DIGITAL TRANSFORMATION AND ADMINISTRATION INNOVATION

Citation: Kavoosi Davoodi, S. M. & Baghbani Gatab, A. (2026). Risk Analysis of Artificial Intelligence Deployment in Administrative Processes Using the Integrated FMEA-DEMATEL Approach: A Case Study of Bank Sepah, Mazandaran Province. *Digital Transformation and Administration Innovation*, 4(5), 1-13.

Received date: 2026-02-10 Revised date: 2026-06-04 Accepted date: 2026-06-11 Initial published date: 2026-06-13 Final published date: 2026-09-01

Risk Analysis of Artificial Intelligence Deployment in Administrative Processes Using the Integrated FMEA-DEMATEL Approach: A Case Study of Bank Sepah, Mazandaran Province

Seyed Mojtaba Kavoosi Davoodi[1], Ali Baghbani Gatab[2*]

1. Assistant Professor, Department of Industrial Engineering, Shomal University, Amol, Iran
2. Master's Degree, Department of Business Administration, Bab.C., Islamic Azad University, Babol, Iran

*****Correspondence: e-mail: ali.baghban49452589@gmail.com

Abstract

Today, the deployment of artificial intelligence (AI) in organizational administrative processes, particularly within the banking industry, has emerged as one of the most significant manifestations of digital transformation. Despite its advantages, including increased speed, accuracy, productivity, and enhanced decision-making, the implementation of AI in administrative processes is associated with numerous risks across data, technical, security, organizational, and governance dimensions. Therefore, the present study aimed to analyze and prioritize the risks associated with AI deployment in administrative processes using the integrated FMEA-DEMATEL approach in Bank Sepah of Mazandaran Province. In terms of purpose, this study is applied research, and in terms of methodology, it adopts a descriptive-analytical approach. Research data were collected through a literature review and expert questionnaires. The study population consisted of experts, managers, and specialists familiar with information technology, risk management, and banking administrative processes, who were selected through purposive sampling. In the first stage, 20 major risks related to AI deployment were identified and evaluated using the Failure Mode and Effects Analysis (FMEA) method based on three criteria: severity, occurrence probability, and detectability. The identified risks were subsequently ranked according to their Risk Priority Numbers (RPNs). The results of this stage indicated that the risks of "information leakage and confidentiality breaches," "lack of model transparency and interpretability," and "model errors and inaccurate predictions" received the highest priority rankings. In the second stage, the top ten risks were analyzed using the Decision-Making Trial and Evaluation Laboratory (DEMATEL) method to examine their causal relationships and mutual influences. The findings revealed that "absence of an AI governance framework," "poor data quality," and "bias in training data" were among the most significant causal and root risks. In contrast, risks such as "information leakage and confidentiality breaches" and "cyber and adversarial attacks on AI models" were primarily consequence-oriented and effect-related risks. The results indicate that effective management of AI-related risks in banks requires simultaneous attention to both critical risks and their underlying root causes at the governance, data, and model design levels. The findings of this study can serve as a foundation for designing AI governance mechanisms, strengthening information security, and improving decision-making processes within banking organizations.

Keywords: Artificial Intelligence, Administrative Processes, Risk Management, FMEA, DEMATEL, Digital Transformation.

1. Introduction

Artificial intelligence has become one of the central components of digital transformation in contemporary organizations, reshaping the way administrative, operational, and managerial processes are designed, executed, monitored, and improved. In recent years, AI-based systems have moved beyond experimental applications and have increasingly been embedded in decision-making workflows, customer service systems, document processing, fraud detection, predictive analytics, risk monitoring, process automation, and managerial control mechanisms. This transition has been particularly important in service-intensive organizations such as banks, where administrative processes depend heavily on data accuracy, procedural reliability, regulatory compliance, information security, and timely decision-making. The expansion of AI in organizational settings is therefore not merely a technological change, but a socio-technical transformation that affects work structures, professional roles, governance mechanisms, accountability systems, and stakeholder trust (Archer, 2025; Cati, 2024; Patil et al., 2024).

In the banking sector, administrative processes involve a complex combination of customer data management, internal documentation, compliance monitoring, credit assessment, operational reporting, risk control, and interdepartmental coordination. These processes are often repetitive, data-intensive, rule-based, and sensitive to errors; therefore, they provide a suitable environment for the deployment of AI and automation technologies. AI can improve administrative performance by increasing processing speed, reducing human error, enabling predictive decision support, identifying abnormal patterns, and enhancing the efficiency of internal workflows. Similar developments have been observed across other data-intensive sectors, where AI and digital technologies have been applied to improve operational efficiency, predictive maintenance, supply chain resilience, and service quality (Brown et al., 2024; E. & B., 2024; Elete et al., 2024; Ugwu & Balogun, 2024). In this regard, AI deployment in banking administration can be understood as part of a broader movement toward intelligent, data-driven, and digitally integrated organizational systems.

Despite these potential benefits, the implementation of AI in administrative processes creates a wide range of risks that cannot be reduced to technical malfunction alone. AI systems operate through data pipelines, algorithmic models, infrastructure platforms, human oversight mechanisms, and organizational rules. Weakness in any of these components can generate adverse consequences for the entire system. Studies on AI implementation have increasingly emphasized that organizations must move from a narrow technology-adoption perspective toward an implementation-governance perspective, in which algorithmic performance, data quality, accountability, monitoring, human capability, and institutional readiness are considered simultaneously (Boag et al., 2024; Rajagopal et al., 2024; Wells et al., 2025). Therefore, AI risk analysis should not only identify visible operational threats but also uncover the underlying causal structure through which risks influence one another.

One of the most fundamental categories of AI-related risk concerns data. AI systems depend on the availability, quality, completeness, consistency, and representativeness of data. In administrative banking processes, inaccurate, incomplete, fragmented, outdated, or biased data may directly reduce the reliability of AI outputs and indirectly intensify other risks, including model error, discriminatory decision-making, compliance failure, reputational damage, and loss of trust. Data-related challenges are particularly significant in organizations with legacy systems, fragmented databases, and heterogeneous information infrastructures. Industrial data platform research has shown that data architecture, interoperability, standardization, and governance are essential conditions for successful digital transformation (Ghosh & Laad, 2024). Similarly, studies on digital ecosystems and smart organizations highlight that the creation of reliable AI-enabled services depends on integrated data environments, coordinated digital infrastructure, and systematic data stewardship (Chow et al., 2025). Hence, in banking administration, poor data quality and data inconsistency should be regarded not as secondary technical issues but as root-level risks that may shape the behavior of the entire AI system.

