Perbanas International Conference on Economics, Business, Management, Accounting and IT (PROFICIENT) 2026
"Integrity-Driven Innovation: Reimagining Business, Finance, and Digital Transformation for Sustainable Impact"

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

Responsible AI in Digital Banking: Balancing Innovation, Cybersecurity, and Ethical Governance

Kesi Hapsari[1,a)], Mariayana Jemita[2,b)], Tiolina Evi Nausta Pardede[3,c)]

[1,2,3]*Department of Economics and Business, Perbanas Institute, Jakarta, Indonesia*

[a)]Corresponding author: kesi.hapsari04@perbanas.id
[b)]mariayana.jemita77@perbanas.id
[c)]tiolina@perbanas.id

Abstract. *Artificial intelligence (AI) is increasingly embedded in digital banking through fraud detection, credit scoring, customer service automation, anti-money laundering monitoring, cyber threat analytics, and personalised financial services. Although these applications improve efficiency and service innovation, they also create risks related to model opacity, algorithmic bias, privacy leakage, third-party dependency, adversarial cyberattacks, inaccurate generative outputs, and unclear accountability. This article develops a systematic empirical framework for examining responsible AI implementation in digital banking by balancing AI-driven innovation capability, cybersecurity readiness, and ethical AI governance. The proposed model is grounded in the Technology-Organization-Environment framework, dynamic capability theory, socio-technical systems theory, and trustworthy AI principles. A quantitative explanatory design is recommended, using survey data from banking professionals involved in digital banking, information technology, cybersecurity, risk management, compliance, data analytics, and customer experience. The model can be tested through partial least squares structural equation modelling or moderated regression analysis. The article contributes by positioning responsible AI as an organisational governance capability rather than merely a technical compliance issue. It argues that responsible AI requires innovation routines that are supported by secure infrastructure, explainable model management, human oversight, fairness controls, privacy protection, audit trails, and board-level accountability.*

Keywords: *responsible AI, digital banking, cybersecurity readiness, ethical governance, financial innovation*

INTRODUCTION

Digital banking has changed the structure of financial services by turning banks into data-intensive platforms. Mobile banking, internet banking, e-KYC, open banking, digital lending, cloud infrastructure, application programming interfaces, real-time payment systems, and automated customer service have expanded how customers access financial products and how banks organise their internal operating models ((Basel Committee on Banking Supervision, 2024; (Financial Stability Board, 2024; Otoritas Jasa Keuangan, 2025). In this environment, artificial intelligence is no longer an experimental technology. Banks increasingly use AI to classify transaction risks, detect fraud patterns, evaluate creditworthiness, recommend products, automate service interactions, screen suspicious transactions, predict customer churn, and support internal productivity. AI therefore promises faster decisions, lower operational costs, better risk analytics, and more personalised customer experiences.

However, the adoption of AI in banking also creates a governance problem. Banking is a high-trust and highly regulated industry because AI decisions may affect customers, operational resilience, consumer protection, and systemic confidence (Crisanto et al., 2024; (Financial Stability Board, 2024; (National Institute of Standards and Technology, 2023). Digital decisions can influence customer rights, data security, access to finance, market confidence, and financial system stability. AI systems that are inaccurate, biased, opaque, insecure, or poorly supervised can harm customers and weaken institutional credibility. For example, an AI credit model may reproduce historical discrimination; a fraud detection model may wrongly block legitimate transactions; a chatbot may provide misleading financial information; a generative AI tool may leak confidential data; and an outsourced AI platform may

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

increase third-party concentration risk. These risks show that AI innovation must be balanced with cybersecurity and ethical governance.

The regulatory context reinforces the urgency of this issue. OJK launched Artificial Intelligence Governance for Indonesian Banks on April 29, 2025 as a minimum benchmark for responsible AI development and implementation in Indonesian banks (Otoritas Jasa Keuangan, 2025). At the international level, the Financial Stability Board identifies AI-related vulnerabilities such as cyber risk, model governance challenges, third-party dependency, and concentration risk in the financial sector (Financial Stability Board, 2024, 2025). NIST positions AI risk management as a practical approach for incorporating validity, reliability, safety, security, resilience, accountability, transparency, explainability, privacy, and fairness into AI governance (National Institute of Standards and Technology, 2023, 2024).

