Received: 13 July 2025 | Revised: 16 January 2026 | Accepted: 1 April 2026 | Published online: 16 April 2026

REVIEW

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–10
DOI: 10.47852/bonviewAIA62026783

BON VIEW PUBLISHING

Ethical Frameworks in AI Decision-Making: Comprehensive Analysis of Management and Human Resource Applications in Contemporary Organizations

Kirill Toropov[1*] and Irina Manakhova[1,2]

[1] *Department of Management, Peoples' Friendship University of Russia named after Patrice Lumumba, Russia*

[2] *Department of Political Economy, Lomonosov Moscow State University, Russia*

Abstract: This literature review examines emerging ethical frameworks for artificial intelligence (AI) decision-making in management and human resource (HR) contexts within contemporary organizations. Through a systematic review of academic literature published between 2021 and 2025, the study synthesizes the theoretical foundations, implementation approaches, and practical challenges associated with the adoption of ethical AI in organizational settings. The analysis focuses on how organizations conceptualize and operationalize ethical principles in AI-driven decision-making. The review shows that most ethical AI frameworks are grounded in core principles such as transparency, fairness, accountability, and respect for human dignity. However, organizations face difficulties in translating these normative principles into operational practices. The findings identify three main barriers to implementation: algorithmic bias, data privacy concerns, and substantial financial costs, often estimated to range from $50,000 to $500,000 for implementation and monitoring. Based on the analysis of 14 academic sources, the study indicates that while organizations often prioritize regulatory compliance when developing ethical AI policies, effective implementation requires broader organizational and cultural transformation. The review integrates theoretical perspectives with examples from industry practices, including emerging metrics used to evaluate ethical AI implementation. The findings suggest that existing ethical AI frameworks must move beyond traditional compliance-oriented approaches as new forms of human–AI interaction reshape organizational decision-making. This is particularly relevant for HR processes such as recruitment algorithms, performance evaluation systems, and workforce management platforms. The study highlights the need for adaptive ethical AI frameworks capable of addressing technological change, evolving regulations, and societal expectations while maintaining organizational efficiency.

Keywords: artificial intelligence ethics, decision-making frameworks, human resource management, algorithmic bias, AI transparency

1. Introduction

The introduction of artificial intelligence (AI) technologies into business decision-making has led to fundamental changes in the field of human resource (HR) selection and management. While this technological breakthrough significantly expands the capabilities of data analysis and predictive behavior modeling in various organizational settings, it also raises new ethical issues that require timely and comprehensive consideration. Modern organizations balance the promise of technological advancement with the constraints of ethical standards. The key challenge in these circumstances is finding optimal solutions that ensure ethical AI use without compromising operational efficiency or undermining the company's competitive advantage. The relevance of this dilemma becomes increasingly apparent given the rapid application of AI in management—from automated and advanced candidate selection systems capable of processing thousands of applications using complex algorithmic reasoning to selection and performance assessment algorithms that will make critical decisions about employee career development and organizational advancement opportunities. AI technologies are permeating virtually every aspect of modern organizational life, exerting a profound impact on a wide range of stakeholders. However, the pace of technological advancement significantly outpaces established, and sometimes rudimentary, ethical norms and associated limitations in the modern world, forcing organizations to continually develop their own ethical principles applicable to the responsible use of AI. At the same time, companies must overcome previously unimaginable ethical and legal challenges. Recently, scientific discourse has begun to fill this critical gap, proposing conceptual frameworks for systematizing AI-based decision-making in organizational contexts and clarifying the

***Corresponding author:** Kirill Toropov, Department of Management, Peoples' Friendship University of Russia named after Patrice Lumumba, Russia. Email: 1142240366@rudn.ru

ethical risks associated with algorithmic governance [1, 4]. Many of these frameworks converge around core values such as transparency, justice, accountability, explicability, and the protection of human dignity, although they differ in how these principles are operationalized across sectors and jurisdictions [2, 3]. In particular, recent studies emphasize the importance of addressing algorithmic bias, ethical HR decision-making, worker dignity, and governance mechanisms in the implementation of AI systems in organizations, especially in HR-related contexts [5–8]. Although much work remains to be done in translating traditional philosophical principles into practical organizational actions, these issues are particularly relevant in the field of management. Large organizations involve multiple stakeholders, each with different interests, perspectives, and priorities regarding the use of AI technologies, which makes ethical governance dependent not only on technical controls but also on managerial judgment and leadership responsibility [9, 10].

In today's business environment, organizations must balance the use of technological innovations where necessity does not trump these specific obligations, but these, in turn, do not harm the competitive advantage and the company's performance. This contradiction is even more acute in HR management, where AI-assisted decisions affect not only individual managers but also organizational culture, employee justice perceptions, and broader employment practices [6–8]. The stakes associated with advanced approaches require that organizations take some of these aspects seriously, and not just the technical capabilities and expected business benefits.

2. Literature Review

The literature on ethical AI decision-making combines insights from moral philosophy, management theory, and organizational practice. Ferrell and co-authors argue that ethically justifiable AI cannot be derived from engineering logic alone and instead requires an integrated framework combining moral reasoning, regulatory awareness, and organizational culture [9]. A similar management-oriented perspective is developed by Manda et al., who show that leadership decisions around AI ethics must balance innovation goals, institutional responsibility, and stakeholder trust [10].

In the architecture formalized by Ferrell and co-authors, transparency turns out to be an impregnable keystone. They express the argument that the AI processes should match the recommended explainability provisions, allowing those impacted by the automated decisions to follow, in human senses, the paths of thought that lead to the algorithmic decision-making. This indicator of dignity preservation in the context of AI-driven decision-making is a valuable learning for organizations that implement the concepts of ethical frameworks in HR-related contexts, not just with regard to the concepts of fairness but also with regard to the necessity to honor human agency, individual worth, and complexity of human factor which governs the relationships in the workplace and career development. Organizations that successfully employ the precepts of dignity in their AI systems experience higher levels of employee trust and more exposure and interest in technological innovations and are generally happier with the way decisions are made in their organizations [8, 9].

The analysis shows that the effective application of HR AI should be based on a balance between technological efficiency and human-centered approaches that will not eliminate employee agency or trust relations that are required to guarantee the high-performing organizational processes.