A second major risk dimension concerns the model and technology layer. AI models may generate inaccurate predictions, fail to generalize across branches or administrative units, become unstable when data patterns change, or operate as opaque "black-box" systems. Lack of transparency and interpretability is particularly problematic in banking because administrative decisions often require justification, documentation, auditability, and accountability. When a model produces an output that cannot be explained, employees, managers, auditors, regulators, and customers may find it difficult to trust or contest the decision. This issue has also been identified in research on generative AI and clinical practice, where safe implementation requires attention to validity, explainability, workflow integration, human oversight, and continuous review (Halamka et al.,

2024; Scott et al., 2025). In software and large-scale AI deployment, similar concerns arise around efficiency, reliability, monitoring, and the practical consequences of embedding AI into real organizational workflows (Mao et al., 2025). Therefore, the technical performance of AI systems cannot be separated from organizational accountability and managerial control.

Security and privacy risks represent another critical domain in AI deployment. Banks process highly sensitive customer, employee, financial, and operational data; consequently, information leakage, confidentiality breaches, unauthorized access, insecure storage, adversarial manipulation, and cyberattacks can create severe legal, financial, operational, and reputational consequences. The increasing complexity of AI-enabled systems can expand the attack surface, particularly when models depend on cloud infrastructures, external platforms, application programming interfaces, or third-party technology providers. Research on AI-powered security frameworks emphasizes the need for zero-trust architectures, continuous authentication, access control, encryption, and monitoring to reduce vulnerabilities in complex digital environments (Dash, 2024). Studies on network security and AI-based intrusion detection similarly show that intelligent systems require robust protection against cyber threats, adversarial attacks, and infrastructure-level vulnerabilities (Cunha et al., 2024; Ibokette et al., 2024). For banking organizations, these insights indicate that AI risk management must be closely integrated with information security governance and privacy protection mechanisms.

In addition to technical and security risks, AI deployment also introduces organizational and human resource challenges. Employees may resist AI-based systems because of fear of job displacement, lack of trust in algorithmic decisions, uncertainty about new responsibilities, or insufficient digital skills. If employees and managers do not understand how AI systems operate, how outputs should be interpreted, and how errors should be escalated, the organization may experience implementation failure even when the technical system is functional. Research on workplace relations shows that AI can create relational and psychosocial risks, including tension in workplace roles, changes in authority relations, and new forms of stress or uncertainty (Cebulla, 2025). From a socio-technical perspective, AI-augmented environments may also generate cognitive pressure, technostress, and new risk-management needs, especially when employees are expected to interact with complex automated systems without adequate training and support (Kereopa-Yorke, 2023). Accordingly, the success of AI in administrative banking processes depends not only on system design but also on change management, employee capability development, organizational communication, and trust-building.

Governance, compliance, and ethical risks form another central dimension of AI implementation. In regulated sectors such as banking, AI systems must comply with legal requirements, privacy regulations, information security standards, audit expectations, and ethical principles. The absence of a clear AI governance framework can lead to ambiguity regarding responsibility, accountability, oversight, model validation, data ownership, risk reporting, and corrective action. This issue has been widely discussed in healthcare AI governance, where legal concerns, regulatory uncertainty, and the need for structured oversight frameworks have become major barriers to safe adoption (Alanazi, 2025; Arnaout, 2025). Practical AI implementation frameworks similarly emphasize the importance of appropriate review, institutional governance, stakeholder engagement, and post-deployment monitoring (Wells et al., 2025). Although healthcare and banking differ in their operational contexts, both sectors involve sensitive data, high-stakes decisions, professional accountability, and public trust; therefore, the governance lessons from one domain can inform risk analysis in the other.

The growing use of AI has also created new demands for auditability and compliance verification. Organizations increasingly need methods to demonstrate that AI-enabled systems comply with internal policies and external regulations without necessarily exposing sensitive data, proprietary algorithms, or confidential operational details. Emerging work on zero-knowledge software auditing highlights the possibility of verifying compliance in AI-enabled systems while preserving confidentiality and reducing disclosure risks (Scaramuzza, 2025). For banks, this issue is highly relevant because AI governance must balance transparency, explainability, data protection, cybersecurity, and competitive confidentiality. In this respect, AI risk management should not be limited to pre-implementation assessment; rather, it must include continuous auditing, monitoring, documentation, and evidence-based assurance throughout the AI system lifecycle.

The literature also shows that AI risk is highly context-dependent. In fire safety, AI risk management requires attention to policy, stakeholders, emerging technologies, and operational readiness (Kumar et al., 2025). In healthcare, perceived threats to rights and safety include privacy violations, unequal access, algorithmic bias, accountability gaps, and unsafe decision

support (Botha et al., 2024). In industrial and chemical contexts, digital twins and intelligent systems raise issues related to model fidelity, real-time monitoring, system integration, and operational reliability (Mane et al., 2024). In robotic process automation, implementation success depends on aligning automation capacity with process characteristics, organizational readiness, and human supervision (Bichel et al., 2023). These examples indicate that AI-related risks should be analyzed according to the operational environment in which AI is deployed. Therefore, in the case of Bank Sepah in Mazandaran Province, risk analysis must consider the specific administrative processes, banking regulations, data structures, organizational culture, and digital transformation conditions of the institution.

A further challenge is that AI risks do not operate independently. Poor data quality may lead to model error; model opacity may weaken accountability; weak governance may increase compliance failure; insufficient employee training may reduce the effectiveness of monitoring; and inadequate access control may intensify privacy breaches. Thus, AI risks form a network of mutual influence in which some risks are root causes and others are consequences. Traditional prioritization methods can identify critical risks based on severity, likelihood, or detectability, but they may not fully explain how risks interact. This distinction is important because a risk with a high immediate priority may be an effect of deeper structural weaknesses, whereas a moderately ranked risk may function as a causal driver of several other risks. For this reason, integrated risk analysis approaches are needed to combine prioritization with causal mapping.