Despite these developments, existing research remains fragmented. Some studies examine AI acceptance from the customer perspective, such as the Indonesian banking study by (Ikhsan et al., 2025), while others review AI ethics, consumer risk, and regulatory perspectives in banking (Fundira & Mbohwa, 2025). Fewer studies integrate innovation capability, cyber resilience, and ethical governance into one empirical model for responsible AI implementation in digital banking. This gap matters because a bank may be innovative but insecure, secure but ethically opaque, or ethically principled but weak in implementation. Responsible AI should therefore be studied as an organisational capability that coordinates technology, governance, human oversight, and regulatory alignment.

This article aims to develop a detailed empirical framework for analysing responsible AI in digital banking. The research questions are: (1) how does AI-driven innovation capability influence responsible AI implementation; (2) how does cybersecurity readiness influence responsible AI implementation; (3) how does ethical AI governance influence responsible AI implementation; and (4) how can ethical governance strengthen the relationship between innovation, cybersecurity, and responsible AI outcomes. The originality and novelty of this research lie in its integrated positioning of responsible AI as a measurable organisational capability that simultaneously combines AI-driven innovation capability, cybersecurity readiness, and ethical AI governance in a single empirical model. Unlike prior studies that mainly examine customer acceptance of AI, general AI ethics, or isolated cybersecurity risks, this article links responsible AI implementation to banking-specific governance controls such as model inventory, explainability review, fairness testing, privacy protection, human oversight, audit trail, vendor risk control, incident response, and board-level accountability. This contribution provides a more operational framework for scholars, bank managers, risk officers, compliance units, and regulators in designing safe, fair, transparent, and accountable AI-based banking services.

LITERATURE REVIEW

Responsible AI in Digital Banking

Responsible AI refers to the development and use of AI systems in ways that create beneficial outcomes while preventing unreasonable harm. In banking, responsible AI should follow trustworthy AI principles, risk-based governance, and an artificial intelligence management system approach (National Institute of Standards and Technology, 2023; OECD, 2024; (Otoritas Jasa Keuangan, 2025). The concept must be interpreted through prudential, technological, and ethical lenses. A responsible AI system should be secure, reliable, explainable, privacy-preserving, fair, auditable, and subject to accountable human oversight.

Responsible AI is not limited to a written code of ethics. It must be operationalised through the AI life cycle: problem definition, data collection, feature engineering, model selection, model validation, deployment, monitoring, incident response, customer communication, and periodic review. In credit scoring, responsible AI requires bias testing, explainable adverse-action logic, human review, and customer recourse. In fraud detection, it requires cyber-resilient architecture, false-positive management, and clear escalation protocols. In chatbots, it requires content guardrails, disclosure, data protection, and handover to human agents. In AML monitoring, it requires model validation and audit trails so compliance decisions can be reviewed.

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

The literature on AI ethics in banking increasingly highlights transparency, fairness, privacy, trust, and regulatory oversight as central concerns (Floridi et al., 2018; (Siau & Wang, 2020). (Fundira & Mbohwa, 2025) show that AI ethics research in banking is moving toward integration between regulatory perspectives, customer experience, and technology governance. This implies that responsible AI cannot be evaluated only by whether a model is technically accurate; it must also be evaluated by whether the model can be justified, challenged, audited, and corrected when it affects customers.

