their mitigation.

The analysis aimed to identify thematic clusters to observe how security and ethics (bias) concepts are integrated into AI-based recruitment techniques. The selection of search terms was designed to ensure a triangulation between technology, the field of application and control mechanisms. The intersection of these terms reflects exactly the logical structure of the article: the problem - the errors (identified through Bias and Ethics) that occur in the selection process (Recruitment) due to technology (AI); the solution - the mitigating role of technical mechanisms (identified through Cybersecurity).

The process of establishing the document corpus the following refinement stages: 1. Initial search - the intersection of the three dimensions generated a raw set of 1558 articles. 2. Temporal filtering - given the accelerated dynamics of AI technologies and the need to analyze the most recent risk mitigation solutions, a temporal filter was applied for the period 2020–2026. 3. Final sample - following the filtering process, 1338 scientific articles were retained, representing the database for the bibliometric analysis itself.

The visual and structural analysis was performed using the VOSviewer software. This tool was used to generate and visualize co-occurrence networks of author-provided keywords and terms extracted from titles/abstracts.

The fact that 1338 results were obtained after temporal filtering (2020-2026) demonstrates that the topic represents a major and current concern within the scientific community. This volume of data allows for a high-resolution bibliometric analysis, capable of highlighting how cybersecurity evolves from a technical necessity into a guarantee of equity in modern

recruitment. We selected these terms to demonstrate that risk mitigation cannot be achieved through ethical regulations alone but requires a technical security foundation (cybersecurity), and the bibliometric analysis of 1338 articles serves as evidence for this need for integration.

## 4. Data Analysis and Interpretation

The visual analysis of the co-occurrence network resulting from the use of VOSviewer software (Figure 1) reveals a complex polycentric structure, organized into 9 thematic clusters, with a variable degree of density (ranging from 37 to 12 items per cluster) and a total of 230 items extracted from the 1338 articles (2020–2026).

The size of the nodes and the density of the connections indicate a clear hierarchy of academic interest, in which artificial intelligence represents the gravitational core, while cybersecurity acts as an emerging field of control. Regarding the conceptual hierarchy and thematic dominance, it is observed that the map is dominated by a massive core, the artificial intelligence node (red), which is the most prominent term and occupies the central position. The absolute centrality of the Artificial Intelligence node (red), located at the intersection of all axes, confirms that the literature no longer treats AI merely as a tool or an emerging technology, but as the standard infrastructure of recruitment.

The distribution of the 1338 articles across 9 thematic clusters, with a clear hierarchy from ”core” to ”periphery” demonstrates that the field is reaching full maturity but remains fragmented between the ethical and technical sides; the fragmentation of the clusters indicates a disconnection between ethical discourse and the implementation of technical security.