Failure Mode and Effects Analysis (FMEA) is a widely used technique for identifying and prioritizing risks based on severity, occurrence, and detection. Its strength lies in its ability to convert expert judgments into a structured Risk Priority Number, thereby helping managers recognize which risks require urgent attention. However, FMEA has limitations when risks are interdependent, because it usually evaluates each risk separately and does not explicitly model causal relationships among them. The Decision-Making Trial and Evaluation Laboratory (DEMATEL) method complements this limitation by analyzing the degree of influence among factors and classifying them into cause and effect groups. When combined, FMEA and DEMATEL provide a more comprehensive approach: FMEA identifies high-priority risks, while DEMATEL reveals which of those risks are root drivers and which are dependent outcomes. Such integrated logic is consistent with contemporary AI implementation research, which emphasizes lifecycle thinking, interdependency analysis, governance mechanisms, and systemic risk management (Boag et al., 2024; Rajagopal et al., 2024).

In the context of AI deployment in banking administrative processes, the integrated FMEA-DEMATEL approach is especially useful because it allows managers to avoid a purely reactive strategy. For example, information leakage may appear as the most severe or urgent risk because of its direct consequences for confidentiality and trust. However, causal analysis may reveal that this risk is driven by deeper factors such as weak AI governance, poor data quality, insufficient access control, model opacity, or inadequate employee training. Similarly, model error may be both a high-priority risk and a central causal factor because inaccurate outputs can affect operational decisions, compliance processes, customer service, internal control, and reputational outcomes. Therefore, effective AI risk management requires simultaneous attention to critical outcomes and root causes.

Overall, the deployment of AI in administrative banking processes offers significant opportunities for improving efficiency, accuracy, responsiveness, and decision quality, but it also creates multidimensional risks across data, technology, security, organization, and governance domains. The reviewed literature suggests that successful AI implementation depends on integrated data infrastructure, explainable and reliable models, cybersecurity safeguards, human capability development, regulatory compliance, and continuous governance. However, limited attention has been paid to the simultaneous prioritization and causal analysis of AI deployment risks in banking administrative processes, particularly through an integrated FMEA-DEMATEL approach. Addressing this gap can provide banking managers with a more actionable understanding of which risks are most critical, which are structurally influential, and which require preventive rather than merely corrective strategies.

The aim of this study is to identify, evaluate, prioritize, and analyze the causal structure of risks associated with artificial intelligence deployment in the administrative processes of Bank Sepah in Mazandaran Province using an integrated FMEA-DEMATEL approach.

2. Methods and Materials

This study was conducted as an applied research project with a descriptive–analytical design. The primary objective was to identify, evaluate, prioritize, and analyze the risks associated with the deployment of artificial intelligence (AI) in the administrative processes of Bank Sepah in Mazandaran Province and to provide managerial recommendations for mitigating these risks. To achieve this objective, an integrated Failure Mode and Effects Analysis (FMEA) and Decision-Making Trial and Evaluation Laboratory (DEMATEL) approach was adopted. The integrated methodology enabled not only the prioritization of AI-related risks based on their severity and likelihood but also the examination of the causal relationships among the identified risks. Consequently, the study provided a comprehensive understanding of both the relative importance of individual risks and the interaction structure among them, thereby supporting more effective managerial decision-making.

The target population consisted of experts and professionals with substantial knowledge and experience in information technology, risk management, banking operations, and administrative processes within Bank Sepah in Mazandaran Province. These individuals were selected because of their direct involvement in digital transformation initiatives, information systems implementation, and organizational process management within the banking sector. Given the expert-oriented nature of multi-criteria decision-making and risk analysis studies, purposive (judgmental) sampling was employed to identify participants who possessed sufficient expertise and practical experience relevant to the research topic. Consistent with previous studies utilizing FMEA and DEMATEL methodologies, a panel of experts was considered appropriate because the quality and depth of professional judgment are more critical than sample size in such analyses (Liu et al., 2020). Accordingly, the expert panel included information technology managers, systems analysts, banking operations managers, and risk management specialists. Eligibility criteria required participants to possess at least five years of professional experience in banking or information technology, familiarity with intelligent systems and digital transformation projects, experience in administrative process management or risk management, and willingness to participate in the study.

Data collection was carried out through both documentary and field research methods. In the documentary phase, an extensive review of the literature was conducted to establish the theoretical foundations of the study and identify potential risks associated with AI deployment in administrative environments. Relevant scientific articles, specialized books, and internationally recognized academic databases, including ScienceDirect, Springer, IEEE Xplore, and Google Scholar, were consulted. The literature review facilitated the identification of an initial list of AI deployment risks across technical, data-related, security, organizational, governance, legal, and ethical dimensions.

Subsequently, field data were collected using expert questionnaires specifically designed for the FMEA and DEMATEL analyses. The initial list of identified risks was subjected to expert validation, during which overlapping risks were merged, irrelevant items were removed, and context-specific risks relevant to Bank Sepah and the operational environment of Mazandaran Province were incorporated. The finalized risk framework consisted of twenty risks categorized into five major dimensions: data risks, model and technology risks, security and privacy risks, organizational and human resource risks, and governance, compliance, and ethical risks.

The FMEA questionnaire was designed to evaluate each identified risk according to three standard dimensions: Severity (S), Occurrence (O), and Detection (D). Experts assessed each risk using a ten-point Likert-type scale ranging from 1 to 10 for each criterion. Severity reflected the potential impact of a risk on banking operations, occurrence represented the likelihood of the risk materializing, and detection measured the ability of the organization to identify the risk before significant consequences occurred.

The DEMATEL questionnaire was developed to investigate the interrelationships among the most critical risks identified through the FMEA stage. Participants evaluated the degree of influence that each risk exerted on other risks using a five-point scale ranging from “no influence” to “very high influence.” These evaluations were subsequently used to construct the direct-relation matrix required for DEMATEL analysis.

To ensure content validity, the preliminary questionnaires were reviewed by university faculty members and banking and information technology experts. Based on their feedback, revisions were made to improve the clarity, relevance, and comprehensiveness of the questionnaire items. Reliability was assessed using Cronbach’s alpha coefficient to examine internal consistency. A Cronbach’s alpha value exceeding 0.70 was considered indicative of acceptable reliability for the measurement instruments.

The data analysis process was conducted in two sequential stages using the integrated FMEA-DEMATEL methodology. In the first stage, Failure Mode and Effects Analysis (FMEA) was employed to identify and prioritize the risks associated with AI deployment in administrative processes. Following the identification and validation of the risk factors, experts evaluated each risk according to the severity, occurrence, and detection criteria. For each risk, a Risk Priority Number (RPN) was calculated by multiplying the corresponding severity, occurrence, and detection scores (RPN = S × O × D). The resulting RPN values were then used to rank the risks, with higher values indicating greater criticality and priority for managerial attention. This stage enabled the identification of the most significant risks affecting the successful implementation of AI technologies within the banking environment.