TABLE 1. AI use-case risk taxonomy in digital banking

AI use case	Main innovation value	Primary risk	Responsible AI control
Credit scoring and digital lending	Faster credit decisions and alternative-data assessment	Bias, explainability gaps, unfair exclusion, privacy risk	Bias testing, adverse-decision explanation, human review, documented model validation
Fraud detection and transaction monitoring	Real-time anomaly detection and lower fraud losses	False positives, adversarial attacks, customer disruption	Threshold review, cyber monitoring, appeal mechanism, incident response testing
AI chatbot and virtual assistant	24/7 service automation and personalised guidance	Hallucination, misleading advice, data leakage, weak escalation	Disclosure, answer guardrails, human handover, confidential-data restriction
AML and compliance analytics	Improved suspicious-transaction prioritisation	Over-flagging, under-detection, opaque compliance logic	Audit trail, explainable alert logic, independent validation, periodic monitoring
Generative AI productivity tools	Faster document drafting, summarisation, and internal support	Confidentiality breach, inaccurate output, excessive reliance	Acceptable-use policy, human verification, secure prompts, access limitation

AI-Driven Innovation Capability

AI-driven innovation capability is the bank's capacity to identify, design, integrate, scale, and improve AI-based services or processes. This capability can be interpreted as a dynamic capability because it enables banks to sense digital opportunities, seize AI-enabled service innovations, and reconfigure organisational routines in response to technological change (Teece, 2007). In the Indonesian banking context, perceived usefulness, security, and trust remain important for AI adoption, which means innovation capability must be supported by credible governance and user confidence (Ikhsan et al., 2025).

Nevertheless, innovation capability alone does not guarantee responsibility. A bank may deploy many AI tools but still fail if the systems are not validated, monitored, explained, secured, or governed. Responsible innovation requires security-by-design, privacy-by-design, fairness-by-design, and governance-by-design. Therefore, this study treats AI-driven innovation capability as a driver of responsible AI only when innovation processes are embedded in controlled and accountable organisational routines.

Cybersecurity Readiness

Cybersecurity readiness is the organisational ability to prevent, detect, respond to, and recover from cyber threats affecting AI-enabled digital banking. AI creates new security challenges because models depend on large datasets, complex infrastructures, third-party platforms, cloud services, APIs, and automated decision pipelines (Basel Committee on Banking Supervision, 2024); (Financial Stability Board, 2024). Generative AI also adds risks such as prompt injection, data leakage, hallucinated outputs, insecure plugins, malicious automation, and difficulty in tracing model behaviour (National Institute of Standards and Technology, 2024).

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

The Financial Stability Board identifies cyber risk, third-party dependency, model governance challenges, concentration risk, and market correlation as important vulnerabilities associated with AI adoption in finance (Financial Stability Board, 2024, 2025). This is particularly relevant for banks because system failure, cyber incidents, and uncontrolled vendor dependence can affect not only individual customers but also operational resilience and financial stability. Therefore, cybersecurity readiness should be embedded into the AI life cycle, not treated as a separate technical function.

Ethical AI Governance

Ethical AI governance refers to formal and informal mechanisms that guide AI decisions according to fairness, explainability, accountability, privacy, transparency, safety, and human-centred values. In banking, ethical governance includes board oversight, AI policy, model inventory, risk classification, model validation, fairness testing, explainability standards, audit trails, customer complaint channels, and escalation rules (International Organization for Standardization, 2023; (Morley et al., 2021; OECD, 2024; (Otoritas Jasa Keuangan, 2025).

Ethical AI governance can also operate as a moderating capability. Innovation capability may increase the number and sophistication of AI applications, but ethical governance determines whether innovation is channelled toward customer protection and accountable decision-making. Cybersecurity readiness may strengthen technical protection, but ethical governance ensures that security practices remain proportional, transparent, and respectful of privacy. Therefore, ethical governance is expected to strengthen the positive influence of innovation and cybersecurity on responsible AI implementation.