Turning the abstract ethics into functional operations entails numerous critical challenges to organizations that aim to develop an efficient ethical AI setup. Farayola and Olorunfemi identify several recurrent barriers to the implementation of ethical AI frameworks, including technical complexity, resource constraints, organizational resistance to change, and cultural barriers to AI acceptance [11]. These problems span across different levels of the organization and have to be addressed technically and culturally in terms of ethical AI application.

Cost remains one of the most significant barriers to the ethical implementation of AI in both large and small organizations, especially when ethical governance requires auditing procedures, employee training, monitoring systems, and ongoing adaptation rather than onetime deployment [11, 12]. Developing sophisticated tools to detect misleading AI, implementing comprehensive transparency regulations, training employees in the use of ethical AI, and monitoring and evaluation systems can typically be expensive for organizations, and this can create unequal access to and application of ethical AI, possibly disadvantaging smaller organizations that lack the resources available to larger technology companies and other multinational organizations, which can afford them.

Brendel et al. propose that sustainable ethical AI management depends on gradual implementation, broad stakeholder involvement, and continuous monitoring mechanisms capable of evolving with organizational needs and technical constraints [12]. Their model emphasizes the following: the process of implementation strategies must be gradual, the process of stakeholder engagement must be thorough, and the monitoring system must be built on the basis of continuous improvement and the alteration of the monitoring systems based on the needs and technical resources of the organization. This approach officially recognizes that AI ethical implementation is not a project with a distinct start and end but rather a long organizational commitment that must be maintained with a focus, resources, and plan over extended time periods. The main implementation challenges and estimated cost ranges associated with ethical AI adoption in organizations are summarized in Table 1.

The constantly changing regulatory landscape of AI ethics offers potential for standardization and a parallel challenge to organizations that operate in diverse jurisdictions that are of varying complexity. Kaur critically analyzes responsible AI governance and shows that organizations must remain flexible as legislative requirements evolve across industries and jurisdictions [13]. The introduction of extensive regulations such as the European Union AI Act introduces new compliance costs that should be considered when enhancing an ethical model of an organization but should not limit its operations and adopt a competitive position.

Attard-Frost et al. review 47 AI ethics guidelines and demonstrate that, although fairness, transparency, and accountability recur across frameworks, implementation depth and enforcement mechanisms vary substantially across documents and jurisdictions [14]. Indeed, most of the guidelines addressed systematically core principles such as fairness, transparency, and accountability though there is a great discrepancy in how or to what extent it deals with implementation, enforcement, and any practical guidance that organizations should adopt. These contradictions are serious hurdles to organizations operating beyond jurisdictions or attempting to learn best practices elsewhere and still make internal policies and procedures consistent.

Industry-specific factors also play important roles in influencing their ethical AI structuring since various industries tend to

Table 1
Implementation challenges and cost analysis

Challenge category	Implementation cost range	Primary barriers	Success factors
Bias mitigation	\$75,000–\$200,000	Technical complexity, data quality	Third-party audits, diverse datasets
Transparency systems	\$50,000–\$150,000	Algorithm complexity, user understanding	XAI tools, clear communication
Privacy protection	\$100,000–\$300,000	Regulatory compliance, data security	Encryption protocols, access controls
Training programs	\$25,000–\$100,000	Skill gaps, resistance to change	Continuous learning, certification
Monitoring systems	\$60,000–\$180,000	Resource allocation, technical expertise	Automated auditing, human oversight

focus on various issues of ethical application with regard to their operating needs, regulatory systems, or stakeholder demands. Baker-Brunnbauer shows that management perspectives on AI ethics differ from purely technical or academic approaches because firms must incorporate legal exposure, operational efficiency, competitive strategy, and stakeholder relations into their ethical choices [15]. Baker-Brunnbauer demonstrates the primacy of practical considerations, including legal liability, comparative advantage sustained outcome, operational efficiency requirement, and stakeholder relationship required practices. These management perspectives prove essential for developing frameworks that achieve both ethical soundness and organizational viability within competitive business environments. Ethical implications of AI adoption have also been explored in knowledge management and broader decision-making practice, where scholars highlight the role of organizational learning, managerial interpretation, and the practical consequences of AI-assisted decisions [16, 17]. Sector-specific studies further demonstrate that ethical risks vary considerably across education, finance, and healthcare environments, where decision stakes, data sensitivity, and professional norms differ substantially [18–20].

Modern methods of quantitative assessment of AI ethics have begun to emerge in multiple fields, but existing limitations remain for their use in organizational management. Current fairness metrics are not lacking technical sophistication, but the practical design limitations of existing fairness measures are prevalent in complex HR decision-making contexts where multiple stakeholder groups hold different fairness expectations or where cultural norms may affect the fairness perceptions of similar treatment.

Existing fairness measurement metrics rely heavily on statistical parity measures such as demographic parity, equalized odds, and individual fairness, as implemented in widely used toolkits such as IBM Fairness 360. These statistical parity measures have clear limitations in HR contexts, where individuals' fairness perceptions are determined by procedural norms rather than statistical comparisons. Demographic parity metrics may be mathematically rigorous but may also violate merit selection criteria critical to the assembling of an effective organizational team. Equalized odds measures are designed to create false positive rates and false negative rates that are the same across groups and do not align with the fairness expectations individuals have for promotion and compensation decisions.

Recent advances in individual fairness metrics have attempted to circumvent procedural fairness design limitations through similarity-based processes by examining whether similar individuals receive similar treatment. However, this faces significant definitional challenges of what "similarity" might be appropriate in positions (and organizations) where job performance is dependent on a wide range of objective and subjective factors that cannot readily be measured to generate a straightforward metric.

Existing methods for measuring AI transparency are based on algorithms for technical explainability measures of model interpretability using quantitative perspectives of feature importance scores, Local Interpretable Model-Agnostic Explanations (LIME) explanations and Shapley Additive Explanations (SHAP) values. While these methods provide valuable insights for technical audiences, they demonstrate limited effectiveness in organizational contexts where nontechnical stakeholders require accessible explanations of automated decisions affecting their careers and workplace experiences.

The Google What-If Tool and similar interpretability platforms offer interactive exploration capabilities, yet their utility remains constrained by the technical expertise required for effective utilization. More alarmingly, while tools like these operate in the behavior explanation realm, they do not address transparency across the board, specifically the ability for employees to learn how data is collected and used in the decision-making process, what criteria were used for evaluation, and how to appeal decisions.