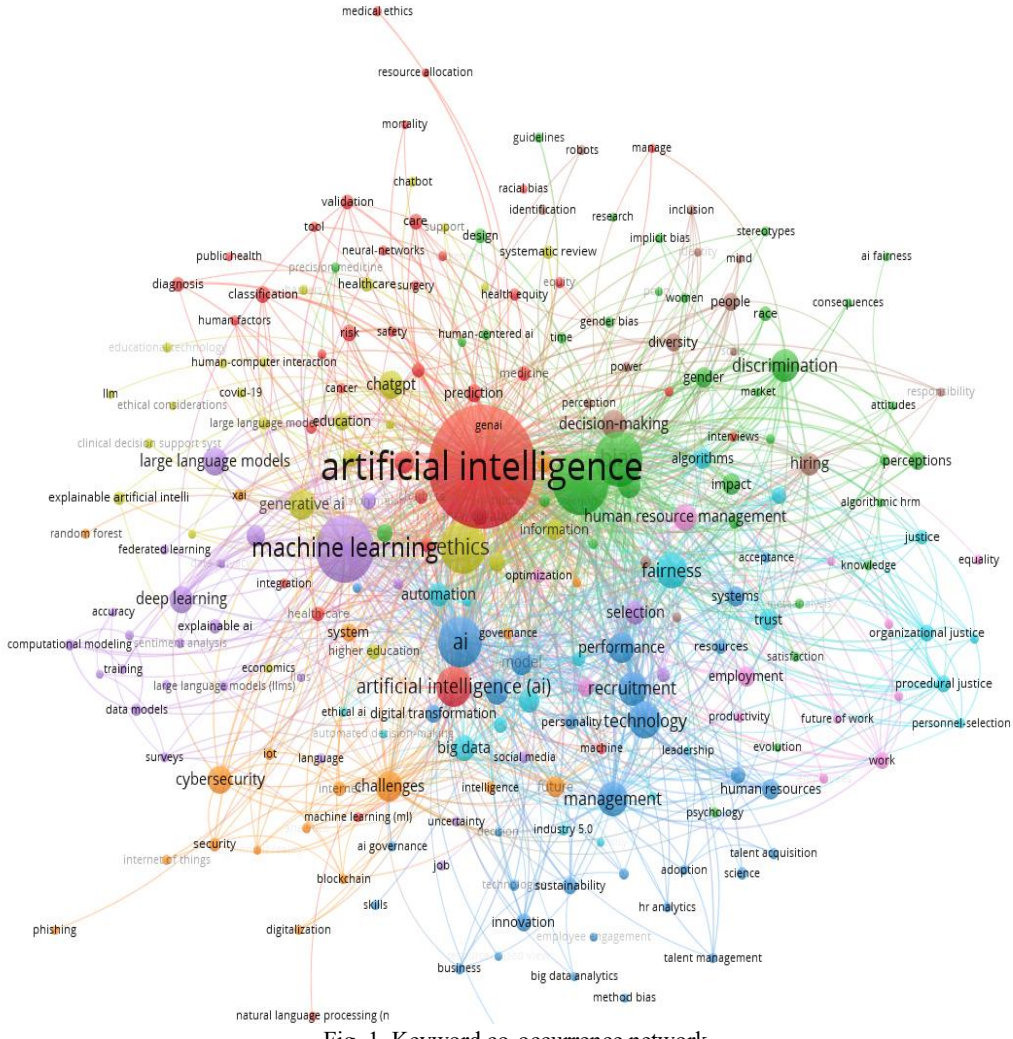



Fig. 1. Keyword co-occurrence network
*Source: own processing with VOSviewer*

For an in-depth understanding of the network density, Table 1 synthesizes the distribution of the 9 clusters based on the number of items and their role within the general theme. A direct correlation is observed between research volume (number of items) and conceptual maturity: while the technological core (C1) and the algorithmic error core (C2) total over 70 items, the pillar dedicated to cybersecurity (C7), although essential for risk mitigation, remains a consolidating field with 20 items. This distribution highlights the research gap that the present study aims to address. Clusters 1, 2, and 3 form the ”backbone” (containing the highest number of articles). Cluster 7 (Cybersecurity) is precisely our point of interest: even though it contains 20 items (nearly half compared to Cluster 1), it is robust enough to be considered a stable trend rather than a mere coincidence.

Table 1. Cluster distribution by number of items, keywords and themes

| Cluster | Colour | Items | Dominant terms (Major nodes) | Theme and conceptual role |
|---|---|---|---|---|
| C1 | red | 37 | artificial intelligence (AI) | The technological core: the computational foundation of recruitment systems and the primary source of risks. |
| C2 | green | 36 | bias, discrimination, gender | Algorithmic risks: identifying errors and negative social impact. |
| C3 | blue | 33 | recruitment, management, performance, technology | Operational implementation: integrating AI into HR processes and their efficiency. |
| C4 | yellow | 29 | ethics, generative AI, chatGPT, information | The current ethical dimension: moral discourse on emerging technologies. |
| C5 | purple | 26 | machine learning, large language models, deep learning, selection | Technical evolution: the shift toward complex automated selection models. |
| C6 | turquoise | 21 | fairness, big data, trust | Standardizing equity: the search for fairness methods in Big Data |
| C7 | orange | 20 | cybersecurity, challenges, security, system | Risk mitigation: cybersecurity as a barrier and technical challenge (the missing link: technical error mitigation) |
| C8 | brown | 16 | hiring, decision making, diversity, people | The decision-making act: human responsibility and diversity in recruitment |
| C9 | pink | 12 | human resource management, employment, work | The institutional framework: strategic workforce management in the digital era, general labor regulatory framework |