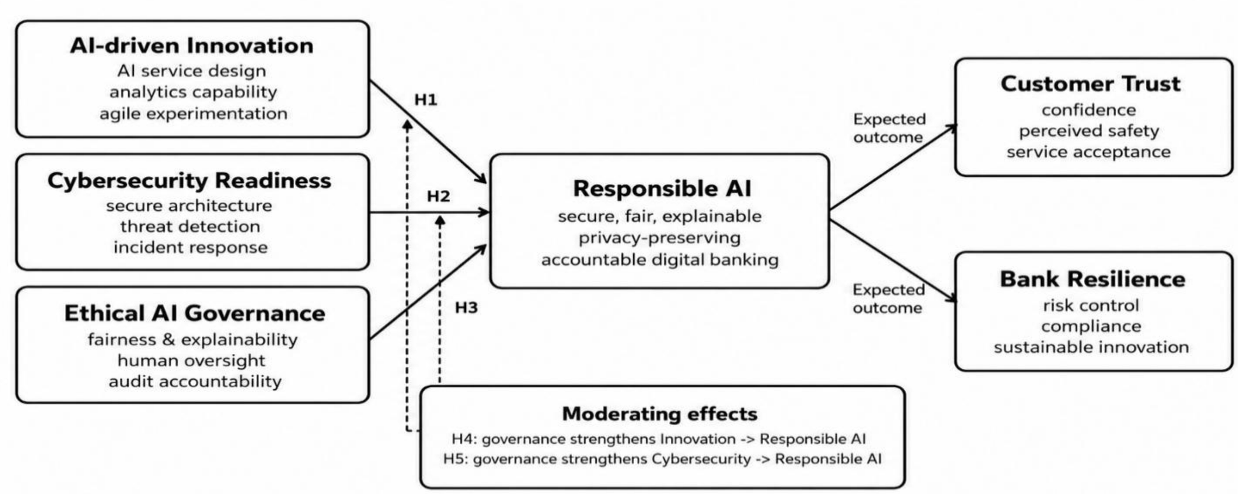

FIGURE 1. Conceptual framework of Responsible AI in digital banking

FIGURE 1. Conceptual framework of responsible AI in digital banking

TABLE 2. Hypothesis development

Code	Hypothesis	Theoretical rationale
H1	AI-driven innovation capability has a positive effect on responsible AI implementation in digital banking.	Banks with stronger AI innovation routines are more able to design, test, and scale responsible AI services.
H2	Cybersecurity readiness has a positive effect on responsible AI implementation in digital banking.	Secure architecture and incident response reduce operational and cyber risks in AI-enabled services.

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

H3	Ethical AI governance has a positive effect on responsible AI implementation in digital banking.	Policies, oversight, fairness testing, explainability, and accountability convert AI principles into practice.
H4	Ethical AI governance strengthens the effect of innovation capability on responsible AI implementation.	Innovation becomes more responsible when embedded in human oversight, audit, and customer protection.
H5	Ethical AI governance strengthens the effect of cybersecurity readiness on responsible AI implementation.	Cybersecurity controls become more trustworthy when aligned with privacy, proportionality, and transparency.

METHODS

This article proposes a quantitative explanatory research design because the objective is to test causal relationships among organisational capabilities and responsible AI implementation. This design is appropriate for examining latent organisational constructs and estimating the strength of direct and moderating effects through survey-based modelling (Hair et al., 2019). The unit of analysis is the bank or banking work unit, while the unit of observation is the professional involved in digital banking, information technology, cybersecurity, risk management, compliance, data analytics, internal audit, product development, or customer experience. The recommended context is commercial banks that have adopted digital channels and are beginning to deploy AI or advanced analytics in operational or customer-facing processes. This scope is consistent with recent banking guidance that treats AI governance, cybersecurity resilience, model risk, third-party risk, and customer protection as interconnected issues in digital finance (Financial Stability Board, 2024; (National Institute of Standards and Technology, 2023; (Otoritas Jasa Keuangan, 2025).

The population consists of banking professionals with direct or indirect knowledge of AI-enabled digital banking. A purposive sampling technique is appropriate because respondents should understand digital transformation, risk controls, or AI governance practices. Respondents may include digital banking managers, IT security officers, AI or data analytics staff, compliance officers, operational risk officers, internal auditors, product managers, and customer experience personnel. The minimum sample can follow a statistical power approach or the ten-times rule in PLS-SEM, where the sample size should exceed ten times the largest number of structural paths pointing at a dependent construct (Hair et al., 2019). For stronger empirical validity, a target of 150-300 respondents is recommended, depending on access to banking institutions. Data can be collected through a structured questionnaire using a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree, because this scale allows respondents to express the degree to which AI innovation, cybersecurity readiness, ethical governance, and responsible AI practices are implemented in their banking units.