The Microsoft Fairlearn framework supports algorithmic fairness assessment through bias detection and mitigation techniques; however, its applicability is largely limited to technically labeled classification and regression problems. Therefore, it is unlikely that organizations using Fairlearn for AI-driven decisions can effectively apply it to the complex, multi-criteria, and context-specific decision-making processes typical of HR management. A trade-off is expected in terms of an organization incorporating a mitigation strategy on outputs from AI decision-making involving fairness and performance, and organizational leaders would find it very difficult to evaluate those trade-offs—without being inundated with performance-ranked decisions from AI systems with no concrete way to evaluate business impact or implications.

Amazon Sagemaker Clarify enables integrated activity to detect bias measurement and model explainability in a predetermined workflow of machine learning and can also generate automated documentation for auditing—functionality that is attractive to organizations that require it for compliance. However, the measurement in SageMaker Clarify is limited and primarily looks at technical quantifications of bias. It fails to plausibly look at ethical dimensions of action including preservation of dignity, procedural justice, and cultural sensitivity, which reflect upon the effectiveness of organizational decision-making.

All of these qualitative measurements are essential for successful organization-wide AI-based decision-making.

All aforementioned frameworks specified in this research take a post hoc notion and decisional model approach to evaluating process, rather than adopting a principal post hoc evaluation framework capable of evaluating observables throughout the decision-making process. Current frameworks look at re-evaluation based upon the AI system outputs being produced, with no foresight of integrating ethical considerations during actual and practical decision-making workflows when intervening could have removed conditions leading to the outcomes. In this case, the AI would not have been used as a punitive mechanism but as a criterion within an algorithm while allowing organizational leaders more evaluative freedom.

In addition to technical fairness and explainability metrics, recent research increasingly emphasizes the importance of integrating human-centered and governance-oriented perspectives into AI evaluation. Human-centered AI approaches highlight that ethical assessment should account for user empowerment, stakeholder understanding, and preservation of human agency rather than relying solely on statistical indicators [21]. Similar concerns are evident in high-stakes domains such as healthcare, where ethical AI deployment requires transparency, human oversight, and risk-sensitive decision-making frameworks [22].

Furthermore, the literature on algorithmic auditing proposes moving beyond isolated performance metrics toward lifecycle-based evaluation models. In particular, end-to-end auditing frameworks suggest embedding accountability and monitoring mechanisms throughout the entire AI system lifecycle, from design to deployment and post-decision evaluation [23]. From a theoretical standpoint, the ethics of algorithms cannot be reduced to technical fixes alone, as issues of opacity, responsibility distribution, and embedded value judgments are inherent to algorithmic systems [24].

Accordingly, organizational accountability remains a critical dimension of ethical AI implementation, requiring firms to justify not only outcomes but also the design logic and governance processes behind automated decision-making [25].

3. Methodology

The Ethical AI Impact Measurement (EAIM) framework is an innovative mixed method of assessing AI system ethics in organizational settings quantitatively, aimed at complying with reporting limitations from current models while introducing the capacity for real-time evaluation in management and HR applications. This method involves the integration of automated monitoring or assessment systems with multi-criteria decision analysis methods to facilitate a broader assessment framework that is adaptable to different organizational contexts and cultural expectations.

The EAIM framework comprises four interrelated modules that together allow for the provision of comprehensive ethical assessments, while being computable for real-time organizational use. The architecture adopts a modular design where the components can be implemented selectively based on organizational needs, aspirations, and resourcing limitations. These integrity modules would allow organizations to implement based on their priorities, but for optimal return on performance, all modules or components should be implemented together.

The **Fairness Assessment Module (FAM)** serves as the primary assessment module, where a distinct multidimensional fairness assessment goes beyond demographic parity to include procedural processes as well as cultural sensitivities. The FAM module takes decisions through performance inputs (each input will go through configurable filters to assess outcomes across the selected demographic dimensions of the value to the organization, organizational effectiveness).

The underpinning mathematical framework for the fairness assessment utilizes a transformed version of individual fairness to include group-level statistical parity. In any decision d affecting individual i, the fairness score $F(d,i)$ can be computed as:

$$\begin{aligned} F(d,i) = {} & \alpha \times IF(d,i) + \beta \times GF(d,i) + \gamma \times PF(d,i) \\ & + \delta \times CF(d,i) \end{aligned} \tag{1}$$

where IF represents individual fairness measured through similarity-based comparisons, GF captures group fairness through demographic parity analysis, PF evaluates procedural fairness based on process consistency, and CF incorporates cultural fairness adjustments based on organizational context. The coefficients α, β, γ, and δ are dynamically adjusted based on decision type and organizational priorities, with default values of 0.3, 0.25, 0.25, and 0.2, respectively, though these can be modified during system calibration phases.

The **Transparency Evaluation System (TES)** assesses and monitors the explainability of AI decisions via multiple assessment dimensions with an emphasis on both technical interpretability and stakeholder comprehension. While current transparency assessment measures focus predominantly on feature importance scores, the TES agencies' explanation quality is from stakeholders who were affected by the decision, rather than from technical audiences. The TES goes further by including mechanisms for stakeholder feedback, allowing for a continuous refinement of explanation value.

The transparency score T(d) for decision d incorporates four primary components:

$$\begin{aligned} \mathrm{T(d)} = {} & w_1 \times \mathrm{E(d)} + w_2 \times \mathrm{C(d)} + w_3 \times \mathrm{A(d)} + w_4 \\ & \times \mathrm{R(d)} \end{aligned} \tag{2}$$

where E(d) measures explanation completeness, C(d) evaluates stakeholder comprehension levels, A(d) assesses explanation accuracy, and R(d) quantifies response timeliness. The weighting coefficients (w_1 through w_4) adapt based on user feedback and decision criticality, with initial values set to 0.4, 0.3, 0.2, and 0.1, respectively.

The **Privacy Protection Mechanism (PPM)** uses differential privacy approaches, in conjunction with data minimization strategies, to enable a proper use of personal information protection throughout the evaluation. This system uses a unique lens on how to strive for a balance between privacy and utility while restricting the ability to obtain accurate assessments, while conforming to applicable regulatory standards across a myriad of disparate jurisdictions.