*Source: own processing.*

The hierarchical analysis of the nodes highlights three primary pillars of knowledge, detailed as follows. The Advanced Technological Pillar (represented by red and purple) is dominated by artificial intelligence, machine learning, large language models (LLM) and deep learning. The concentration of these terms indicates a research orientation toward “black-box” models, where process complexity increases the risk of difficult-to-detect algorithmic errors. This automated selection (purple) signifies a technological evolution that creates the very complexity generating the ”errors” referenced in this work's title.

The Strategic and Management Pillar (represented by blue and pink): the terms human resource management, recruitment, technology and performance reflect the functional integration of AI within organizations. It is noteworthy that human resource management is perceived both as a global strategy (pink) and as a set of organizational processes focused on efficiency (blue).

The Ethico-Social Pillar (represented by green, yellow, and turquoise) refers to the area dedicated to recruitment errors. Large nodes such as bias, ethics, discrimination, fairness, gender, impact demonstrate that the primary concern of current literature is the subjective and potentially discriminatory nature of algorithms. Notably, human resource management

appears both in Cluster 2 (green – 36 items), closely linked to bias and discrimination and in Cluster 9 (pink – 12 items). This separation shows that, in contemporary literature, human resource management is more concerned with the negative consequences of AI (algorithmic errors) than with its strategic integration.

A significant result of the present research is the analysis of Cluster 7 (orange, with 20 items), which groups the terms cybersecurity and challenges. Although cybersecurity is a visible and well-defined node, it is situated in the lower half of the size hierarchy, significantly outweighed by bias and ethics. This size disparity is significant to our interpretation, indicating an implementation gap regarding cybersecurity in mitigating AI-based recruitment errors within organizations. More precisely, current literature is extremely rich in identifying errors (bias), but far more reserved in proposing technical protection solutions. The fact that cybersecurity shares a cluster with challenges indicates that cyber defense is not yet viewed as a routine component in HR, but rather as a complex technical barrier to be overcome. A visual discrepancy is observed between the size of the bias node (large) and the cybersecurity node (small-to-medium). This suggests that while risks are identified, technical mitigation mechanisms through cybersecurity are only in a consolidation phase; specifically, contemporary literature extensively documents errors but offers few security-based mitigation solutions. Although cybersecurity is a visible node, it sits at a considerable distance from the recruitment (blue) and hiring (brown) nodes, a visual distance that shows cybersecurity is perceived as an external challenge (challenges) rather than an internal control instrument for the selection process.

The emergence of the terms generative AI, ChatGPT (yellow) and large language models (purple) among the most visible circles indicates a paradigm shift during the 2020–2026 period: as recruitment transitions from simple algorithms to generative models, security risks (e.g. training data integrity) become much more sophisticated. The positioning of the terms decision making and hiring (brown) at the edge of the central core suggests that decisional accountability requires a more solid anchoring in security protocols to ensure effective risk mitigation.

Recruitment errors (the central subject of this paper) are distributed across three key clusters: Cluster 2 (green), which focuses on the direct manifestation of errors (bias, discrimination), is the second-largest cluster (36 items), demonstrating that the social impact of recruitment errors is the primary concern of researchers; and in Clusters 4 (yellow) and 6 (turquoise), which introduce the modern nuances of errors through generative AI, ChatGPT and the need for fairness. The presence of ChatGPT and generative AI as visible nodes indicates an escalation of error risks alongside the adoption of large language models.