The instrument should be adapted from prior studies and standards on AI adoption, cybersecurity readiness, digital transformation capability, technology governance, and responsible AI principles. The items for AI-driven innovation capability should capture experimentation, analytics infrastructure, cross-functional collaboration, AI service design, and model deployment routines. The cybersecurity readiness items should capture secure architecture, access control, anomaly detection, API protection, vendor security review, incident response, and recovery capability. The ethical AI governance items should capture AI policy, model inventory, fairness testing, explainability, privacy protection, human oversight, auditability, board accountability, and customer recourse, in line with trustworthy AI and AI management system principles (Morley et al., 2021; (OECD, 2024). Before main data collection, expert review should be conducted with digital banking practitioners, IT risk officers, compliance specialists, or academics to assess content validity. A pilot test with 20-30 respondents is recommended to examine item clarity, reliability, and completion time. Ethical procedures should include informed consent, anonymous responses, voluntary participation, and assurance that no confidential bank data are collected.

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

Data analysis can be performed using Partial Least Squares Structural Equation Modelling (PLS-SEM) because the model is prediction-oriented, includes latent constructs, and contains interaction effects between ethical AI governance and the two predictor variables (Hair et al., 2019). The technical analysis should proceed in two stages. First, the measurement model is evaluated using indicator loadings, Cronbach's alpha, composite reliability, average variance extracted, and discriminant validity through the heterotrait-monotrait ratio (Henseler et al., 2015). Indicators with outer loadings below the acceptable threshold should be reviewed carefully and removed only when deletion improves construct reliability and does not damage theoretical meaning. Second, the structural model is evaluated using collinearity statistics, path coefficients, coefficient of determination, effect size, predictive relevance, and bootstrapped significance levels. Bootstrapping with at least 5,000 resamples is recommended to obtain t-statistics, p-values, and confidence intervals for H1-H5. The moderating effects in H4 and H5 can be tested using a two-stage or product-indicator approach by creating interaction terms between ethical AI governance and AI-driven innovation capability, and between ethical AI governance and cybersecurity readiness.

To reduce common method bias, the questionnaire should use clear wording, mixed item order, psychological separation between predictors and outcomes, and anonymity. Statistical checks such as full collinearity VIF, Harman's single-factor test, or marker variable procedures may be used as robustness tests. The full collinearity approach is useful in PLS-SEM because it detects potential common method bias by examining whether latent variable VIF values exceed acceptable thresholds (Kock, 2015). The final interpretation of the data should not rely only on statistical significance. It should also consider the size and direction of the coefficients, theoretical consistency, predictive relevance, and whether the findings support responsible AI implementation as a governance capability rather than merely a technological adoption outcome.

TABLE 3. Operational definition of research variables

Variable	Operational definition	Main indicators
AI-driven innovation capability (X1)	Ability to design, integrate, and scale AI-based banking innovation	AI experimentation, data infrastructure, analytics talent, cross-functional collaboration, service personalisation, model deployment routines
Cybersecurity readiness (X2)	Ability to protect AI-enabled digital banking from cyber threats	secure architecture, access control, anomaly detection, API protection, vendor security review, incident response, recovery capability
Ethical AI governance (X3)	Formal and practical mechanisms for fair, transparent, accountable, and human-supervised AI	AI policy, model inventory, fairness testing, explainability, human oversight, audit trail, board accountability, customer recourse
Responsible AI implementation (Y)	Extent to which AI systems are deployed safely, fairly, transparently, and accountably in digital banking	privacy protection, reliability, explainability, fairness, cyber resilience, human-in-the-loop, compliance alignment, continuous monitoring

TABLE 4. Suggested questionnaire blueprint

Code	Measurement item
X1.1	Our bank uses AI to improve digital banking services and operational processes.
X1.2	AI initiatives are supported by adequate data infrastructure and analytics capability.
X1.3	Business, risk, IT, and compliance units collaborate in AI-based innovation projects.
X2.1	AI-enabled digital banking systems are protected by strong access control and monitoring.
X2.2	The bank has procedures to detect and respond to AI-related cyber incidents.