Privacy protection functions primarily through the mechanism of noise injection and data aggregation, with this model being controllable through a privacy parameter ε, with smaller values leading to stronger privacy guarantees although reduced accuracy in assessing. The privacy–utility balance is considered via:

$$\mathrm{P}(\varepsilon) = \max(\mathrm{U}(\varepsilon) - \lambda \times \mathrm{R}(\varepsilon)) \tag{3}$$

where $\mathrm{U}(\varepsilon)$ represents utility (assessment accuracy) as a function of privacy parameter ε, $\mathrm{R}(\varepsilon)$ measures privacy risk, and λ serves as the organization's privacy–utility trade-off parameter, typically ranging from 0.1 to 0.5 based on data sensitivity levels.

The **Human Dignity Preservation Subsystem (HDPS)** is perhaps the most innovative component of the system because it attempts to quantify and monitor the preservation of human dignity throughout an AI-mediated decision process. This subsystem evaluates decisions based on entities' respect for human agency, recognition of individual value, and maintenance of a meaningful human presence in processes that could impact personal outcomes.

The core EAIM evaluation algorithm then organizes the organizational decisions by processing through a highly sophisticated multi-criteria assessment in which outputs from all four subsystems are combined to assign comprehensive ethical impact scores. The algorithm is able to process individual decisions, compare them to previous outputs, and assess aggregate patterns of decisions to identify systemic ethical considerations that may not have been disclosed through the individual decision assessment process.

The integrated ethical score S for decision d is calculated using:

$$cS(d) = \Sigma_i w_i \times M_i(d) \times C_i(d) \tag{4}$$

where $M_i(d)$ represents the output from module i (FAM, TES, PPM, or HDPS), W_i is the associated weight coefficient, and $C_i(d)$ is a context adjustment factor that considers decision type, organizational culture, and regulatory considerations.

The algorithm is temporal by design to identify trends and patterns that can signal systematic ethical issues and require organizational attention. Temporal assessment examines rolling averages of ethical scores over defined time frames, triggers alerts when the average falls below established threshold scores, or demonstrates worrisome downward trends.

The appropriateness of establishing weight coefficients for various ethical criteria is an essential factor when implementing EAIM and requires the right balance between organizational values, regulatory requirements, and stakeholder expectations. The framework supports an adaptive weighting approach, changing weightings in response to organizational feedback, regulatory requirements, and performance outcomes.

The initial weight determination uses a standard operating procedure that includes engaging stakeholders, addressing regulatory requirements, and considering the culture of the organization. Initial weights are based on established ethical frameworks and regulatory instruments. The weight values are then adjusted based on organization-specific factors such as industry context, culture, and risk tolerance levels.

The weight optimization algorithm uses gradient-based processes that continuously optimize the weight coefficients based on feedback from multiple sources:

$$w_i(t+1) = w_i(t) + \eta \times \nabla F(w_i(t)) \tag{5}$$

where η represents the learning rate (typically 0.01–0.05) and F represents a composite feedback function incorporating user satisfaction scores, regulatory compliance measures, and business outcome metrics.

The EAIM system implementation leverages a microservices architecture built on Python 3.9+ with FastAPI for service orchestration and Redis for real-time data caching. The choice of Python reflects its extensive machine learning ecosystem and widespread adoption in organizational AI environments, while FastAPI provides the performance characteristics necessary for real-time ethical assessment in high-volume decision environments. This architectural choice aligns with recent work on auditable AI governance and ethical system design, where modular infrastructures are valued because they support traceability, monitoring, and lifecycle-level accountability in organizational decision environments [23, 25].

Core libraries and dependencies:

1) Scikit-learn 1.3.0: Primary machine learning library for bias detection algorithms and fairness metrics computation
2) NumPy 1.24.0 and Pandas 2.0.3: Mathematical operations and data manipulation frameworks
3) PyTorch 2.0.1: Deep learning components for advanced pattern recognition in bias detection
4) Matplotlib 3.7.1 and Seaborn 0.12.2: Visualization libraries for dashboard and reporting components
5) SQLAlchemy 2.0: Database abstraction layer supporting PostgreSQL and MySQL backends
6) Celery 5.3.1: Distributed task queue for background processing of computationally intensive assessments

The system divides the work into three parts: one part for what the user sees and the entry point Application Programming Interface (API), the second part for the logic of the ethics assessment, and the third for storing data on assessments and settings. This division gives each part a chance to grow separately, depending on what the company needs and how the system loads.

In the system's work, we mixed two types of databases—a regular one for clear data and NoSQL for flexible settings and audit trails. PostgreSQL became the main database for recording decisions, assessments, and reviews, and MongoDB stores settings, logs, and reviews in a flexible form, where the schema can be easily changed.

Data on decisions is stored in a clear schema, which provides a connection between them and enables easy search for analyzing trends and reports. The main table contains the essence of the decisions, and the others contain ethics assessments, people's reviews, and the settings that were used during the assessment.

The system can monitor both its own work and trends in ethics assessments, which gives a chance to find problems before they affect how the company makes choices. Performance monitoring sees response times, accuracy of estimates, and workload, while ethics monitoring catches strange trends in estimates or decision outcomes. The system sends different types of alerts: when it is not working well, when ethics assessments are hitting thresholds, when decisions are going wrong, or when bias is visible. Alerts can come via email, SMS, or through whatever the company already has for monitoring systems, through common protocols and APIs.

4. Results and Analysis

Evaluating EAIM on four different company datasets clearly shows that it is better at assessing AI ethics than older methods, although the complexity of implementation varies greatly depending on where and for what it is used. Our tests found that EAIM consistently outperforms and solves problems that older methods struggle to solve.

When EAIM was tested on TechCorp hiring data (15,427 people), it turned out to be very good at catching subtle biases across different groups of people that traditional tools often miss. EAIM scored 89.4% for ethics, while IBM Fairness 360 only scored 78.3%, which is definitely not a fluke ($p < 0.001$, Cohen's $d = 1.23$). It can look at fairness from multiple angles and find hidden problems for people with multiple protected traits—things that older tools often miss or consider normal. Further analysis indicates that EAIM is effective because it evaluates not only outcome parity across groups but also the procedural fairness of the decision-making process. Old methods mostly make sure everyone gets an equal share, but EAIM also checks whether the process is the same for everyone and whether it takes into account cultural differences. This way, it was possible to find 347 cases where everything seemed fair in terms of numbers, but the process itself violated principles of fairness, which really affects people's trust and the company's reputation. EAIM outperformed the comparative tools in terms of user-rated explanation quality, achieving an average score of 7.8 out of 10, compared to 5.2 for the Google What-If Tool. It seems that the point is that EAIM tailors its explanations to people's knowledge, rather than giving everyone the same thing, no matter whether you are a techie or not. The comparative performance of the EAIM framework and existing tools across several evaluation metrics is presented in Table 2.