## 5. Research Conclusions

The co-occurrence map generated via VOSviewer reveals a dense structure with marked centrality; however, it also highlights significant peripheral areas that indicate a lack of security integration within HR processes. The fact that large nodes exist for technology and ethics, while cybersecurity occupies a smaller position, demonstrates that security is the

missing piece required to transition from the awareness of errors to their practical resolution.

The prominence of the bias and discrimination nodes confirms that errors in AI-assisted recruitment are not isolated incidents, but rather a mature and urgent research theme. The presence of cybersecurity (orange), despite being smaller in scale than the ethical pillar, demonstrates a "security deficit' in HR processes. There is a clear awareness of the problems, yet a lack of integration regarding technical security solutions within the functional workflows of HR.

While fairness is predominantly addressed at a theoretical level, from philosophical and social perspectives, this article proposes a technical solution: cybersecurity as a pragmatic method for mitigating the risks identified on the map, thereby bridging the gap between ethical intent and technical reality. The big data node is linked to fairness, creating a paradox: the graph suggests that researchers hope to achieve equity through the analysis of large datasets, overlooking the fact that massive data is precisely what is most exposed to cyberattacks capable of manipulating recruitment outcomes.

Although artificial intelligence (red) acts as a "central gravitator", the visual hierarchy reveals a "fracture" between technology and risk control:

- the development axis (center-left): The close link between artificial intelligence, machine learning (purple) and large language models (purple) indicates a purely technical research trend centered on performance. Their proximity suggests these themes are discussed "as a block", as a self-sustaining ecosystem.

- the social-ethical axis (center-top right): The discrimination (green) and bias (green) nodes are large and situated in the upper section. Their positioning relative to the "hard core" of deep learning suggests that discrimination issues are treated more as sociological effects and less as technical vulnerabilities that require security patches.

Cybersecurity remains the "great absentee" within the HR core. In the orange cluster, security is grouped with challenges and digitalization, suggesting that cyber defense is perceived more as an administrative burden than as a real support mechanism. Although generative AI and ChatGPT (yellow) are prominent nodes relatively close to the red center, they lack thick link lines to cybersecurity. This indicates a major risk: we are utilizing generative AI in recruitment without having security protocols to verify the authenticity or neutrality of these models' outputs.

The results obtained through VOSviewer validate the necessity of our research. The hierarchy of nodes (Bias > Cybersecurity) demonstrates that while the academic community understands the risks and describes the errors, there is an implementation gap regarding their mitigation; we are only at the beginning of building the security "shield". The prominence of the bias node (green) shows that errors are recognized as intrinsic to AI systems in HR. However, their grouping with ethics (yellow) suggests that remediation is

sought in philosophical principles, which are often difficult to quantify technically. By positioning the cybersecurity node on the same plane as challenges, the data suggests that transforming cybersecurity from a "challenge" into a "core solution" is the necessary step for bias mitigation. This article proposes using the tools from Cluster 7 (Security) to correct the vulnerabilities identified in Clusters 2 and 4 (Bias and Ethics).

The considerable visual distance on the map between cybersecurity and hiring or decision-making shows that, currently, hiring decisions are made without considering the cyber integrity of the system providing the data. The bibliometric analysis highlights the gap between the awareness of algorithmic errors (green/yellow clusters) and the implementation of cybersecurity mechanisms (orange cluster), supporting the need for an integrated approach to risk mitigation in HR and a transition from passive ethics to active security. Our work proposes "pulling" the cybersecurity node from the map's periphery toward the decisional center (decision-making). Only through this junction can we mitigate recruitment errors, transforming an abstract discussion about ethics into a technical safety procedure.

## 6. Integrating Cybersecurity into AI-Assisted Recruitment

Since the results of the bibliometric analysis confirm a fracture between technological performance and risk control, it becomes imperative to articulate a technical mediation framework based on cybersecurity, defined as a sum of rules, technologies, processes, and controls to protect systems, networks, programs, devices and data from cyberattacks [15]. The model presented hereafter materializes this vision, transforming cybersecurity from a defensive function into an active layer for protecting data integrity and decisional equity.