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

X2.3	Third-party AI or cloud providers are evaluated based on security and resilience criteria.
X3.1	The bank has a formal policy or guideline for responsible AI use.
X3.2	High-risk AI models are reviewed for fairness, explainability, and customer impact.
X3.3	There is clear accountability for AI decisions, model monitoring, and incident escalation.
Y1	AI systems used in digital banking are explainable, auditable, and monitored after deployment.
Y2	The bank provides human oversight for AI decisions that may affect customers.
Y3	AI implementation strengthens customer trust while maintaining privacy, security, and compliance.

RESULTS AND DISCUSSION

Expected Measurement Model Evaluation

Because this article is prepared as a journal-ready empirical framework, this section explains how the results should be interpreted after field data are collected. In the measurement model, each construct should demonstrate adequate reliability and validity based on the PLS-SEM reporting standards proposed by (Hair et al., 2019)Hair et al. (2019). Indicator loadings above 0.70 indicate that items represent their intended constructs, while composite reliability and Cronbach's alpha above 0.70 indicate internal consistency. Average variance extracted above 0.50 supports convergent validity, and HTMT values below 0.85 or 0.90 support discriminant validity (Henseler et al., 2015). The statistical interpretation should be reported in a table that includes outer loadings, reliability values, AVE, HTMT ratios, VIF values, R-square, f-square, Q-square, path coefficients, t-statistics, p-values, and confidence intervals. These values allow the researcher to explain whether the empirical data support the proposed theoretical model, rather than only presenting descriptive results.

A strong measurement model would indicate that responsible AI can be measured through multiple dimensions: security, fairness, explainability, privacy, accountability, human oversight, and monitoring. This is important because responsible AI is not a single technical feature. It is a multidimensional organisational capability. If ethical AI governance indicators show strong loadings, it means that respondents perceive AI policy, fairness testing, audit trail, model inventory, and accountability as coherent aspects of governance. If cybersecurity indicators show strong loadings, it means that AI risk management is not separated from cyber resilience. This interpretation is consistent with NIST's AI risk management framework, which emphasises valid, reliable, safe, secure, accountable, transparent, explainable, privacy-enhancing, and fair AI systems (National Institute of Standards and Technology, 2023). It is also consistent with OJK's AI governance direction for Indonesian banks, which places AI accountability, human oversight, and risk control within banking governance rather than treating AI as only an IT project (Otoritas Jasa Keuangan, 2025).

Expected Structural Model Interpretation

If H1 is supported, AI-driven innovation capability positively influences responsible AI implementation. This would suggest that banks with mature data infrastructure, AI talent, experimentation routines, and cross-functional collaboration are more likely to implement AI responsibly. The interpretation should be linked to the size of the path coefficient and its significance level; a positive and significant coefficient would show that innovation capability contributes to responsible AI, while the f-square value would indicate whether the contribution is small, medium, or large. However, the interpretation should avoid assuming that innovation automatically produces responsibility. The managerial implication is that innovation teams must incorporate responsible design practices from the beginning of AI projects, not only after system deployment. This finding would be aligned with dynamic capability theory, which argues that organisations need sensing, seizing, and reconfiguring capabilities to transform technological opportunities into sustainable performance (Teece, 2007).