In the second stage, the DEMATEL method was applied to the highest-priority risks identified through the FMEA analysis. Expert evaluations were used to construct a direct-relation matrix representing the influence of each risk on other risks. The matrix was subsequently normalized to prevent excessive numerical magnitudes and to facilitate meaningful comparison among factors. Using the normalized direct-relation matrix, the total-relation matrix was computed to capture both direct and indirect relationships among the risks.

The DEMATEL analysis further involved calculating the D and R indices, where D represents the total influence exerted by a risk on other risks and R represents the total influence received from other risks. The sum of these values (D + R) indicates the overall prominence or importance of a risk within the system, while the difference (D − R) determines whether a risk functions primarily as a causal factor or an effect factor. Positive values of D − R identify causal risks that influence other risks, whereas negative values indicate effect risks that are predominantly influenced by other factors.

The integration of FMEA and DEMATEL provided a comprehensive framework for risk assessment. Initially, FMEA facilitated the prioritization of AI-related risks based on their criticality. Subsequently, DEMATEL enabled the examination of the causal structure underlying these risks, distinguishing root causes from consequential risks. This combined approach allowed for a deeper understanding of the risk network and supported the development of more effective risk mitigation strategies by directing managerial attention toward influential root risks rather than focusing solely on observable consequences.

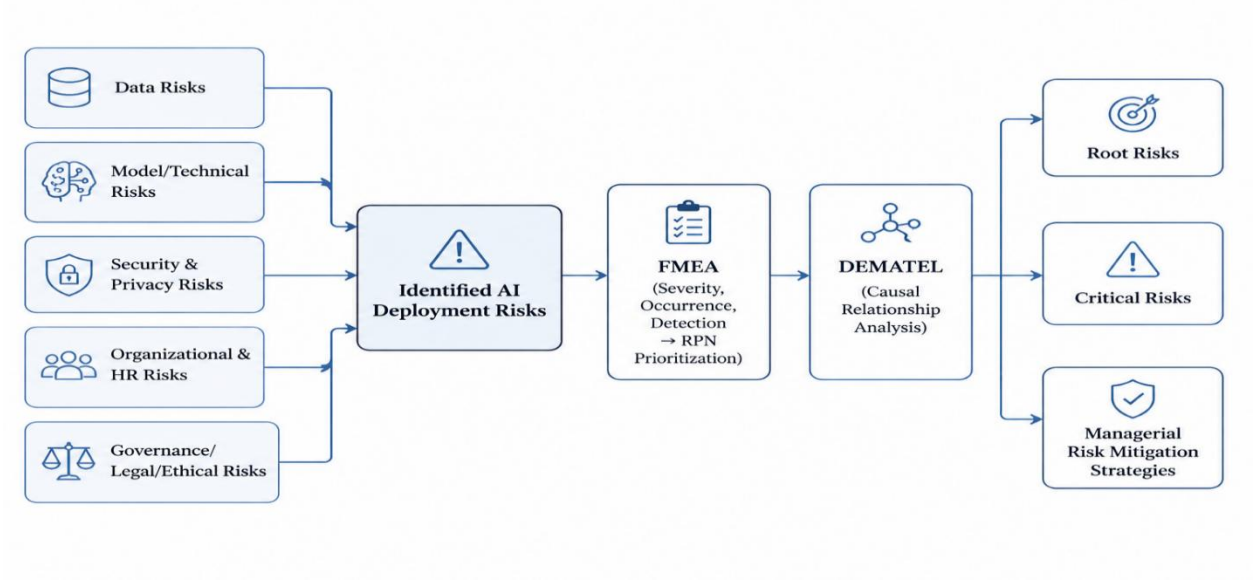

Figure 1. Conceptual Model of the Study

3. Findings and Results

After identifying and classifying the risks of artificial intelligence deployment in the administrative processes of Bank Sepah in Mazandaran Province, this section analyzes the data using the integrated FMEA-DEMATEL approach. The purpose of this stage is first to evaluate and prioritize the identified risks based on severity, occurrence probability, and detectability, and then to analyze the reciprocal relationships and causal structure among the higher-priority risks. The combined use of these two

methods enables the study not only to identify critical risks but also to determine root and influential factors, thereby providing the necessary basis for developing effective managerial strategies.

In this section, the implementation of the FMEA method, the calculation of the Risk Priority Number, and the ranking of risks are first explained. Then, based on the obtained results, priority risks are selected for causal analysis, and the relationships of influence and dependence among them are examined using the DEMATEL method. Finally, the integrated output of the two methods is presented as the basis for the final analysis of the risks of artificial intelligence deployment in the administrative processes of the bank.

Based on expert evaluations, the final score of each identified risk across different dimensions was calculated, and the Risk Priority Number was then determined. Table 1 presents the results of risk evaluation using the FMEA method.

Table 1. Results of Risk Evaluation and Prioritization Using the FMEA Method

Risk Code	Risk Title	Severity (S)	Occurrence (O)	Detection (D)	Risk Priority Number (RPN)	Rank
R1	Poor data quality	8	7	6	336	4
R2	Bias in training data	8	6	7	336	5
R3	Data inconsistency and lack of integration	7	8	6	336	6
R4	Variability in data patterns	7	6	6	252	12
R5	Lack of model transparency and interpretability	8	7	7	392	2
R6	Model error and inaccurate prediction	9	7	6	378	3
R7	Weak model generalizability	7	6	7	294	9
R8	Dependence on infrastructure or technology supplier	6	5	6	180	17
R9	Information leakage and confidentiality breach	10	7	6	420	1
R10	Cyber and adversarial attacks on the model	9	6	7	378	4
R11	Weak access control	8	6	6	288	10
R12	Weak storage and processing security	8	6	7	336	7
R13	Employee resistance to technology	7	7	5	245	13
R14	Lack of skills and training	8	7	6	336	8
R15	Weak change management	7	7	6	294	11
R16	Ambiguity in responsibility and accountability	8	5	7	280	14
R17	Non-compliance with regulations	9	5	7	315	9
R18	Absence of an artificial intelligence governance framework	8	6	7	336	6
R19	Weak internal control and monitoring	8	6	6	288	15
R20	Risk of damage to reputation and trust	9	5	7	315	10

The results of the above table show that information leakage and confidentiality breach (R9), lack of model transparency and interpretability (R5), and model error and inaccurate prediction (R6) had the highest priorities among the identified risks. This finding indicates that, in the process of deploying artificial intelligence in the administrative activities of the bank, the most important threats are related, on the one hand, to information security and the protection of sensitive data, and on the other hand, to the performance, interpretability, and accuracy of artificial intelligence models.