The proposed model extends the classic AI-assisted recruitment pipeline by introducing a security overlay, applied throughout the entire data and model lifecycle. The cybersecurity level should be addressed on a case-by-case basis, this model positions it as a continuous mechanism for controlling integrity and fairness. However, we can anticipate that, at the current state of the art, no algorithm can be considered fully autonomous without human supervision [16].

### *Proposed architecture*

From an architectural perspective, effective risk mitigation requires adopting the security by design principle, in which cybersecurity is integrated into all stages of the recruitment pipeline. This includes data collection, preprocessing, model training, and decision-making. Relevant technical measures include access control, data encryption, logging and continuous auditing of decisions, and avoiding non-transparent (black-box) decisions [14]. These mechanisms help ensure the integrity and traceability of the system, facilitating the identification and correction of errors. Cybersecurity provides these mechanisms, allowing the transformation of abstract principles into operational solutions.

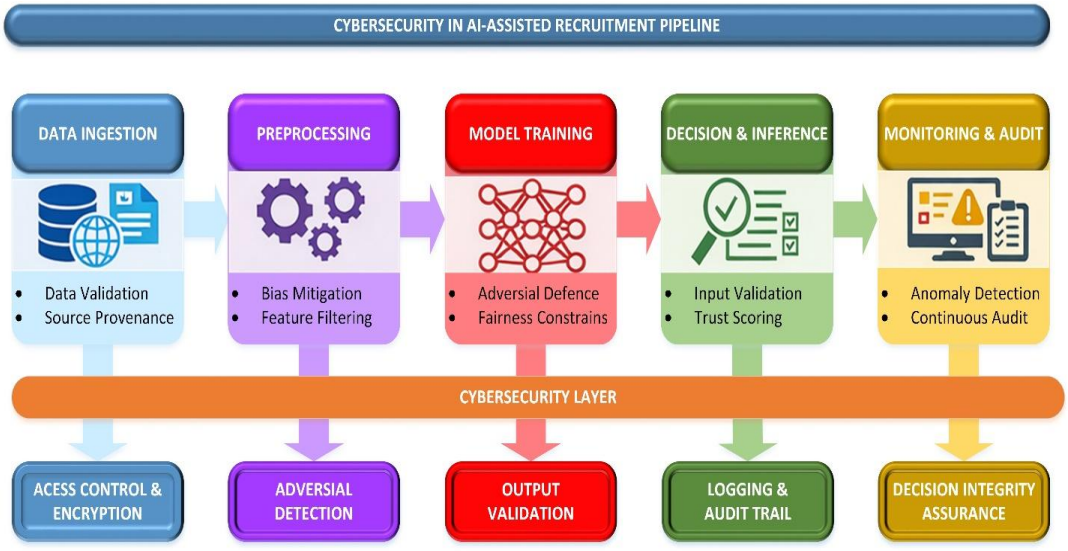



Fig. 2. Flow chart for cybersecurity in AI-assisted recruitment
*Source: own processing*

The proposed model is structured on five operational levels:
**(1) Data Ingestion Layer**
Sources: CVs, job boards, internal databases.

Risk: compromised or manipulated data (data poisoning).

Control: syntactic and semantic validation, data provenance verification, automatic filtering.
**(2) Preprocessing and feature engineering (Processing Layer)**
Risk: amplification of existing bias in the data [17].

Control: normalization techniques, removal of sensitive attributes, statistical audit of distributions.
**(3) Model training (Model Training Layer)**
Risk: learning of discriminatory or maliciously injected correlations.

Control: adversarial training, cross-set validation, fairness constraints [18].
**(4) Inference and decision (Decision Layer)**
Risk: incorrect or manipulable decisions in real time.

Control: validation of inputs, confidence scoring, explicit business rules.
**(5) Monitoring and audit (Monitoring & Audit Layer)**
Risk: degradation of performance (model drift) or emergence of new types of bias.

Control: logging, continuous auditing, anomaly detection, model recalibration [19].