If H2 is supported, cybersecurity readiness positively influences responsible AI implementation. This would confirm that secure infrastructure, access control, threat monitoring, incident response, and vendor security review are essential for trustworthy AI-based banking. AI systems are only responsible when they are resilient against cyber

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

manipulation, data leakage, and operational disruption. For digital banking, cybersecurity is therefore not merely a back-office control but a core component of customer protection and institutional trust. The result should be interpreted by connecting the statistical evidence to recent financial-sector risk reports showing that AI adoption may intensify cyber risk, third-party concentration risk, model governance challenges, and operational resilience concerns (Basel Committee on Banking Supervision, 2024); (Financial Stability Board, 2024, 2025)

If H3 is supported, ethical AI governance positively influences responsible AI implementation. This would indicate that formal policy, board oversight, fairness evaluation, explainability, audit trails, and human accountability are central to responsible AI. The finding would support the argument that banks cannot rely only on technical model performance. They need governance mechanisms that decide which AI applications are acceptable, how high-risk AI is reviewed, who is accountable for outcomes, and how customers can challenge or understand automated decisions. A strong and significant H3 path would also support the idea that AI ethics must be operationalised through concrete organisational controls, as argued in prior studies on AI ethics implementation and trustworthy AI governance (Floridi et al., 2018); Morley et al., 2021); OECD, 2024)

If H4 and H5 are supported, ethical AI governance strengthens the relationship between both innovation capability and cybersecurity readiness toward responsible AI. This would show that governance operates as an enabling condition. The interaction effect should be interpreted through the sign, significance, and plotted slope of the interaction term. A positive interaction would mean that innovation capability and cybersecurity readiness generate stronger responsible AI outcomes when ethical AI governance is high. Innovation becomes safer when guided by ethical principles, and cybersecurity becomes more trustworthy when linked to privacy, fairness, proportionality, and accountability. In practical terms, banks should not separate AI innovation committees, cyber risk teams, and ethics or compliance functions. These units should work as an integrated responsible AI governance architecture.

Managerial and Policy Discussion

The proposed framework has several managerial implications. First, banks should develop an AI inventory that classifies AI systems according to risk level, business function, data sensitivity, customer impact, and vendor dependency. Second, high-risk AI systems such as credit scoring, fraud blocking, investment recommendation, and AML prioritisation should undergo enhanced validation, explainability review, fairness testing, cybersecurity assessment, and periodic audit. Third, generative AI tools should be governed through acceptable-use rules, data leakage prevention, prompt and output monitoring, human verification, and restrictions on confidential information processing.

Fourth, cybersecurity readiness should be integrated into the AI life cycle. Cybersecurity teams need to participate in model development, API integration, cloud deployment, third-party review, and incident simulation. Fifth, ethical governance should be made measurable. Banks should define key risk indicators and key control indicators for fairness, explainability, customer complaints, model drift, false positives, privacy incidents, and unresolved AI exceptions. Sixth, board and senior management oversight should be strengthened because responsible AI involves strategic trade-offs between innovation speed, operational risk, customer trust, and regulatory compliance.

From a policy perspective, regulators can use the framework to encourage banks to report AI adoption in a structured manner. Supervisory assessment should not only ask whether banks use AI, but also which use cases are high risk, how models are validated, how customer harms are prevented, how third-party AI vendors are controlled, and how cyber incidents involving AI are managed. This approach is consistent with the need for risk-based, proportionate, and technology-neutral AI governance. It also supports the view that responsible AI in banking must protect financial innovation without weakening prudential standards and consumer protection.

TABLE 5. Responsible AI implementation roadmap for digital banking

Stage	Key action	Expected outcome
1. Governance foundation	Establish AI policy, model inventory, roles, accountability, and risk classification.	Clear ownership for AI risk and decision accountability.

2. Secure-by-design development	Embed access control, secure data pipelines, vendor security review, and API protection.	Lower exposure to cyber incidents and third-party failures.
3. Ethical validation	Test fairness, explainability, privacy, and customer impact before deployment.	Reduced bias, opacity, and consumer protection risk.
4. Controlled deployment	Use human-in-the-loop controls, escalation procedures, and documented approval gates.	AI decisions remain supervised and auditable.
5. Continuous monitoring	Track model drift, cyber anomalies, false positives, complaints, and incident response.	Responsible AI becomes a continuous governance process.

CONCLUSIONS

This article develops a detailed empirical framework for responsible AI in digital banking by integrating AI-driven innovation capability, cybersecurity readiness, and ethical AI governance. The central argument is that banks cannot treat AI as only a technological innovation. AI in digital banking directly affects trust, security, privacy, fairness, and regulatory compliance. Therefore, responsible AI should be understood as an organisational capability that aligns innovation speed with cyber resilience, ethical accountability, and customer protection.