On the other hand, risks such as dependence on infrastructure or technology supplier, variability in data patterns, and employee resistance, although still important, were ranked lower than the other risks. This may indicate that, in the studied environment, experts are more concerned about the direct security, control, and decision-making consequences arising from the deployment of artificial intelligence.

Since the DEMATEL method is used to analyze causal relationships among factors, and its implementation for a large number of variables can increase the complexity of both response and analysis, the present study selected higher-priority risks based on the FMEA results for inclusion in this stage. Accordingly, the top ten risks with the highest RPN values were selected as key and critical risks. These risks were R9: information leakage and confidentiality breach; R5: lack of model transparency and interpretability; R6: model error and inaccurate prediction; R10: cyber and adversarial attacks on the model; R1: poor data quality; R2: bias in training data; R3: data inconsistency and lack of integration; R12: weak storage and processing security; R14: lack of skills and training; and R18: absence of an artificial intelligence governance framework.

These risks were considered the main variables in the DEMATEL stage.

At this stage, experts evaluated the intensity of the influence of each selected risk on the other risks using a five-point scale ranging from zero to four. Accordingly, the initial direct-relation matrix was formed for the selected risks. After receiving the individual matrices from the experts, the average of their opinions was considered the final direct-relation matrix. Table 2 presents the direct-relation matrix among the priority risks.

Table 2. Direct-Relation Matrix among Key Risks

Risk	R9	R5	R6	R10	R1	R2	R3	R12	R14	R18
R9	0	2	1	3	1	1	1	3	1	2
R5	1	0	3	1	2	2	1	1	1	2
R6	1	3	0	1	3	2	2	1	1	2
R10	3	1	2	0	1	1	1	3	1	2
R1	1	2	3	1	0	3	2	1	1	1
R2	1	2	2	1	3	0	2	1	1	1
R3	1	1	2	1	3	2	0	1	1	1
R12	3	1	1	3	1	1	1	0	1	2
R14	1	1	2	1	1	1	1	1	0	2
R18	2	2	2	2	2	2	2	2	2	0

To perform the DEMATEL calculations, the direct-relation matrix was first normalized. After normalization, the total-relation matrix was calculated. In this matrix, both direct and indirect effects are considered. Therefore, the total-relation matrix provides a more accurate picture of the interaction structure among the key risks.

For the final DEMATEL analysis, the sums of the rows and columns of the total-relation matrix were calculated. The row sum, denoted by D, indicates the degree to which each risk influences other risks, while the column sum, denoted by R, indicates the degree to which each risk is influenced by other risks. In addition, the D + R index indicates the importance and prominence of each risk in the overall system, whereas the D − R index determines whether the risk plays a causal or effect-based role.

Table 3. Results of DEMATEL Analysis for Key Risks

Risk Code	Risk Title	D	R	D + R	D − R	Factor Type
R9	Information leakage and confidentiality breach	5.84	6.91	12.75	-1.07	Effect
R5	Lack of model transparency and interpretability	6.42	6.10	12.52	0.32	Cause
R6	Model error and inaccurate prediction	6.88	6.45	13.33	0.43	Cause
R10	Cyber and adversarial attacks on the model	6.15	6.70	12.85	-0.55	Effect
R1	Poor data quality	6.74	6.02	12.76	0.72	Cause
R2	Bias in training data	6.37	5.89	12.26	0.48	Cause
R3	Data inconsistency and lack of integration	6.11	5.78	11.89	0.33	Cause
R12	Weak storage and processing security	6.03	6.58	12.61	-0.55	Effect
R14	Lack of skills and training	5.66	5.94	11.60	-0.28	Effect
R18	Absence of an artificial intelligence governance framework	7.12	5.95	13.07	1.17	Cause

The results of Table 3 show that the risks of absence of an artificial intelligence governance framework (R18), poor data quality (R1), bias in training data (R2), model error and inaccurate prediction (R6), and lack of model transparency and interpretability (R5) had positive D − R values. Therefore, they are classified as causal and influential factors. In contrast, information leakage and confidentiality breach (R9), cyber and adversarial attacks on the model (R10), weak storage and processing security (R12), and lack of skills and training (R14), due to their negative D − R values, are classified as effect-based and dependent factors.

From the perspective of the D + R index, model error and inaccurate prediction (R6) and the absence of an artificial intelligence governance framework (R18) had the greatest importance within the overall risk structure. This result indicates that these two risks not only have a high degree of influence but also play a central role in the network of reciprocal relationships among the risks.

The integration of the FMEA and DEMATEL results provides a more comprehensive perspective on the structure of risks associated with artificial intelligence deployment in the administrative processes of the bank. While the FMEA method identifies critical risks based on severity, occurrence probability, and detectability, the DEMATEL method determines which of these risks serve as the origins and drivers of other risks.

Based on the results of this study, although information leakage and confidentiality breach (R9) received the highest priority in the FMEA ranking, it was classified as an effect factor in the DEMATEL analysis. This finding indicates that information leakage is more a consequence of deeper weaknesses in governance, data quality, model transparency, and technical controls than a root risk in itself. In contrast, the absence of an artificial intelligence governance framework (R18) and poor data quality (R1), although they may not rank first or second in terms of Risk Priority Number, play a more fundamental role in the formation of other risks from a causal-structural perspective.

Furthermore, lack of model transparency (R5) and model error (R6) are two key risks that held high positions in both the FMEA ranking and the DEMATEL analysis. This indicates that technical and model-related risks have a strategic position in the deployment of artificial intelligence in the administrative processes of the bank, and managing them can be effective in reducing a wide range of security, operational, and reputational risks.

Overall, the integrated results of the study show that governance, data-related, and technical risks have a more root-oriented nature than other dimensions, and controlling them can reduce consequential risks such as information leakage, cyberattacks, weak processing security, and declining customer trust.