***Cybersecurity cross-layer***
This layer includes mechanisms applied at all levels: access control and data encryption; auditability (traceability); adversarial attack detection; output validation for generative systems.

Automation systems based on artificial intelligence can restructure processes in cyberspace only to reduce manual effort and increase efficiency, but continuous verification of these processes must be taken into account [20].

The model's major contribution resides in the fact that it transforms security from a defensive function into an active mechanism for reducing bias and guaranteeing the integrity of decisions. The concepts presented can also be analyzed in the table below.

Table 2. Correlating risks with security controls

| Risk category | Technical description | Cybersecurity control | Impact on bias and decision |
|---|---|---|---|
| Data Poisoning | Malicious data injection into training sets | Data validation, data provenance, anomaly detection | Reduces artificially induced bias and increases dataset integrity |
| Historical bias | The data reflects existing discrimination | Fairness constraints, dataset rebalancing | Decreases systemic discrimination |
| Adversary attacks | Manipulated inputs to induce errors | Adversarial training, input validation | Increases the robustness of real-time decisions |
| Lack of auditability | Non-transparent decisions (black-box) | Logging, audit trail, explainability tools | Allows identification and correction of bias |
| Unvalidated generative output | Potentially erroneous or biased AI content | Output validation, cross-model verification | Reduces the propagation of errors in decision making |

*Source: own processing*

In conclusion, mitigating errors in AI-assisted recruitment cannot be achieved solely through ethical or statistical approaches. Deep integration of cybersecurity into the systems architecture is required to ensure data integrity, model robustness, and decision correctness. Depending on the use case, distorted performance can lead to serious social equity and security issues [21].

All algorithms make mistakes, but an algorithm is considered biased if it systematically makes incorrect predictions about one group or another. This bias can be caused by the design of the algorithm itself or by the biased data on which the algorithm is trained [22]. On the other hand, data integrity must be constantly verified.

## 7. Discussions

The bibliometric analysis and the proposed architectural model provide the theoretical foundation for reconciling the identification of ethical errors with the implementation of technical control solutions. This approach moves beyond the fragmentation observed in current literature by integrating the principle of security by design directly into the recruitment workflow. This conceptual convergence demonstrates that mitigating algorithmic bias and discrimination cannot remain a mere philosophical desideratum; instead, it must evolve into an approach rooted in technical resilience. Consequently, the

transition from passive ethics to active security transforms cybersecurity from an administrative barrier, often perceived as an external challenge, into a structural pillar, essential for guaranteeing data integrity and full equity in hiring decisions.

Thus, by ”pulling” the cybersecurity node from the periphery of the bibliometric map toward the decisional center, the present research offers a pragmatic solution for the "security deficit" observed in human resources. This convergence between ethics and technology marks the necessary transition from the simple identification of algorithmic errors to their active mitigation through technical resilience.

## References


[1] O. Kravchuk, “Human-centric digital leadership for HRM 5.0 ecosystem readiness”. Smart Cities and Regional Development Journal, vol. 10, nr. 1, 55-85, 2026. DOI: https://doi.org/10.25019/5ykgm661

[2] A.-C. Alexandru, “Building a smart economy: Training the workforce in multi-source data and AI-enhanced economic analysis”. Smart Cities International Conference (SCIC). Vol. 12: Resilient Communities Empowered by Collective Intelligence. 265-283, 2025.

[3] A. Hunkenschroer, C. Luetge, “Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda”. Journal of Business Ethics. 178:977–1007, 2022. https://doi.org/10.1007/s10551-022-05049-6

[4] A. Kaplan, M. Haenlein, “Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and impli-cations of artificial intelligence”. Business Horizons, 62(1), 15-25, 2019. https://doi.org/10.1016/j.bushor.2018.08.004

[5] E. T. Albert, “AI in talent acquisition: a review of AI-applications used in recruitment and selection”. Strategic HR Review, 18(5), 215-221, 2019. https://doi.org/10.1108/SHR-04-2019-0024