When we looked at promotion decisions in 23 countries, we found that the concept of fairness and laws are different everywhere. EAIM achieved a score of 92.1%, which is much better than other methods, which failed to take into account cultural differences and the fact that evaluation criteria depend on the context.

For this data, the part of EAIM that can adapt to cultures turned out to be very important. It itself changed the fairness thresholds, taking into account local laws and the company's cultural environment. This made it possible to evaluate decisions equally clearly all over the world but, at the same time, respect the real cultural differences in how companies work and what employees expect.

This result is consistent with prior studies showing that ethical AI governance cannot be fully standardized across jurisdictions because local regulatory expectations, organizational cultures, and stakeholder perceptions shape what counts as fair and acceptable decision-making in practice [13–15].

The biggest gap from other methods was where performance evaluation was based on personal opinion. EAIM found 156 cases where automatic evaluation could hurt people's feelings, although other methods considered them normal. In these cases, everything seemed fair according to the numbers, but the approach itself was too dry or impersonal, which can have a strong impact on the mood of employees and the culture in the company. Because the data was large (21,356 decisions), we were able to get a good understanding of how EAIM performed under real-world load. It handled a variety of decision types reliably and was less resource-intensive than older methods.

Each decision took an average of 0.34 s, which was 67% faster than other methods (1.03 s). This is critical for companies that need to evaluate the ethics of decisions on the fly, when there may be thousands of decisions in an hour.

The scheduling algorithms were also something EAIM could discover that was causing harm to part-timers and employees in some areas that the algorithm could not see. Such biases were months old, yet they were not noticed by older approaches, whereas EAIM detected them through a study of long-term patterns.

Healthcare statistics posed their own difficulties: patient safety, licenses of physicians, and work–life balance. Older tools struggled to address these. However, EAIM, with its medicine-specific settings, achieved a performance level of 91.7% and continued to satisfy essential patient care needs.

The welfare of staff and safety of patients could be balanced, and EAIM identified 23 instances when scheduling algorithms produced dangerous working conditions when they appeared to be only making things more efficient. The legacy techniques failed to do so due to the priority of maintaining equality in all groups and neglecting the concerns of safety and competence.

In terms of all the key outcome measures, it is evident that EAIM is better in terms of its statistics. We conducted paired t-tests with multiple-comparison Bonferroni correction (= 0.01), and all differences in performance exceeded thresholds to reject the hypothesis that the methods are the same. The statistical significance of the performance differences between EAIM and the alternative frameworks is summarized in Table 3.

Effect sizes indicate large practical significance for all comparisons, with Cohen's d values ranging from 0.84 to 1.23, well above the threshold for large effects ($d > 0.8$). These results suggest that performance differences translate into meaningful organizational benefits rather than merely statistical artifacts.

Table 2
TechCorp hiring dataset performance metrics

Metric	EAIM	IBM Fairness 360	Google What-If	Microsoft Fairlearn	Amazon Clarify
Ethical compliance score	89.4%	78.3%	81.7%	84.2%	86.1%
Bias detection accuracy	94.7%	87.2%	89.1%	91.3%	88.6%
False positive rate	4.2%	12.8%	15.7%	9.4%	11.2%
Transparency score	7.8/10	6.1/10	5.2/10	6.7/10	5.9/10
Processing time (seconds)	0.34	1.12	0.89	0.67	0.91
Resource usage (CPU%)	23.1%	45.7%	38.2%	29.4%	41.8%

Table 3
Statistical significance testing results

Comparison pair	t-statistic	p-value	Cohen's d	95% CI lower (%)	95% CI upper (%)
EAIM vs IBM Fairness 360	8.47	< 0.001	1.23	8.9	13.4
EAIM vs Google What-If	7.92	< 0.001	1.15	6.1	9.8
EAIM vs Microsoft Fairlearn	6.73	< 0.001	0.97	3.7	7.2
EAIM vs Amazon Clarify	5.84	< 0.001	0.84	2.1	4.9

Detailed analysis reveals EAIM's efficiency advantages stem from architectural optimizations and intelligent caching strategies rather than reduced analytical sophistication. The framework's modular design enables selective activation of assessment components based on decision criticality, allowing organizations to balance thoroughness with computational efficiency according to their specific requirements. A comparison of computational resource utilization among the evaluated frameworks is shown in Table 4.

Tests under load up to 10,000 decisions per hour showed that EAIM grows linearly and maintains the same speed and accuracy even with a large flow. But the old methods start to fail after 3000–4000 decisions per hour, especially the search for distortions suffers—where the number of false alarms increases significantly when the system is under load.

When we systematically evaluated EAIM parameters, we found notable patterns in which values are better for different types of companies and decisions. We used the grid method to search for ideal coefficients so that the accuracy of the ethics assessment was higher, and the calculations and acceptance by people did not suffer.

Analysis of the weight of the fairness module (α): Values from 0.25 to 0.35 give the best results on all datasets, and $\alpha = 0.30$ was the most reliable choice for common tasks. But for companies that are more afraid of distortions or under strict supervision, it is better to take a little higher ($\alpha = 0.33$–0.35), and for those who think more about fast work, average values ($\alpha = 0.27$–0.30) are suitable.

Transparency weight (β): Here we were in for a surprise—the module worked stably with different weights from 0.20 to 0.30. Such strength suggests that for better transparency, the algorithm itself is more important than its weight in the overall system. It seems that the subsystem that creates explanations is made very well.

Influence of the cultural adaptation factor (δ): The analysis revealed that the best cultural adaptation settings are highly dependent on the region and culture of the company. Medical companies need higher sensitivity to culture ($\delta = 0.25$–0.30), and tech companies work better with average adaptation ($\delta = 0.15$–0.20). Financial services have two modes: old banks like high adaptation, and fintech companies like low. As we removed parts of EAIM one by one and looked at the results, we realized how each of them affected the overall performance of the system. This gave us a hint about which parts were most important for assessing the ethics of AI in different settings. The impact of different EAIM module configurations on ethical performance metrics is presented in Table 5.