4. Discussion and Conclusion

The present study aimed to identify, prioritize, and analyze the causal structure of risks associated with the deployment of artificial intelligence (AI) in the administrative processes of Bank Sepah in Mazandaran Province using an integrated FMEA-DEMATEL approach. The findings demonstrated that AI deployment risks are not isolated phenomena but rather constitute an interconnected network of technical, data-related, organizational, security, and governance factors. The FMEA results revealed that information leakage and confidentiality breaches, lack of model transparency and interpretability, and model error and inaccurate prediction represented the highest-priority risks. Furthermore, the DEMATEL analysis indicated that the absence of an AI governance framework, poor data quality, bias in training data, model error, and lack of model transparency functioned as causal factors, whereas information leakage, cyberattacks, weak storage and processing security, and insufficient employee skills were primarily effect factors. These findings provide important insights into the nature of AI-related risks in banking environments and contribute to a more comprehensive understanding of how these risks interact and reinforce one another.

One of the most significant findings of this study was the identification of information leakage and confidentiality breaches as the highest-priority risk according to the FMEA analysis. This result is understandable given the highly sensitive nature of banking data and the severe consequences that privacy violations can have for customer trust, regulatory compliance, financial stability, and institutional reputation. AI systems rely heavily on large-scale data collection, storage, processing, and sharing mechanisms, thereby increasing exposure to unauthorized access and information security vulnerabilities. Similar concerns have been highlighted in studies examining AI implementation in healthcare and public services, where privacy protection and information security are considered among the most critical challenges associated with intelligent systems (Botha et al., 2024; Halamka et al., 2024). Research on cybersecurity frameworks further emphasizes that advanced AI environments require sophisticated security architectures capable of addressing emerging threats that traditional protection mechanisms may not adequately manage (Cunha et al., 2024; Dash, 2024). Therefore, the prominence of information leakage in the present study reflects a broader international concern regarding the protection of sensitive organizational and customer information in AI-enabled systems.

Another important finding was the high priority assigned to the lack of model transparency and interpretability. This result highlights the fact that banking organizations increasingly require AI systems that can explain the rationale behind their recommendations and decisions. Administrative processes often involve accountability, auditing, documentation, and regulatory review; therefore, black-box models may create uncertainty among employees, managers, auditors, and customers. The findings align with previous studies emphasizing that explainability and transparency are fundamental requirements for responsible AI implementation, particularly in environments characterized by high-stakes decisions and regulatory oversight (Scott et al., 2025; Wells et al., 2025). Similarly, research on healthcare AI governance has shown that the inability to explain algorithmic outputs can undermine trust and hinder organizational acceptance of AI systems (Alanazi, 2025; Arnaout, 2025). Consequently, the current findings suggest that interpretability should be viewed not merely as a technical characteristic but as a strategic organizational requirement for successful AI adoption.

The results also revealed that model error and inaccurate prediction constitute one of the most critical risks associated with AI deployment in banking administrative processes. AI systems support decision-making activities that directly influence operational performance, customer services, compliance activities, and managerial planning. When model outputs are inaccurate, the resulting decisions may lead to operational inefficiencies, resource misallocation, compliance violations, and financial losses. Previous studies have similarly noted that the reliability and validity of AI outputs are essential determinants of implementation success (Boag et al., 2024; Rajagopal et al., 2024). Large-scale deployment experiences have further demonstrated that even highly sophisticated AI systems may generate inaccurate recommendations if model assumptions, data quality, or contextual factors are not adequately considered (Mao et al., 2025). Therefore, the high ranking of model error observed in this study underscores the importance of rigorous validation, continuous monitoring, and performance auditing of AI systems throughout their operational lifecycle.

Perhaps the most important contribution of this study emerges from the DEMATEL analysis. Although information leakage achieved the highest risk priority score, it was classified as an effect factor rather than a root cause. This finding suggests that confidentiality breaches often arise as consequences of deeper structural weaknesses rather than functioning as independent risks. Specifically, weaknesses in governance, data quality, model design, and organizational controls appear to create conditions that facilitate information security failures. This result reinforces contemporary perspectives that view cybersecurity incidents as manifestations of systemic deficiencies rather than isolated technical problems (Dash, 2024; Ibokette et al., 2024). Accordingly, organizations that focus exclusively on security technologies without addressing governance and data management issues may fail to eliminate the underlying sources of vulnerability.

The absence of an AI governance framework emerged as the most influential causal factor in the entire risk network. This finding is particularly noteworthy because governance-related risks are often less visible than technical or security risks. Nevertheless, the results indicate that governance deficiencies influence a wide range of downstream risks, including model errors, compliance failures, information leakage, and accountability ambiguities. These findings strongly support previous research emphasizing the need for structured governance mechanisms, institutional oversight, policy frameworks, accountability structures, and lifecycle management practices for AI systems (Arnaout, 2025; Wells et al., 2025). Studies examining legal concerns surrounding AI implementation similarly argue that organizations require clear regulatory frameworks to manage risks effectively and ensure responsible deployment (Alanazi, 2025). The present findings therefore suggest that governance should be regarded as the foundation upon which all other AI risk management efforts are built.

Poor data quality was identified as another highly influential causal factor. This result is consistent with the fundamental principle that AI systems are only as reliable as the data on which they are trained and operated. Inaccurate, incomplete, inconsistent, or outdated data can propagate errors throughout the entire AI system, affecting model performance, decision quality, compliance outcomes, and customer experiences. Similar conclusions have been reported in studies on industrial data platforms and digital ecosystems, which emphasize that successful digital transformation depends on robust data governance, integration, and quality management practices (Chow et al., 2025; Ghosh & Laad, 2024). The significance of data quality in the current study indicates that investments in AI technology alone are insufficient unless accompanied by systematic efforts to improve organizational data infrastructure and governance.

The finding that bias in training data functions as a causal risk is equally important. Training data bias can systematically distort AI outputs, producing unfair, inaccurate, or discriminatory decisions. In banking environments, such distortions may affect customer assessments, operational prioritization, resource allocation, and risk evaluations. Previous studies have repeatedly warned that biased datasets can compromise the fairness and reliability of AI systems, thereby generating legal, ethical, and reputational consequences (Botha et al., 2024; Scott et al., 2025). The present study extends this understanding by demonstrating that training data bias not only represents a risk in itself but also contributes to the emergence of other risks within the organizational system.

Another noteworthy finding concerns the role of employee skills and training. While lack of skills and training was categorized as an effect factor rather than a causal factor, its position within the risk network remains significant. This result suggests that workforce capability gaps may partly reflect broader organizational shortcomings in governance, strategic planning, and digital transformation management. Similar observations have been made in studies examining technology adoption and organizational transformation, where employee readiness is influenced by leadership support, communication

quality, change management effectiveness, and institutional commitment to digital innovation (Patil et al., 2024; Ugwu & Balogun, 2024). Consequently, training initiatives should not be implemented in isolation but rather integrated into broader organizational transformation strategies.