The proposed framework contributes to the literature by connecting technology capability, risk management, and ethical governance in one testable model. It also provides practical indicators for measuring responsible AI implementation through secure architecture, explainability, fairness controls, privacy protection, human oversight, auditability, and continuous monitoring. For bank managers, the study highlights the need to integrate innovation teams, cybersecurity units, compliance functions, risk management, internal audit, and board oversight. For regulators, the framework supports risk-based supervision of AI use cases, third-party dependency, model governance, and AI-related cyber resilience.

Future research should test the model using survey data from multiple banks and compare results across bank size, digital maturity, ownership type, and regulatory environment. Qualitative interviews may also be added to capture deeper insights from risk officers, cybersecurity teams, data scientists, and compliance managers. Longitudinal research would be valuable to examine how responsible AI governance evolves as banks move from pilot projects to large-scale AI deployment.

Acknowledgments

The authors acknowledge academic and professional discussions on digital banking governance, cybersecurity risk, and responsible AI implementation. No external funding was received for the preparation of this article.

References

Basel Committee on Banking Supervision. (2024). *Digitalisation of finance*. Bank for International Settlements. https://www.bis.org/bcbs/publ/d575.pdf

Crisanto, J. C., Leuterio, C. B., Prenio, J., & Yong, J. (2024). *Regulating AI in the Financial Sector: Recent Developments and Main Challenges* (Number FSI Insights on Policy Implementation No. 63). Bank for International Settlements. https://www.bis.org/fsi/publ/insights63.htm

Financial Stability Board. (2024). *The Financial Stability Implications of Artificial Intelligence*. Financial Stability Board. https://www.fsb.org/uploads/P14112024.pdf

Financial Stability Board. (2025). *Monitoring Adoption of Artificial Intelligence and Related Vulnerabilities in the Financial Sector*. Financial Stability Board. https://www.fsb.org/2025/10/monitoring-adoption-of-artificial-intelligence-and-related-vulnerabilities-in-the-financial-sector/

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & 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

Fundira, M. A., & Mbohwa, C. (2025). AI ethics in banking services: A systematic and bibliometric review of regulatory and consumer perspectives. *Discover Artificial Intelligence*, *5*, 319. https://doi.org/10.1007/s44163-025-00432-4

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Ikhsan, R. B., Fernando, Y., Prabowo, H., Yuniarty, Gui, A., & Kuncoro, E. A. (2025). An empirical study on the use of artificial intelligence in the banking sector of Indonesia by extending the TAM model and the moderating effect of perceived trust. *Digital Business*, *5*(1), 100103. https://doi.org/10.1016/j.digbus.2024.100103

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of E-Collaboration*, *11*(4), 1–10. https://doi.org/10.4018/ijec.2015100101

Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mokander, J., & Floridi, L. (2021). Ethics as a service: A pragmatic operationalisation of AI ethics. *Minds and Machines*, *31*(2), 239–256. https://doi.org/10.1007/s11023-021-09563-w

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (Number NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

National Institute of Standards and Technology. (2024). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile* (Number NIST AI 600-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.600-1

OECD. (2024). *OECD artificial intelligence principles*. Organisation for Economic Co-operation and Development. https://oecd.ai/en/ai-principles

Otoritas Jasa Keuangan. (2025). *Artificial intelligence governance for Indonesian banks*. Otoritas Jasa Keuangan. https://ojk.go.id/en/Publikasi/Roadmap-dan-Pedoman/Perbankan/Pages/Indonesia-Artificial-Intelligence-Governance-for-Banking.aspx

Siau, K., & Wang, W. (2020). Artificial intelligence ethics: Ethics of AI and ethical AI. *Journal of Database Management*, *31*(2), 74–87. https://doi.org/10.4018/JDM.2020040105

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of sustainable enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350. https://doi.org/10.1002/smj.640