The FAM is the most critical element, and its removal results in a significant performance decrease, estimated at 22.2% for the ethical rating. This is understandable, as fairness plays a key role in the moral aspects of AI applications in organizations.

Regardless, the magnitude of the decrease thus indicates that the multidimensional EAIM fairness acquisition strategy yielded greater benefits than simple demographic parity checks in single-dimensional approaches. Similarly, the transparency assessment system has a moderate impact on overall ethical ratings but undermines user understanding. This suggests that transparency influences ethical assessments through stakeholder belief and acceptance, rather than through explicit fairness. This underscores the relevance of holistic AI ethical practices that consider both the human aspects of the technology and its impact on the human being.

An analysis of the interactions between the components reveals that the effects arising between them create synergistic results that exceed the simple summation of their contributions. The overall ethics score for FAM and TES is 8.5%, significantly higher than the sum of the individual components' contributions, which is 0.5% (0.5 + 5.3).

This interaction is beneficial in that the transparency and fairness assessment capabilities mutually reinforce each other by increasing stakeholder trust and more accurately identifying bias. Through error analysis, failures can be categorized into technical constraints, data quality, and issues concerned with contextual interpretation.

Technical limitation failures (12.3% of errors): In some cases, with more complex models of machine learning, EAIM cannot

Table 4
Resource utilization comparison

Framework	CPU Usage (%)	Memory (GB)	Storage (MB)	Network (KB/s)
EAIM	23.1 ± 4.2	1.8 ± 0.3	12.4 ± 2.1	8.7 ± 1.9
IBM Fairness 360	45.7 ± 8.9	3.2 ± 0.7	28.6 ± 5.3	15.2 ± 3.4
Google What-If	38.2 ± 7.1	2.7 ± 0.5	19.8 ± 3.9	12.1 ± 2.8
Microsoft Fairlearn	29.4 ± 5.6	2.1 ± 0.4	16.2 ± 3.1	9.8 ± 2.2
Amazon Clarify	41.8 ± 7.8	2.9 ± 0.6	22.3 ± 4.2	13.6 ± 3.1

Table 5
Impact of EAIM module configurations on ethical performance metrics

Configuration	Ethical score	Bias detection	Transparency	Processing time	Resource usage
Complete EAIM	89.4%	94.7%	7.8/10	0.34s	23.1%
Without FAM	67.2% (–22.2%)	78.3% (–16.4%)	7.9/10 (+0.1)	0.28s (–17.6%)	18.7% (–19.0%)
Without TES	84.1% (–5.3%)	93.2% (–1.5%)	4.2/10 (–46.2%)	0.31s (–8.8%)	20.8% (–9.9%)
Without PPM	86.7% (–2.7%)	94.1% (–0.6%)	7.6/10 (–2.6%)	0.33s (–2.9%)	21.9% (–5.2%)
Without HDPS	85.8% (–3.6%)	93.8% (–0.9%)	7.4/10 (–5.1%)	0.32s (–5.9%)	22.3% (–3.5%)
FAM + TES only	78.9% (–10.5%)	89.4% (–5.3%)	6.1/10 (–21.8%)	0.19s (–44.1%)	14.2% (–38.6%)
Minimal (FAM only)	71.3% (–18.1%)	85.7% (–9.0%)	3.8/10 (–51.3%)	0.16s (–52.9%)	11.8% (–48.9%)

explain them, particularly when it is trained in an ensemble learning model, or when features have large numbers of dimensions in deep neural networks. Such cases usually feature advanced predictive models in which meaningful transparency would consume computation resources that are out of a reasonable deployment range.

Such transparency constraints are well documented in the ethics and auditing literature on algorithmic systems, particularly where model opacity, post hoc explanation limits, and organizational accountability requirements create tensions between technical performance and meaningful interpretability [23–25]. These limitations reflect fundamental explainability trade-offs rather than deficiencies of the EAIM framework itself.

Example case: A financial services promotion algorithm based on gradient boosting on 847 features and complex interaction terms would not provide an intuitive explanation to affected employees, although it was highly accurate in identifying bias and determining procedural fairness.

Data quality problems (31.7% of errors): Partial or discrepant historical data make it difficult to set fairness expectations in the baseline, and it also makes it difficult to set cultural adaptation parameters. Companies that have ineffective data governance policies have a rate of errors that is especially elevated when a cross-cultural environment is involved, where normal variation assumptions are not valid.

Example case: Scheduling system of a retail chain seemed to favor particular demographics of employees regionally; however, further analysis revealed that the apparent favoritism was due to a lack of full capturing of employee availability choices and regional labor market variations and not algorithmic discrimination.

Contextual interpretation challenges (56.0% of errors): The most common type of error is one where the organizational context or even the expectations of the stakeholders are misinterpreted, and therefore, their assessments are accurate but inappropriate in a specific business environment or culture.

Example case: EAIM was presented with a low score of dignity preservation in a healthcare staffing algorithm that was assigning overnight shifts based on seniority and specialization, which the EAIM interpreted as potentially unfair to younger employees. However, the research conducted within the organization demonstrated that the top staff was valuing these changes due to the differentiated pay and career growth, and thus, the transfers were consistent with the workers' occupational and professional development concerns.

Error mitigation strategies: When the experience of effective error recovery is examined, it is possible to identify several trends that may be adopted by an organization in order to make EAIM as effective as possible. Domain-specific feedback calibration has been shown to be most effective, with error rates 23.4 times lower on average after three months of calibration. In addition, the local stakeholder input in the initial configuration reduces the contextual interpretation errors by approximately 31%.

The 6-month performance follow-up on implementation in pilot organizations shows that EAIM performance evolves over time in intriguing ways. Unlike the conventional AI systems that may degenerate unless re-trained, the adaptive learning capabilities of EAIM offer a chance to produce improved output with regard to the experience attained and changes in the stakeholder responses.

Month-over-month accuracy gains of 0.8–1.2% in the first four months and plateau, reflecting that organizational adaptation is a critical component of effective implementation of ethical AI. The organizations making systematic feedback and parameter refinement collection will perform 15–20% better in the long term as compared to those adopting static collections.