The study also revealed that cyber and adversarial attacks on AI models occupy a prominent position among the priority risks. As AI systems become increasingly integrated into organizational operations, malicious actors gain new opportunities to manipulate inputs, exploit vulnerabilities, and compromise outputs. Previous cybersecurity research has demonstrated that AI-enabled systems face unique threats, including adversarial manipulation, model poisoning, and infrastructure exploitation (Cunha et al., 2024; Ibokette et al., 2024). The classification of cyberattacks as an effect factor suggests that improving governance, data quality, model design, and security controls may significantly reduce organizational exposure to such threats.

The findings further highlight the interconnected nature of governance, technical, and organizational dimensions of AI deployment. The risk network identified through DEMATEL illustrates that technological risks cannot be managed effectively without considering organizational structures, human capabilities, and institutional policies. This perspective aligns with socio-technical approaches that emphasize the interaction between technology, people, processes, and organizational environments (Cebulla, 2025; Kereopa-Yorke, 2023). Similarly, studies on public-sector digitization and digital transformation stress that successful implementation depends on balancing technological innovation with organizational adaptation and governance reform (Cati, 2024; Elete et al., 2024).

From a broader perspective, the study contributes to the growing body of literature on AI risk management by demonstrating the value of integrating prioritization and causal analysis methods. While FMEA effectively identifies the most critical risks, DEMATEL reveals the structural relationships that shape the overall risk environment. This integrated perspective enables managers to move beyond symptom-focused interventions and address the root causes that generate multiple downstream risks. Such an approach is consistent with recent discussions of AI lifecycle management, implementation frameworks, and organizational readiness models, which emphasize proactive governance, continuous monitoring, and systemic risk management (Boag et al., 2024; Rajagopal et al., 2024; Wells et al., 2025).

Overall, the results indicate that governance, data, and model-related risks constitute the foundational drivers of AI risk within banking administrative processes. Although security incidents and operational failures may attract the most immediate attention, these outcomes are frequently rooted in deficiencies related to governance structures, data management practices, model design, and organizational preparedness. Consequently, effective AI risk management requires a holistic approach that simultaneously addresses technological, organizational, human, and regulatory dimensions of implementation.

This study has several limitations that should be considered when interpreting the findings. First, the research was conducted within a single banking institution and a specific regional context, which may limit the generalizability of the results to other banks or industries. Second, the analyses relied on expert judgments, which, although appropriate for FMEA and DEMATEL methodologies, may be influenced by subjective perceptions and professional experiences. Third, the study examined risks at a particular point in time, whereas AI technologies and associated threats evolve rapidly. Finally, the causal relationships identified in the DEMATEL analysis reflect expert assessments rather than empirically observed interactions, and therefore should be interpreted as informed representations of the risk structure rather than definitive causal mechanisms.

Future studies could expand the scope of investigation by including multiple banking institutions and conducting comparative analyses across different organizational contexts. Longitudinal research designs may provide deeper insights into how AI-related risks evolve over time as technologies mature and regulatory environments change. Researchers may also integrate quantitative performance indicators with expert evaluations to validate causal relationships identified through DEMATEL. Furthermore, future studies could investigate specific categories of AI applications, such as generative AI, intelligent customer service systems, automated compliance monitoring, or predictive risk assessment tools, in order to develop more context-specific risk management frameworks. The application of alternative multi-criteria decision-making methods and hybrid analytical approaches may also enrich the understanding of AI risk structures.

Bank managers should establish comprehensive AI governance frameworks that clearly define responsibilities, accountability mechanisms, monitoring procedures, and risk management protocols before large-scale AI deployment. Organizations should invest in data quality improvement programs, including data integration, standardization, validation, and governance initiatives. Continuous model auditing, explainability assessments, and performance monitoring should be

incorporated into AI lifecycle management processes. Employee training programs should focus not only on technical competencies but also on responsible AI use, risk awareness, and decision oversight. Finally, banks should strengthen cybersecurity controls, enhance access management systems, develop incident response capabilities, and adopt proactive risk monitoring practices to ensure the secure and sustainable deployment of artificial intelligence within administrative processes.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

Acknowledgments

Authors thank all who helped us through this study.

Conflict of Interest

The authors report no conflict of interest.

Funding/Financial Support

According to the authors, this article has no financial support.

References

Alanazi, A. (2025). Assessing Clinicians' Legal Concerns and the Need for a Regulatory Framework for AI in Healthcare: A Mixed-Methods Study. *Healthcare*, *13*(13), 1487. https://doi.org/10.3390/healthcare13131487

Archer, T. A. (2025). Reckoning With Power: Why Distance Matters in the Age of Ai and Global Disruption. *Leader to Leader*, *2026*(119), 8-15. https://doi.org/10.1002/ltl.70003

Arnaout, A. (2025). Are You Leading an Artificial Intelligence-Capable Healthcare Organization? *Healthcare Management Forum*. https://doi.org/10.1177/08404704251375388

Bichel, A., ȘIȘU, J.-A., & Tîrnovanu, A. (2023). Innovative Trends Through Robotic Process Automation. A Case Study. 456-463. https://doi.org/10.24818/basiq/2023/09/046

Boag, W., Hasan, A., Kim, J. Y., Revoir, M., Nichols, M., Ratliff, W., Gao, M., Zilberstein, S., Samad, Z., Hoodbhoy, Z., Ali, M., Khan, N. S., Patel, M. R., Balu, S., & Sendak, M. (2024). The Algorithm Journey Map: A Tangible Approach to Implementing AI Solutions in Healthcare. *NPJ Digital Medicine*, *7*(1). https://doi.org/10.1038/s41746-024-01061-4

Botha, N. N., Segbedzi, C. E., Dumahasi, V. K., Maneen, S., Kodom, R. V., Tsedze, I. S., Akoto, L. A., Atsu, F. S., Lasim, O. U., & Ansah, E. W. (2024). Artificial Intelligence in Healthcare: A Scoping Review of Perceived Threats to Patient Rights and Safety. *Archives of Public Health*, *82*(1). https://doi.org/10.1186/s13690-024-01414-1