[6] P. Tiwari, M.S. Natal, “The Role of AI in Talent Acquisition: A Study of Recruitment Processes in IT Companies”. International Journal of Research Publication and Reviews, Vol 6, Issue 5, pp 2215-2221, 2025. https://doi.org/10.5281/zenodo.15354623

[7] P. Tambe, P. Cappelli, V. Yakubovich, “Artificial Intelligence in Human Resources Management: Challenges and a Path Forward”. California Management Review, 61(4), 1-28, 2019. https://doi.org/10.1177/0008125619867910

[8] A. M. Votto, R. Valecha, P. Najafirad, H. R. Rao, “Artificial Intelligence in Tactical Human Resource Management: A Systematic Literature Review”, International Journal of Information Management Data Insights, Volume 1, Issue 2, 100047, 2021. https://doi.org/10.1016/j.jjimei.2021.100047.

[9] M.F. Gonzalez, W. Liu, L. Shirase, D. Tomczak, “Allying with AI? Reactions toward human-based, AI/ML-based, and augmented hiring processes”. Computers in Human Behavior 130(4):107179, 2022. DOI: 10.1016/j.chb.2022.107179

[10] A. Köchling, M. C. Wehner, “Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR”. Business Research, 13(3), 795–848, 2020. https://doi.org/10.1007/s40685-020-00134-w

[11] M. Raghavan, S. Barocas, J. Kleinberg, K. Levy, “Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices”. FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability and Transparency. 469-481, 2020. https://doi.org/10.1145/3351095.3372828

[12] D. Pessach, E. Shmueli, “A Review on Fairness in Machine Learning”. ACM Computing Surveys, 55(3), 1-44, 2022. https://doi.org/10.1145/3494672

[13] I. Ajunwa, “The “black box” at work”. Big Data & Society. Volume 7, Issue 2, 2020. https://doi.org/10.1177/2053951720938093

[14] Z. Chen, “Ethics and discrimination in artificial intelligence-enabled recruitment practices”. Humanities and Social Sciences Communications 10, 567, 2023. https://doi.org/10.1057/s41599-023-02079-x

[15] C. Sîrbu, “Cybersecurity, the challenge of the present”. Machine Intelligence & Security for Smart Cities (TRUST). Vol. 2: Ensuring Trust and Security in Intelligent Urban Ecosystems, 25-28, 2025.

[16] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, M. Marchetti, “On the effectiveness of machine learning for cybersecurity”, 10th International Conference on Cyber Conflict (CyCon), 2018. DOI:10.23919/CYCON.2018.8405026

[17] D. R. Chaudhari, “Bias in AI Recruitment Systems: An Ethical Evaluation of Algorithmic Hiring Tools” ResearchGate, 2025. DOI:10.13140/RG.2.2.11411.80163

[18] C. Zhang, S. Li, Y. Liu, Y. Li, J. Zhu, "Artificial Intelligence Recruitment - A Literature Review Based on Equity and Efficiency Perspectives", Journal of Human Resource Development, Vol 7, Issue 1, 2025, DOI: 10.23977/jhrd.2025.070103.
[19] U. Mmaduekwe, "Bias and Fairness Issues in Artificial Intelligence-driven Cybersecurity", Current Journal of Applied Science and Technology, Volume 43, Issue 6, 109-119, 2024, 10.9734/cjast/2024/v43i64391
[20] European Commission, EU AI Act, 2024, https://eur-lex.europa.eu/eli/reg/2024/1689/oj
[21] Y. Xu, P. Terhörst, M. Pedersen, K. Raja, "Analyzing Fairness in Deepfake Detection With Massively Annotated Databases" , IEEE Transactions on Technology and Society, PP(99):1-1, 2024, DOI:10.1109/TTS.2024.3365421
[22] E. Ip, "Fair AI in hiring: Experimental evidence on biased hiring algorithms and debiasing methods affect the quality and diversity of applicants", Behavioral Science & Policy 11(1), 2025, DOI:10.1177/23794607251353585