The analysis demonstrates that ethical AI evaluation should be treated as a continuous organizational capability requiring ongoing monitoring, adaptation, and improvement rather than as a static tool with fixed performance characteristics.

5. Discussion and Implications

A massive study has demonstrated that the new ethical AI in the field of management and HR demands clever strategies that can strike a balance among new technologies, rapid work, legislation, and trust of the people. Although good frameworks can be used to understand ethical rules in theory, there is often a great deal that these requirements need, in terms of practice, to be adjusted to meet the needs of the company, industry, and culture of the place, which general frameworks do not necessarily offer.

Everyone is discussing transparency, meaning that everyone is in agreement that AI in companies should provide some explanation on how it works. However, big technical challenges lie on the way to this goal, in particular, with sophisticated machine learning algorithms that operate by some elaborate mathematical tricks that are hard to understand in simple terms. Businesses are often faced with a decision of the effectiveness of a system versus its complexity to comprehend, and the various ethical theories address this issue differently, which does not always work in favor of the work. The issue of algorithm bias shows that ethical AI does not happen once but rather is a journey of time, unlike traditional regulations, which can just write a policy, train a system, and tweak it. In comparison to the traditional regulations, where the end point and criteria of the readiness of everything are clear, ethical AI presupposes testing the system and its constant improvement, as well as additional adjustments of the system over time. It means that the companies must plan resources, strategies, and how they will fund these initiatives in the years.

The identified transparency limitations in complex models should not be interpreted solely as deficiencies of the EAIM framework, but rather as reflections of broader structural constraints inherent to contemporary algorithmic systems, including opacity, explanation limits, and the difficulty of assigning responsibility in automated decision environments [23–25]. Recent research suggests that lifecycle auditing procedures, modular oversight mechanisms, and design-stage ethical controls may partially mitigate these constraints by embedding transparency and accountability into system governance rather than treating them as purely technical add-ons [23, 25]. However, such approaches remain bounded by trade-offs between explanation depth, computational cost, and stakeholder usability.

The excessive disparities in the laws between the countries pose a challenge to international companies that have to comply with the complex requirements but still have to be capable of working together in a rational and logical manner. One of the differences in laws is reflected in different values, legal traditions, and priorities but too much variation can hinder the development of responsible AI since it can make compliance costly and the task more complicated. These issues can be addressed in the long term through the formulation of international standards and best practices, but much effort will be required to ensure that they can actually be of use to all. Companies interested in employing ethical AI should have elaborate strategies that encompass many aspects including technical expertise, organizational culture, laws, and interaction with stakeholders. The figures reveal that the success should be accompanied by strong executive buy-in, finances to survive more than several years, and permanent consideration of the new challenges together with the appearance of technology and the change of laws.

The research established that phase implementation is the most appropriate approach as firms get to experience, develop processes, and gradually build trust and have control over the risks and costs. Beginning with low-priority origin locations by pilots can assist pilots in learning in smaller groups prior to addressing larger issues with greater complexity and risk.

Ethical training of employees is a crucial part of ethical AI success—businesses must develop the ability to comprehend and address AI ethics on all levels and in all departments. This is especially important because ethical AI implementation depends not only on formal policies but also on managerial interpretation, organizational learning, and employees' ability to understand how automated decisions affect professional life and workplace dignity [16, 17, 21]. It is not only technical training but also the set of ethics, communication, and change management skills that make it possible to drive change successfully.

6. Limitations and Future Research Directions

There are gaps in our review that provide opportunities for new research in the area of AI ethical frameworks. Our analysis mostly took English-language texts from Western academia and business, which may have limited the range of ideas and approaches. Future work should more thoroughly examine AI ethical frameworks from different cultures, economies, and legal systems to better understand how the world is approaching the implementation of ethical AI. Comparative cross-sector analysis is also needed because evidence from education, finance, and healthcare indicates that ethical priorities and acceptable governance mechanisms differ markedly depending on domain-specific risks and institutional constraints [18–20].

AI and its laws are changing so quickly that our findings may soon become outdated as new features emerge and laws mature. Future work should create ways to continuously monitor how AI ethical frameworks work and how they change over time.

Our analysis relied heavily on theories and texts about how things should be but took little data on how things actually work in companies. Research is needed that examines the experience of real-world implementation, measures both successes and failures, and provides advice on what really works.

7. Conclusion

Our full analysis of ethical frameworks in AI for management and HR shows that the field is growing rapidly and there are advances in theory, but it remains difficult to implement across companies. The theoretical foundation for ethical AI is now firmly rooted in principles such as clarity, fairness, accountability, and respect for people, providing important guidance for decisions about how to implement and govern AI. However, translating these principles into practice remains complex and resource-intensive because ethical AI requires governance structures, monitoring procedures, managerial accountability, and continuous adaptation rather than onetime compliance measures. Companies must address technical challenges, comply with regulations, and change their structure without losing edge or reducing operational efficiency. The most important success factors of ethical AI implementation include the commitment of the leaders to the success, sufficient funding over the long term, and the continuous consideration of new issues as technology advances and society demands more of it.

These effective workflows recognize that ethical AI implementation is a journey that should be an ongoing and constantly improving process and not a place where the sense of the meaning of what is done is clear. As AI technologies continue to advance and regulatory frameworks evolve, companies will need to remain adaptable, uphold ethical principles, and continuously optimize their business processes. The involvement of technologists, ethicists, regulators, and business leaders is mandatory to design the framework that will help reach a balance between innovation and responsibility, efficiency and fairness, and technical skills and human values to ensure ethical AI in business.

Recommendations

Our full analysis offers some clear advice for companies implementing ethical AI frameworks in the areas of management and HR:

Companies should implement AI in stages, allowing them to learn and adapt over time. Start with pilot projects in less critical areas, and then move on to more critical tasks where the stakes and requirements are higher.

Investing in full workforce training is essential to ensure that the company has the ability to understand and manage the ethical side of AI systems at different levels and across departments.

Systems should be in place for regular audits and controls to ensure ethical principles are being followed, including both automated controls and human oversight that can adapt to changing conditions and requirements.

Companies should be actively involved in the development of industry laws and standards to influence how ethical AI governance evolves and to be prepared for new requirements that arise.