Brown, W. L., Johnson, O., & Wilson, G. (2024). The Impact of 5G Technology on Retail Marketing and Supply Chain Operations. https://doi.org/10.20944/preprints202407.2073.v1

Cati, M. M. (2024). The Digitization Process in the Italian Public Administration: Future Challenges. https://doi.org/10.5772/intechopen.1003909

Cebulla, A. (2025). AI and Workplace Relations: A WHS Framework for Managing Relational Risks in Workplaces. *Journal of Industrial Relations*, *67*(5), 743-761. https://doi.org/10.1177/00221856251392987

Chow, W., Venkataraman, N., Oh, H. C., Ramanathan, S., Sridharan, S., Arish, S. M., Wong, K.-C., Hoo, J. F., & Tan, W. (2025). Building an Artificial Intelligence and Digital Ecosystem: A Smart Hospital's Data-Driven Path to Healthcare Excellence. *Singapore medical journal*, *66*(Suppl 1), S75-S83. https://doi.org/10.4103/singaporemedj.smj-2025-066

Cunha, J., Ferreira, P., Barbero, E. M. C., Oliveira, P., Nicolau, M. J., Núñez, I., Sousa, X. R., & Serôdio, C. (2024). Enhancing Network Slicing Security: Machine Learning, Software-Defined Networking, and Network Functions Virtualization-Driven Strategies. *Future Internet*, *16*(7), 226. https://doi.org/10.3390/fi16070226

Dash, B. (2024). Zero-Trust Architecture (ZTA): Designing an AI-Powered Cloud Security Framework for LLMs' Black Box Problems. *Current Trends in Eng Sc*, *4*(2), 1-5. https://doi.org/10.54026/ctes/1058

E., R., & B., K. (2024). Resilient Supply Chains in Industry 5.0: Leveraging AI for Predictive Maintenance and Risk Mitigation. *International Journal for Multidisciplinary Research*, *6*(4). https://doi.org/10.36948/ijfmr.2024.v06i04.25116

Elete, T. Y., Nwulu, E. O., Erhueh, O. V., Akano, O. A., & Aderamo, A. T. (2024). Digital Transformation in the Oil and Gas Industry: A Comprehensive Review of Operational Efficiencies and Case Studies. *International Journal of Applied Research in Social Sciences*, *6*(11), 2611-2643. https://doi.org/10.51594/ijarss.v6i11.1692

Ghosh, A., & Laad, A. (2024). Empowering Industry 4.0 With Industrial Data Platforms: Architecture, Challenges, and Innovations. https://doi.org/10.21428/e90189c8.07732e12

Halamka, J., Kirsh, S., Liu, V. X., & Simon, L. (2024). Applications of Artificial Intelligence in Medicine: An Expert Panel Discussion. *The Permanente Journal*, *28*(3), 3-12. https://doi.org/10.7812/tpp/24.068

Ibokette, A. I., Ogundare, T. O., Anyebe, A. P., Alao, F. O., Odeh, I. I., & Okafor, F. C. (2024). Mitigating Maritime Cybersecurity Risks Using AI-Based Intrusion Detection Systems and Network Automation During Extreme Environmental Conditions. *International Journal of Scientific Research and Modern Technol*, *3*(10), 65-91. https://doi.org/10.38124/ijsrmt.v3i10.73

Kereopa-Yorke, B. (2023). Safeguarding Cognitive Well-Being in an AI-Augmented World: A Socio-Technical Systems Perspective on Neuroethics, Technostress, and Risk Management. https://doi.org/10.31234/osf.io/97fyg

Kumar, R., Kaur, A., Dangi, H. K., Kumari, P., & Kumar, N. (2025). Artificial Intelligence in Fire Safety: A Critical Perspective on Policy, Stakeholders and Emerging Technologies in India. *Fiib Business Review*, *15*(1), 11-19. https://doi.org/10.1177/23197145251342714

Mane, S., Dhote, R. R., Sinha, A., & Raja, T. (2024). Digital Twin in the Chemical Industry: A Review. *Digital Twins and Applications*, *1*(2), 118-130. https://doi.org/10.1049/dgt2.12019

Mao, K., Kapus, T., Åhs, C. T., Marescotti, M., Ip, D., Hajdu, Á., & Cela, S. (2025). WhatsCode: Large-Scale GenAI Deployment for Developer Efficiency at WhatsApp. https://doi.org/10.48550/arxiv.2512.05314

Patil, D., Rane, N. L., & Rane, J. (2024). Acceptance of ChatGPT and Generative Artificial Intelligence in Several Business Sectors: Key Factors, Challenges, and Implementation Strategies. https://doi.org/10.70593/978-81-981367-8-7_5

Rajagopal, A., Ayanian, S., Ryu, A. J., Qian, R., Legler, S. R., Peeler, E. A., Issa, M., Coons, T. J., & Kawamoto, K. (2024). Machine Learning Operations in Health Care: A Scoping Review. *Mayo Clinic Proceedings Digital Health*, *2*(3), 421-437. https://doi.org/10.1016/j.mcpdig.2024.06.009

Scaramuzza, F. (2025). "Show Me You Comply... Without Showing Me Anything": Zero-Knowledge Software Auditing for AI-Enabled Systems. https://doi.org/10.48550/arxiv.2510.26576

Scott, I., Reddy, S., Miller, T., & Vegt, A. v. d. (2025). Using Generative Artificial Intelligence in Clinical Practice: A Narrative Review and Proposed Agenda for Implementation. *The Medical Journal of Australia*, *223*(11), 664-672. https://doi.org/10.5694/mja2.70057

Ugwu, K. E., & Balogun, O. S. (2024). Navigating Supply Chain Management: Technology Adoption in Southeast Nigerian Breweries. *Journal of Commerce Management and Tourism Studies*, *3*(3), 231-243. https://doi.org/10.58881/jcmts.v3i3.231

Wells, B. J., Nguyen, H., McWilliams, A., Pallini, M., Bovi, A., Kramer, J., Chou, S. H., Hetherington, T., Corn, P., Taylor, Y. J., Cuison, A., Gagen, M., Isreal, M., Akbilgiç, O., Barr, K., Caffrey, R. G., Carroll, M. S., CiRullo, M., Downs, S. M., . . . Setliff, E. (2025). A Practical Framework for Appropriate Implementation and Review of Artificial Intelligence (FAIR-AI) in Healthcare. *NPJ Digital Medicine*, *8*(1). https://doi.org/10.1038/s41746-025-01900-y