We need to work with everyone affected to maintain transparency, build trust, and gather feedback from employees, customers, and others about how AI works and the ethical implications it has.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Kirill Toropov: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization. **Irina Manakhova:** Conceptualization, Validation, Resources, Writing – review & editing, Supervision, Project administration.

References

[1] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1–21. https://doi.org/10.1177/2053951716679679

[2] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., . . . , & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

[3] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

[4] Osasona, F., Amoo, O. O., Atadoga, A., Abrahams, T. O., Farayola, O. A., & Ayinla, B. S. (2024). Reviewing the ethical implications of AI in decision making processes. *International Journal of Management & Entrepreneurship Research*, *6*(2), 322–335. https://doi.org/10.51594/ijmer.v6i2.773

[5] Chowdhury, T., & Oredo, J. (2023). AI ethical biases: Normative and information systems development conceptual framework. *Journal of Decision Systems*, *32*(3), 617–633. https://doi.org/10.1080/12460125.2022.2062849

[6] Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2023). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. *Human Resource Management Review*, *33*(1), 100925. https://doi.org/10.1016/j.hrmr.2022.100925

[7] Bankins, S. (2021). The ethical use of artificial intelligence in human resource management: A decision-making framework. *Ethics and Information Technology*, *23*(4), 841–854. https://doi.org/10.1007/s10676-021-09619-6

[8] Bankins, S., Formosa, P., Griep, Y., & Richards, D. (2022). AI decision making with dignity? Contrasting workers' justice perceptions of human and AI decision making in a human resource management context. *Information Systems Frontiers*, *24*(3), 857–875. https://doi.org/10.1007/s10796-021-10223-8

[9] Ferrell, O. C., Harrison, D. E., Ferrell, L. K., Ajjan, H., & Hochstein, B. W. (2024). A theoretical framework to guide AI ethical decision making. *AMS Review*, *14*(1-2), 53–67. https://doi.org/10.1007/s13162-024-00275-9

[10] Manda, V. K., Christy, V., & Jitta, M. R. (2024). Ethical AI and decision-making in management leadership. In P. Bhattacharya, A. Hassan, H. Liu, & B. Bhushan (Eds.), *Ethical dimensions of AI development* (pp. 197–226). IGI Global. https://doi.org/10.4018/979-8-3693-4147-6.ch009

[11] Farayola, O. A., & Olorunfemi, O. L. (2024). Ethical decision-making in IT governance: A review of models and frameworks. *International Journal of Science and Research Archive*, *11*(2), 130–138. https://doi.org/10.30574/ijsra.2024.11.2.0373

[12] Brendel, A. B., Brennecke, J. T., Zapadka, P., & Kolbe, L. M. (2021). Ethical management of artificial intelligence. *Sustainability*, *13*(4), 1974. https://doi.org/10.3390/su13041974

[13] Kaur, J. (2024). Responsible artificial intelligence (AI) governance: Ethical frameworks for business decision-making. In D. A. M. Roache (Ed.), *Ethical quandaries in business practices: Exploring morality and social responsibility* (pp. 337–368). IGI Global. https://doi.org/10.4018/979-8-3693-3948-0.ch014

[14] Attard-Frost, B., Ríos, de los., A, & Walters, D. R. (2023). The ethics of AI business practices: A review of 47 AI ethics guidelines. *AI and Ethics*, *3*(2), 389–406. https://doi.org/10.1007/s43681-022-00156-6

[15] Baker-Brunnbauer, J. (2021). Management perspective of ethics in artificial intelligence. *AI and Ethics*, *1*(2), 173–181. https://doi.org/10.1007/s43681-020-00022-3

[16] Rhem, A. J. (2021). AI ethics and its impact on knowledge management. *AI and Ethics*, *1*(1), 33–37. https://doi.org/10.1007/s43681-020-00015-2

[17] Li, J. L., Litvinova, Y., Marabelli, M., & Newell, S. (2024). Ethical implications of AI use in practice for decision-making. In I. Constantiou, M. P. Joshi, & M. Stelmaszak (Eds.), *Research handbook on artificial intelligence and decision making in organizations* (pp. 359–375). Edward Elgar Publishing. https://doi.org/10.4337/9781803926216.00030

[18] Khreisat, M. N., Khilani, D., Rusho, M. A., Karkkulainen, E. A., Tabuena, A. C., & Uberas, A. D. (2024). Ethical implications of AI integration in educational decision making: Systematic review. *Educational Administration: Theory and Practice*, *30*(5), 8521–8527. https://doi.org/10.53555/kuey.v30i5.4406

[19] Owolabi, O. S., Uche, P. C., Adeniken, N. T., Ihejirika, C., Islam, R. B., & Chhetri, B. J. T. (2024). Ethical implication of artificial intelligence (AI) adoption in financial decision making. *Computer and Information Science*, *17*(1), 49–56. https://doi.org/10.5539/cis.v17n1p49

[20] Elendu, C., Amaechi, D. C., Elendu, T. C., Jingwa, K. A., Okoye, O. K., John Okah, M., . . . , & Alimi, H. A. (2023). Ethical implications of AI and robotics in healthcare: A review. *Medicine*, *102*(50), e36671. https://doi.org/10.1097/MD.0000000000036671

[21] Usmani, U. A., Happonen, A., & Watada, J. (2023). Human-centered artificial intelligence: Designing for user empowerment and ethical considerations. In *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications*, 1–7. https://doi.org/10.1109/HORA58378.2023.10156761

[22] Wang, C., Liu, S., Yang, H., Guo, J., Wu, Y., & Liu, J. (2023). Ethical considerations of using ChatGPT in health care. *Journal of Medical Internet Research*, *25*, e48009. https://doi.org/10.2196/48009

[23] Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., . . . , & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33–44. https://doi.org/10.1145/3351095.3372873

[24] Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835–850. https://doi.org/10.1007/s10551-018-3921-3

[25] Biondi, G., Cagnoni, S., Capobianco, R., Franzoni, V., Lisi, F. A., Milani, A., & Vallverdú, J. (2023). Editorial: Ethical design of artificial intelligence-based systems for decision making. *Frontiers in Artificial Intelligence*, *6*, 1250209. https://doi.org/10.3389/frai.2023.1250209

How to Cite: Toropov, K., & Manakhova, I. (2026). Ethical Frameworks in AI Decision-Making: Comprehensive Analysis of Management and Human Resource Applications in Contemporary Organizations. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA62026783